RESEARCH ARTICLE

Fossil-calibrated molecular clock data enable reconstruction of steps leading to differentiated multicellularity and anisogamy in the Volvocine algae

Charles Ross Lindsey[1], Andrew H. Knoll[2], Matthew D. Herron[1] and Frank Rosenzweig[1,3]*

Abstract

Background Throughout its nearly four-billion-year history, life has undergone evolutionary transitions in which simpler subunits have become integrated to form a more complex whole. Many of these transitions opened the door to innovations that resulted in increased biodiversity and/or organismal efficiency. The evolution of multicellularity from unicellular forms represents one such transition, one that paved the way for cellular differentiation, including differentiation of male and female gametes. A useful model for studying the evolution of multicellularity and cellular differentiation is the volvocine algae, a clade of freshwater green algae whose members range from unicellular to colonial, from undifferentiated to completely differentiated, and whose gamete types can be isogamous, anisogamous, or oogamous. To better understand how multicellularity, differentiation, and gametes evolved in this group, we used comparative genomics and fossil data to establish a geologically calibrated roadmap of when these innovations occurred.

Results Our ancestral-state reconstructions, show that multicellularity arose independently twice in the volvocine algae. Our chronograms indicate multicellularity evolved during the Carboniferous-Triassic periods in Goniaceae + Volvocaceae, and possibly as early as the Cretaceous in Tetrabaenaceae. Using divergence time estimates we inferred when, and in what order, specific developmental changes occurred that led to differentiated multicellularity and oogamy. We find that in the volvocine algae the temporal sequence of developmental changes leading to differentiated multicellularity is much as proposed by David Kirk, and that multicellularity is correlated with the acquisition of anisogamy and oogamy. Lastly, morphological, molecular, and divergence time data suggest the possibility of cryptic species in Tetrabaenaceae.

Conclusions Large molecular datasets and robust phylogenetic methods are bringing the evolutionary history of the volvocine algae more sharply into focus. Mounting evidence suggests that extant species in this group are the result of two independent origins of multicellularity and multiple independent origins of cell differentiation. Also, the origin of the Tetrabaenaceae-Goniaceae-Volvocaceae clade may be much older than previously thought. Finally, the possibility of cryptic species in the Tetrabaenaceae provides an exciting opportunity to study the recent divergence of lineages adapted to live in very different thermal environments.

*Correspondence:
Frank Rosenzweig
frank.rosenzweig@biology.gatech.edu
Full list of author information is available at the end of the article

Background

Major evolutionary transitions occur when multiple autonomous units (e.g., genes) combine to form an interdependent autonomous unit (e.g., chromosomes) capable of storing and transmitting information in a novel way [1]. Over the past four billion years a relatively small number of such transitions have resulted in a myriad of innovations that have contributed to the diversification of life on Earth [1] [2]. Among the most conspicuous of these is the transition from organisms whose individuals consist of one cell to individuals that consist of many.

Multicellularity has independently evolved from a unicellular ancestor at least 45 times [3], and has been documented in all three domains of life: Bacteria[4], Archaea [5], and Eukarya [6]. In most cases, this transition opened the door to evolution of division of labor, which in turn paved the way for cellular differentiation, where cells in a multicellular body take on specific tasks.

Task specialization has the potential to boost a multicellular organism's fitness, provided that the programmed partitioning of functions provides metabolic, structural, and/or genetic advantages over retaining all functions in all cells. Cellular differentiation occurs in most but not all multicellular eukaryotes.

Though multicellularity and differentiation have repeatedly evolved, we have limited knowledge of what selects for these features and the genetic steps that enable their initial evolution [7]. To fill this knowledge gap, we can use comparative genomics to study evolutionary lineages, or clades, in which some species are unicellular and others multicellular, and in which the multicellular species exhibit varying degrees of cellular differentiation. Several eukaryotic clades fulfill these requirements [8, 9], notably the volvocine algae [10], a small extant group of freshwater green algae nested within Viridiplantae in the Chlorophycean order Volvocales (Fig. 1).
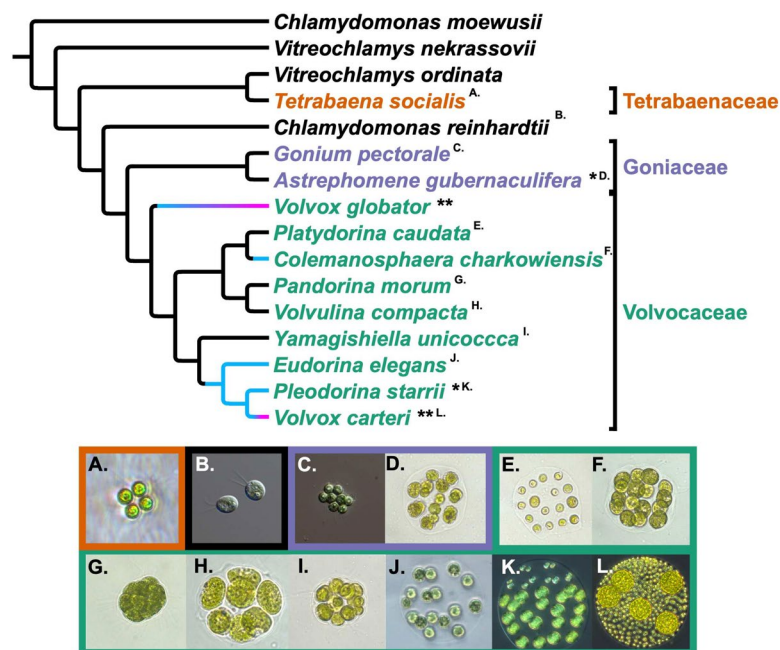

Fig. 1. Cladogram and photographs of volvocine genera. Depicted evolutionary relationships derived from Lindsey et al. [11] with Goniaceae modified to be monophyletic based upon more recent evidence from Ma et al. [12]. The three multicellular volvocine families (Tetrabaenaceae, Goniaceae, and Volvocaceae) are represented by orange, purple, and green font and borders. Unicellularity is denoted by black font and borders. Blue lines correspond to anisogamous lineages, and blue to pink gradient lines represent lineages in which oogamy has evolved. A single asterisk (*) denotes somatic cell differentiation, and double asterisks (**) denote complete germ-soma differentiation. Superscript letters correspond to their respective photographs. Photographs are arranged in the following order: (**A**) *Tetrabaena socialis*, (**B**) *Chlamydomonas reinhardtii*, (**C**) *Gonium pectorale*, (**D**) *Astrephomene gubernaculifera*, (**E**) *Platydorina caudata*, (**F**) *Colemanosphaera charkowiensis*, (**G**) *Pandorina morum*, (**H**) *Volvulina compacta*, (**I**) *Yamagishiella unicocca*, (**J**) *Eudorina elegans*, (**K**) *Pleodorina starrii*, (**L**) *Volvox carteri*. Photos are not to scale. Figure Credit for (**B**) and (**C**): Deborah Shelton

The volvocine algae are facultatively sexual [13]; thus, in addition to being a model system for the study of multicellularity and differentiation, they can also provide insight into another major transition: the evolution of male and female sexes via gametic differentiation. The definition and elaboration of the sexes, as well as the evolution of sexually selected traits, is predicated on the advent of anisogamy, gamete types that differ in size. Oogamy constitutes a specialized form of anisogamy wherein small male gametes are motile and large female gametes are not (e.g., vertebrates and land plants are oogamous). Anisogamy is hypothesized to have evolved from isogamy [14], in which the genetic determinants of mating type segregate during meiosis but the resulting gametes are identical in size and shape [15]. Because multicellular lineages such as animals and plants lack extant isogamous forms, the origins of anisogamy cannot be deduced using these groups. By contrast, the volvocine algae range from unicellular, isogamous species like *Chlamydomonas* to multicellular, oogamous species like *Volvox*, which exhibits complete germ-soma differentiation. Following in the vein of previous volvocine authors, terms such as "multicellular", "colonies", and "colonial" are used interchangeably. The use of this terminology (e.g., "colonies" to describe multicellular volvocine taxa) for this group can be traced back to the early 20th century [16]. We refer to any volvocine species with more than one cell as "colonial" or "multicellular". For additional context on terminology used to describe volvocine algae, see footnote in Kirk, 1998 pp. 115-116) [17].

Multicellular volvocine species are grouped within three families: Tetrabaenaceae, Goniaceae, and Volvocaceae. According to recent phylogenetic evidence [11], these families are paraphyletic. The unicellular *Vitreochlamys ordinata* is sister to the Tetrabaenaceae. This clade of *V. ordinata* + Tetrabaenaceae is sister to several *Chlamydomonas* species plus the multicellular Goniaceae + Volvocaceae. Both *Vitreochlamys* and *Chlamydomonas* are isogamous, undifferentiated unicells. The genera composing the Tetrabaenaceae, *Tetrabaena* and *Basichlamys*, are 4-celled colonies that are isogamous and undifferentiated. The Goniaceae comprises two genera, *Gonium* and *Astrephomene*. *Gonium* consists of flattened 8-32 celled colonies that are isogamous and undifferentiated. *Astrephomene*, also isogamous, includes 32-128 celled spheroidal colonies that exhibit somatic cell differentiation. Somatic cells engage in specific cellular functions, principally motility, are mortal, and do *not* pass on genetic information to subsequent generations. Among the 8 genera that comprise the Volvocaceae, *Pandorina*, *Volvulina*, and *Yamagishiella* are isogamous and undifferentiated; *Colemanosphaera*, *Eudorina*, and *Platydorina* are anisogamous and undifferentiated; *Pleodorina*

is anisogamous with somatic differentiation, and *Volvox* is oogamous with specialized germ and somatic cells. Specialized germ cells pass on genetic information and are distinct from undifferentiated cells in that they do not significantly contribute to colony motility.

Because the volvocine algae are closely related, but vary in cellularity, differentiation, and sexuality, comparative genomics offers the prospect of understanding how and when each of these traits evolved. Nearly 15 years ago Herron et al [18]. estimated divergence times for this group while analyzing the temporal sequence of certain developmental traits. Since then, a wealth of molecular, paleontological, and phylogenetic data has come to light, motivating us to reevaluate their conclusions. For example, multiple volvocine genomes [19–23] have been published, as well as the transcriptomes of all known extant volvocine genera [11, 24], including a new genus [25]; furthermore, genomes and transcriptomes have been sampled among different strains within many volvocine species [11, 24]. Also, the fossil record has been enriched by the discovery of a billion-year-old multicellular chlorophyte [26], firm dates of more than a billion years have been established for simple, multicellular red algae [27], and red algal fossils have been observed that may be 1.6 billion years old [28]. Using a dataset of 40 nuclear-protein coding genes, Lindsey et al [11]. recently reported that *(i)* multicellularity evolved at least twice in this group, *(ii)* cell differentiation evolved at least four times, *(iii)* anisogamy evolved from isogamous ancestors at least three times, and *(iv)* oogamy evolved at least three times. The first two of these findings have since been bolstered by Ma et al. [12].

When divergence times were estimated by Herron et al. [18], their molecular dataset consisted of 5 chloroplast protein-coding genes plus the 18S nuclear, small ribosomal subunit across 35 volvocine taxa. Subsequent studies undertaking ancestral-state reconstruction [29–32] of this group have relied solely on the 5 chloroplast gene dataset for their analyses. Unsurprisingly, the resulting trees exhibited essentially the same branching order, none of which reflect the new insights gained by Lindsey et al. [11] As noted, Ma et al. [12] also inferred divergence times of the volvocine algae using large nuclear datasets, employing a single relaxed clock model. However, these researchers did not use their phylogenetic inferences to perform ancestral-state reconstruction of sexual and developmental traits.

Here, we provide a geologically-calibrated roadmap of when multicellularity, differentiation, and anisogamy arose in the volvocine algae. Since no reliable fossils exist for this group, we sampled fossils across the Archaeplastida, or kingdom Plantae (*sensu lato*), where reliable fossils are plentiful. Specifically, we sampled 14 fossil taxa
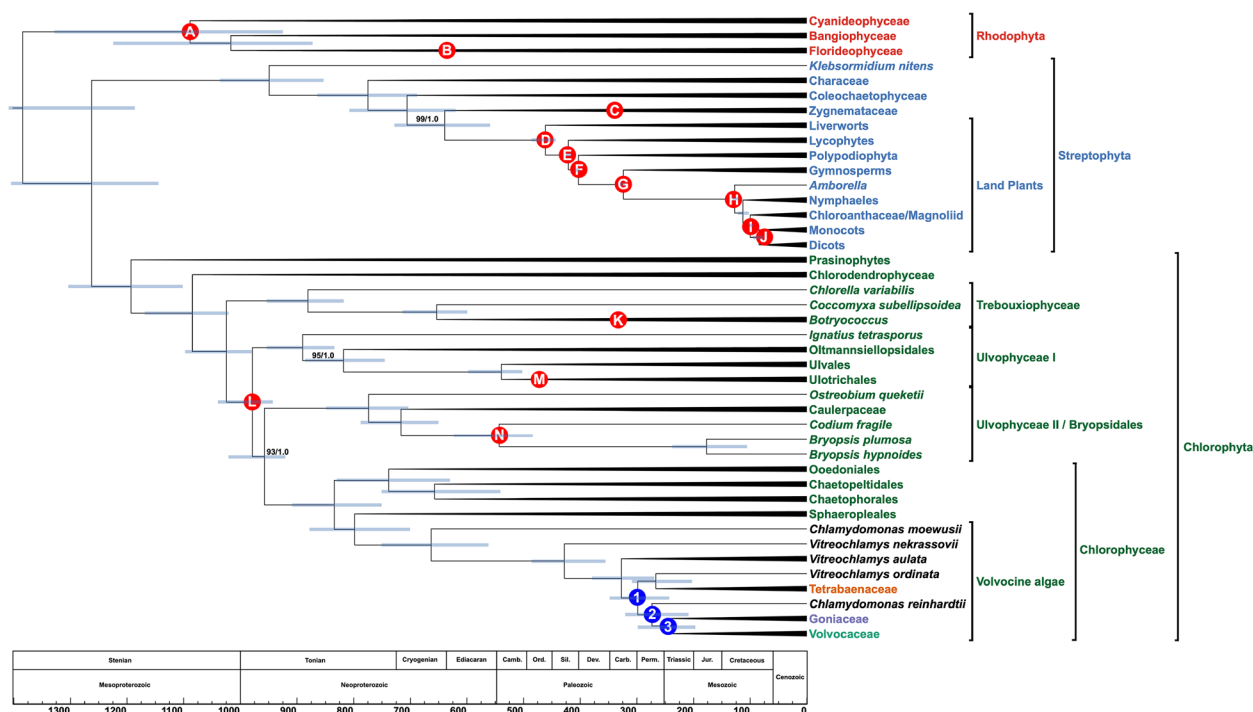
Fig. 2. Time-calibrated phylogeny of the Archaeplastida. Branching order of the tree was inferred under maximum-likelihood analysis from an aligned amino acid, concatenated dataset of 263 nuclear genes. Numbers on branches represent bootstrap and posterior probability values, respectively. Branch lengths, corresponding to time, were inferred under the CIR relaxed clock model using 16 most clock-like genes as determined by Sortadate. Blue bars correspond to the inferred 95% HPD interval for each node. Red bubbles correspond to calibrated nodes (Table 1), and blue bubbles correspond to key divergences among the volvocine algae (Figure S4B). Members of the multicellular volvocine algae (Tetrabaenaceae, Goniaceae, and Volvocaceae) are denoted in orange, purple, and green. Taxa in black font are unicellular volvocine algae

across the three major Archaeplastida clades (Rhodophyta, Streptophyta, and Chlorophyta) selecting them so as to calibrate our time-tree over an interval of at least one billion years. Our molecular data consist of amino acid sequences for 263 single-copy nuclear genes drawn from 164 taxa across the Archaeplastida. Sequences were obtained both from publicly available genomic and transcriptomic datasets as well as from our own published RNA-Seq dataset [11]. Our goal was to reanalyze the divergence times and the gain and/or loss of traits related to multicellularity, differentiation, and sexuality for the volvocine algae. Our data, presented as ancestral-state reconstructions, lead us to conclude that there were two independent origins of multicellularity in the volvocine algae, as predicted by Lindsey et al. [11] and Ma et al. [12] Moreover, under one of our four relaxed clock models reassessment of divergence times indicates that multicellularity may have originated in the Goniaceae + Volvocaceae sometime between the Carboniferous and Triassic, much older than previous inferrences [12, 18]. Among the Tetrabaenaceae we find that multicellularity may have arisen as early as the Carboniferous or as recently as the Cretaceous. Using our chronogram, we

also reassessed the developmental milestones proposed in David L. Kirk's 12-Step Program [33] to differentiated multicellularity in the volvocine algae. Through ancestral state reconstructions, we show that the temporal sequence of events he hypothesized almost 20 years ago is essentially correct. To investigate multicellularity's role in the evolution of sex in the volvocine algae, we reconstructed 7 sexual traits. When the gain and loss of these traits are mapped onto our time tree, multicellularity's role as a driver of sexual evolution becomes evident, consistent with the results of Hanschen et al [32]. Finally, our analyses of morphological, molecular, and divergence time data suggest the possibility of a cryptic species in the multicellular Tetrabaenaceae.

Results and Discussion

Selection of orthologues in taxa across the Archaeplastida provides the basis for inferring how multicellularity and cellular differentiation evolved in the volvocine algae. We sampled a total of 164 taxa representing all major clades of Archaeplastida: Rhodophyta, Streptophyta, and Chlorophyta. Specifically, we sampled 12 rhodophytes across 3 red algal subclades,

45 streptophytes representing 13 major green algal and land plant subclades, and 107 chlorophytes, representing prasinophytes, Ulvophyceae, Trebouxiophyceae, and Chlorophyceae, including 68 unicellular and multicellular volvocine algae (Fig. 2, Additional File 1: Figure S1, Figure S2, and Figure S3). We chose the red algae for our outgroup, as multiple studies indicate that they are sister to the green algae + land plants (Embryophyta), and that the red and green algae share a common plastid ancestor [34, 35]. Although our chief aim was to discern, on a geological timescale, the evolution of traits leading to multicellularity and differentiation in the volvocine algae, we included other green algal genera such as *Coleochaete* and *Tetraselmis* to further illuminate the history of Viridiplantae. The former is generally considered to be relatively closely related to land plants [36, 37], while the latter is believed to be sister to the three major Chlorophyta clades [38]. Because all multicellular lineages necessarily evolved from unicellular ancestors, clarifying which lineages are sister to which multicellular clades will enable a deeper understanding of how this major transition occurred.

A single dataset consisting of 263 single-copy, protein coding genes was compiled and analyzed under Maximum-Likelihood (ML), Bayesian inference (BI), and coalescent-based (CB) phylogenetic methods. These 263 genes were conserved across the three main clades of Archaeplastida. Our concatenated alignment for ML and BI analyses represents an aggregate of 79,844 amino acids, equivalent to 239,532 nucleotide positions, with a total of 62,106 parsimony-informative sites. All raw reads used to complete our single-gene and concatenated alignments encompassing 164 total taxa were mined from previously published data located in public repositories (Additional File 2: Table S1) [11, 20–22, 24, 39–83].

Phylogenetic analyses of the volvocine algae indicate multiple independent origins of multicellularity and differentiation. Our ML, BI, and CB analyses all indicate two independent origins of multicellularity among the volvocine algae: one in the lineage leading to the Tetrabaenaceae and another in the lineage leading to the Goniaceae + Volvocaceae (Fig. 2, Additional file 1: Figure S1, Figure S2, and Figure S3). These results bolster the findings of Lindsey et al. [11] and Ma et al. [12] As before, our ML and BI analyses have identical branching orders. The Tetrabaenaceae + *Vitreochlamys ordinata* are shown to be sister to several *Chlamydomonas* species + the Goniaceae and Volvocaceae in our ML and BI trees (Maximum-likelihood bootstrap support (MLBS) = 100, Bayesian posterior probablility (BPP) = 1.0) (Fig. 2, Additional file 1: Figure S1, and Figure S2), and this topology both corroborates and expands on with those presented in several earlier studies [11, 12]. Our CB tree,

however, indicates a slightly different branching order for the Tetrabaenaceae. The resulting coalescent-based tree shows the Tetrabaenaceae + *V. ordinata* forming a clade with *Chlamydomonas reinhardtii* and its relatives, and this clade is shown to be sister to the Goniaceae + Volvocaceae (Coalescent posterior probability (CPP) = 1.0) (Additional file 1: Figure S3). Of note, the sister relationship between the Tetrabaenaceae + *V. ordinata* and the *Chlamydomonas* clade is poorly supported in our CB analysis (CPP = 0.34).

In accordance with the findings of Lindsey et al. [11], all three of our phylogenetic analyses indicate a minimum of 4 independent origins of somatic cellular differentiation and a minimum of 3 origins of anisogamy among the volvocine algae. Origins of somatic cell differentiation occur in the following lineages: (*i*) *Astrephomene*, (*ii*) section *Volvox*, (*iii*) *Pleodorina thompsonii*, and (*iv*) the *Pleodorina japonica* + *Volvox carteri* clade. Anisogamy evolved from isogamous ancestors at least three times in the following lineages: (*i*) *Astrephomene,* (*ii*) section *Volvox*, and (*iii*) in the *Eudorina* + *Volvox* + *Pleodorina* (EVP) clade.

In contrast to several recent volvocine studies [11, 23], all three of our phylogenetic analyses support the major conclusion that the Goniaceae are monophyletic, albeit with varying support values (MLBS = 95, BPP = 1.0, CPP = 0.67). This conclusion was reached by multiple earlier investigations [12, 18, 31, 32, 84–91]. Consistent with results of Lindsey et al. [11] and Ma et al. [12], we find that *Volvox* section *Volvox* is sister to the *Pandorina* + *Volvulina* + *Colemanosphaera* (PVC) and the (EVP) clades within the Volvocaceae. Our ML + BI and CB trees contain minor branching order differences in the EVP clade, but these have no bearing on our major findings.

Variation in divergence times inferred under different relaxed molecular clock models necessitate validation tests. All divergence time estimates were inferred by Phylobayes 4.1b [92] under a Bayesian approach for our inferred ML and CB species trees. For each topology, a total of 14 nodes were calibrated across Rhodophyta, Streptophyta, and Chlorophyta (Fig. 2 and Table 1), and four relaxed clock models: autocorrelated lognormal (LN) [93] and Cox-Ingersoll-Ross (CIR) [94] models, and the uncorrelated gamma (UGAM) and white -noise (WN) models [95]. Autocorrelated models (i.e., LN and CIR) allow the rate of evolution to vary across branches, and the rate of evolution is more similar along branches for closely related species compared to distantly related taxa [93, 94]. Uncorrelated models (i.e., UGAM and WN) assume each branch of will have its own unique rate of evolution irrespective of rates across branches for closely and distantly related species [94, 95]. Inferred divergence times established under all clock models are

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 6 of 21

Table 1 Fossils used to calibrate nodes in Archaeplastida tree

Node	Calibration	Age	References
A	*Bangiomorpha*	1047 +13/-17	Butterfield et al. (2000) [96]; Gibson et al. (2018) [27]
B	Oldest Florideophycidae	609 +/-5	Xiao et al. (2004) [97]; Zhou et al. (2017) [98]
C	Oldest Zygnemataceae	MIN: 350	Colbath and Grenfell (1995) [99]
D	First Land Plants	MIN: 480	Strother & Foster (2021) [100]
E	Oldest Tracheophytes (*Cooksonia* with tracheids)	423-419	Morris et al. (2018) [101]
F	Fern-Seed Plant Split	MIN: 385	Morris et al. (2018) [101]
G	Extant Gymnosperms-Angiosperms Split (based on oldest *Cordaites*)	330-323	Morris et al. (2018) [101]
H	*Amborella* and Nymphaeles Split from other Plants (Barremian flowers and pollen)	129-125	Morris et al. (2018) [101]
I	Chloroanthaceae/Magnoliid versus monocots plus eudicot split (Chloranthaceous fossils)	MIN: 125	Morris et al. (2018) [101]
J	Monocot-Eudicot Split (*Liliacidites* monocot pollen)	MIN: 113	Morris et al. (2018) [101]
K	*Botryococcus braunii*	358-356	Colbath and Grenfell (1995) [99]
L	*Proterocladus*	1056-948	Tang et al. (2020) [26]
M	Ulotrichales: *Vermiporella*	470-458	Kozlowski & Kazmierczak (1968) [102]
N	*Protocodium*	MIN: 541	Chai et al. (2022) [103]

largely consistent across the three major red and green algal clades (Figure S4). However, there are nodes such as the root age, earliest rhodophyte divergence, and major divergences within the volvocine algae where one model infers a date markedly younger or older than the others (Fig. 3).

The estimated mean root age of all tested clock models varies by ~630 million years (MY) with the WN model estimating the earliest red algal divergence as ~2016 million years ago (MYA), and the CIR model estimating it as late as ~1385 MYA (Additional file 1: Figure S4B). Averaging the mean root ages of the LN, CIR, and UGAM models produces an average root age of ~1503 MYA, reducing the variance in estimated mean root age between WN and the others to ~500 MY. Similarly, for the earliest rhodophyte divergence, the WN model inferred a mean date significantly older than all other models by at least 300 MY, whereas the LN, CIR, and UGAM models estimated mean ages within <100 MY of each other. Large date discrepancies such as these may be solely due to differences in clock algorithms [104, 105].

For all volvocine algae divergences, the UGAM model estimated ages markedly younger than the three other models we tested (Additional file 1: Figure S4A, B (divergences 9-11)). It is noteworthy that the inferred ages of the UGAM model for the volvocine algae are very similar to the dates inferred by Ma et al. [12], who used a single relaxed clock model. When taking the average of the mean dates inferred by the LN, CIR, and WN models, there is a consistent ~130 MY difference between the average and a mean node date inferred by the UGAM model for this group. Significantly, the UGAM 95% highest posterior density (HPD) intervals for the volvocine algal divergence do not overlap with any of the 95% HPD intervals inferred by other models, conversely the other three models see a comfortable overlap in their 95% HPD intervals for volvocine divergence estimates.

Given that the WN and UGAM models inferred outlier dates for certain key divergence events, we decided to exclude both as clock models. Ages inferred by the CIR model in this study were never observed to be outliers for key divergence events, and the ages inferred by the CIR model for the volvocine algae are nearly identical to the WN model's estimates for this group. Additionally, divergence estimates inferred under the CIR model for the volvocine algae were the most conservative

(See figure on next page.)

Fig. 3. Estimated divergence times of the volvocine algae. Branching order of the tree was inferred under maximum-likelihood analysis from an aligned amino acid, concatenated dataset of 263 nuclear genes. Numbers on branches represent bootstrap and posterior probability values, respectively. Branch lengths, corresponding to time, were inferred under the CIR relaxed clock model using 16 most clock-like genes as determined by Sortadate. Blue bars correspond to the inferred 95% HPD interval for each node. Green bubbles correspond to a developmental trait gain, and red bubbles corresponds to loss of a trait. The figure table lists the 12 developmental traits identified by Kirk in their original order. Blue bubbles indicate key divergences in the volvocine algae. Members of the multicellular volvocine algae (Tetrabaenaceae, Goniaceae, and Volvocaceae) are denoted in orange, purple, and green. Taxa in black font are unicellular volvocine algae
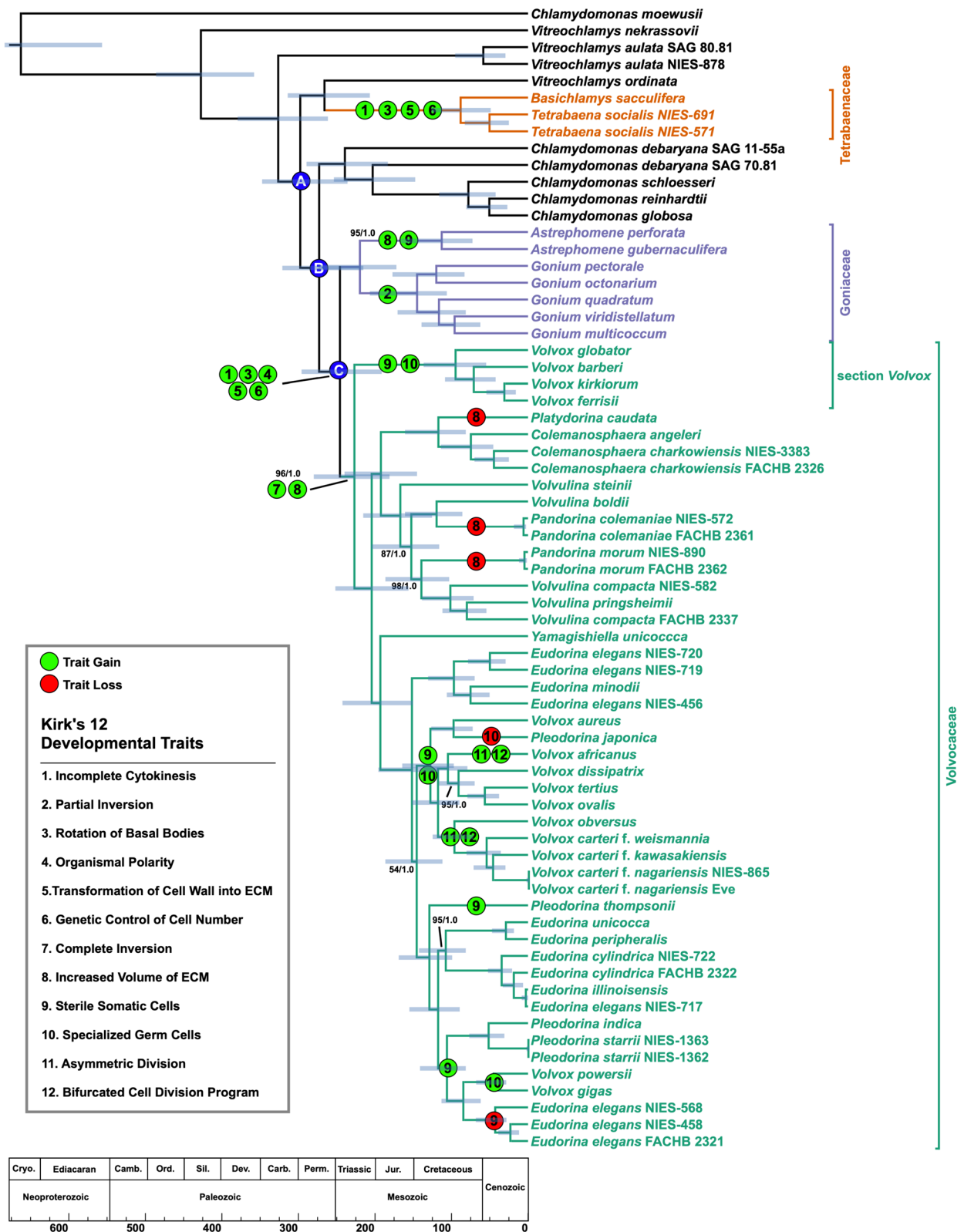
Fig. 3. (See legend on previous page.)

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 8 of 21

among the LN, WN, and CIR results. Altogether, the foregoing considerations prompted us to report dates for major divergence events using data produced by the CIR relaxed clock model (Additional file 1: Figure S4A, B). Lepage et al. [94] performed model comparisons of the CIR, LN, UGAM, and WN relaxed clock models using real, as opposed to simulated, molecular datasets that varied in type (nuclear vs mitochondrial and nucleotide vs amino acid) and number of taxa. Their tests concluded that autocorrelated clock models performed best against the various datasets used in their study. Furthermore, LN and CIR models outperformed UGAM and WN models when taxonomic sampling was high, and the CIR model was recommended by the authors for dense taxonomic datasets. Fossil cross validation tests were performed as described earlier for each relaxed clock model, with the objective of identifying a best-performing analysis. However, no relaxed clock model performed markedly better than any of the others (Additional file 2: Table S2).

Fossil selections and Archaeplastida phylogenetic results and divergence times. Since no reliable fossils exist for the volvocine algae, we selected 14 fossil taxa across the Archaeplastida where reliable fossils are abundant (Fig. 2 and Table 1). For each primary fossil calibration used in this study, each calibration point was constrained to a range rather than a fixed-point estimate, acknowledging the inherent uncertainty in fossil ages. With each date range, we specified a soft bound where 2.5% of the total probability mass is positioned outside of the specified lower and/or upper bounds. Detailed information regarding each fossil taxon and calibration point may be found in Additional file 3: Supplementary Methods [18, 26, 27, 75, 96, 97, 99–142]. Moreover, our results indicate that divergence times among the Archaeplastida inferred under the CIR model are largely congruent with those of prior studies. Maximum-Likelihood (ML), Bayesian inference (BI), and coalescent-based (CB) phylogenies for Archaeplastida are largely congruent, well-supported, and, overall, confirm relationships inferred during past studies (Additional file 1: Figure S1, Figure S2, and Figure S3). The ML and BI phylogenies have identical branching orders (Additional file 1: Figure S1 and Figure S2), and the only differences between the ML + BI and CB trees occur in Chlorophyta, the major green algal and seaweed clade. With the volvocine algae being our principal clade of interest, we elected to report general Archaeplastida phylogenetic inferences and divergence times in Additional file 3: Supplementary Methods.

Molecular clock analysis reveals that the crown Tetrabaenaceae-Goniaceae-Volvocaceae clade arose sometime between the Carboniferous and Triassic, and ancestral state reconstructions confirm our prior inference that multicellularity arose twice. According

to our ancestral state reconstructions (ASR) (Additional file 1: Figures S5 and 6), multicellular lineages share a common unicellular ancestor ~298 MYA (CIR relaxed clock model, 95% HPD interval 349-237 MYA), marking the emergence of the crown Tetrabaenanceae-Goniaceae-Volvocaceae (TGV) clade sometime in the Carboniferous to the Triassic periods (Figs. 2 and 3, node A). Our 95% HPD interval for the TGV clade partly overlaps with ranges reported in previous studies [12, 18]. For example, the range reported by Herron et al. [18] is mostly confined to the Triassic (260-209 MYA), whereas that reported by Ma et al. [12] extends across the Early Triassic and Early Jurassic (251-180 MYA).

From the TGV clade's common ancestor (Fig. 3, node A), our ASR results show that two unicellular lineages diverged (Additional file 1: Figures S4 and 5), each of which led to independent origins of multicellularity in the volvocine algae (Figs. 2 and 3). From one branch of this split, a multicellular ancestor of the Tetrabaenaceae diverged from a unicellular ancestor of *Vitreochlamys* ~267 MYA (316-206 MYA), and from the other branch a multicellular ancestor of the Goniaceae + Volvocaceae diverged from a unicellular ancestor of *Chlamydomonas* ~274 MYA (323-214 MYA (Fig. 3B and Fig. 3). These results confirm two independent origins of multicellularity among the volvocine algae as indicated by previous phylogenies [11, 12, 86]. According to our data, multicellularity arose in Tetrabaenaceae sometime between the Jurassic and Cretaceous (126-49 MYA), while multicellularity arose in the Goniaceae + Volvocaceae during the Permian-Jurassic (297-190 MYA).

Ancestral state reconstructions suggest that the steps leading to differentiated multicellularity in the volvocine algae occur in the order envisioned by Kirk. In 2005, David L. Kirk, a developmental and cell biologist, hypothesized how differentiated multicellularity, as manifest in *Volvox carteri*, might have evolved from a unicellular, *Chlamydomonas*-like ancestor [33]. In his essay, attention was brought to 12 ontogenetic traits (Fig. 3) that occur in the developmental cycle of *V. carteri* as it matures from a single-cell to a >500-celled organism. These 12 traits were then shown to be present in various extant Goniaceae and Volvocaceae, allowing them to be positioned in a stepwise manner along a simplified volvocine phylogeny, beginning with the most recent common ancestor of *C. reinhardtii* and *V. carteri*. Through the ordered mapping of these traits, Kirk suggested a developmental program that could explain how multicellularity with a division of labor might evolve. This hypothesis, known as Kirk's 12-Step Program to differentiated multicellularity, has been formally tested once by Herron et al. [18] using ancestral-state reconstructions from a previous study [84]. Herron et al. [18] suggested that the

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 9 of 21

12-Step Program was not as orderly as hypothesized due to the phylogenetic position of the Tetrabaenaceae, which are absent from Kirk's hypothesis. Below, we show that, using ancestral-state reconstructions based on our more robust phylogeny, Kirk's 12-Step Program occurs temporally much as originally hypothesized.

For continuity, Kirk's original order of 12 developmental traits has been retained in our analysis (Fig. 3). Approximately 298 MYA (349-237 MYA), two lineages emerged from the common unicellular ancestor of the colonial volvocine algae (Fig. 3, node A). One of these lineages eventually gave rise to the Tetrabaenaceae (Fig. 3). In this lineage, *Steps 1, 3, 5* and *6* evolved in the multicellular ancestor of *Basichlamys* and *Tetrabaena* (Fig. 3 and Additional file 1: Figures S7-14). According to our ASR, this multicellular ancestor was established ~89 MYA (126-49 MYA); no other traits identified by Kirk are present either in this ancestor or in any of the known extant Tetrabaenaceae. Importantly, the number of Kirk's traits gained in the lineage leading to the Tetrabaenaceae is twice that reported by Herron et al. [18] This conclusion is supported by recent studies confirming incomplete cytokinesis (*Step 1*) and basal body rotation (*Step 3*) in *Tetrabaena* [143, 144].

The second lineage diverged from a *Chlamydomonas*-like ancestor ~274 MYA (323-214 MYA), and eventually gave rise to the multicellular Goniaceae and Volvocaceae (Fig. 3, node B). ASR indicates that the common ancestor to the Goniaceae and Volvocaceae was multicellular (Additional file 1: Figures S5 and 6), that it had evolved 5 of the 12-Step Program traits over a period of ~27 MY (Fig. 3, node C) (Additional file 1: Figures S7-16), and was most likely non-spheroidal (Additional file 1: Figures S17 and 18). By ~247 MYA (297-190 MYA), the non-spheroidal, multicellular ancestors to the Goniaceae and Volvocaceae had evolved: (*Step 1*) an incomplete cytokinesis where daughter cells were left connected to each other via cytoplasmic bridges, (*Step 3*) peripheral colony cells with rotated flagellar basal bodies beating in parallel allowing the colony to efficiently mobilize through water, (*Step 4*) central-to-peripheral polarity similar to what is seen in *Gonium*, (*Step 5*) daughter cells embedded into an extracellular matrix with two distinct boundaries, and (*Step 6*) colony cell number controlled by its genetic code rather than the environment (Fig. 3).

For some volvocine algae, a developmental process known as "inversion" must occur. During development, certain bowl-shaped and spheroidal colonies have their flagella pointing inside the bowl or spheroid, preventing the colony from being able to swim. The process of inversion flips the curvature for bowl-shaped volvocines and turns spheroidal colonies inside out; thus, allowing the flagella to provide motility to the colony. In the volvocine

algae, inversion occurs in the following three states: *i*) does not occur, *ii*) partial inversion, and *iii*) complete inversion. For unicellular volvocines, the Tetrabaenaceae, and *Astrephomene*, inversion does not occur in their developmental cycle. Due to the way volvocine cells divide, embryos resulting in *Gonium* and Volvocaceae colonies must undergo "partial" and "complete" inversion, respectively, so that the eventual colonies have the ability swim. For *Gonium*, embryos are in the shape of a bowl with their flagellar ends positioned in the concave region, and volvocacean embryos, in similar fashion, have their flagellar ends located inside the developing colony. By ~146 MYA (206-106 MYA), the ancestor to *Gonium* had developed (*Step 2*) "partial" inversion whereby a curvature reversal in the embryo occurs, allowing the flagellar ends to be situated on the exterior of the colony. By ~228 MYA (279-175 MYA), volvocacean embryos developed "complete" inversion (*Step 7*), where cells, in a coordinated fashion, move and curl outward along the anteroposterior axis, resulting in an embryo whose flagellar ends point outside the colony (Additional file 1: Figure S19 and 20) (*Step7*). In both examples, cells in colonies are held in position by cytoplasmic bridges formed via incomplete cytokinesis (*Step 1*). Along with *Step 7*, by ~228 MYA (279-175 MYA) spheroidal volvocacean ancestors had evolved body plans with increased volume. Diminished colony volume occurred secondarily in lineages leading to *Platydorina* and the paraphyletic *Pandorina* (Fig. 3).

The spheroidal body plan in the Volvocaceae and in the lineage leading to *Astrephomene* likely consisted of cells positioned along a colony's periphery, resulting in a transparent organism with an interior filled with extracellular matrix. Despite the morphological similarity among spheroidal colonies, earlier studies [145, 146] and our ASR results indicate that such a body plan arose independently twice in the volvocine algae (Additional file 1: Figures S17 and 18). Yamashita et al. [145] elucidated the developmental mechanisms that underlie the two independently evolved forms. During *Astrephomene* embryogenesis daughter protoplasts gradually rotate their apical ends so that flagella extend towards the posterior of the growing colony after each successive cell division. Development proceeds differently in the Volvocaceae. In volvocacean early embryos, chloroplasts are positioned at one end of each protoplast, facing outward. As the embryo matures, these cells change shape, with the chloroplast ends forming acute edges [146]. This cell shape change results in bending the epithelium at an opening of the embryo known as the phialopore [147]. Epithelial bending causes the colony to turn itself inside-out allowing flagella to be located on the exterior of the developing colony. This process in the Volvocaceae

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 10 of 21

is known as "complete" inversion [145]. Yamashita and Nozaki [146] showed that neither of these developmental mechanisms occur in *Tetrabaena* or *Gonium* embryogenesis and thus concluded that spheroidal body plans in the volvocine algae are independently derived via separate mechanisms.

In our updated phylogeny, the most recent common ancestor of the Volvocaceae exhibited 7 of the 12 traits identified by Kirk (Fig. 3). Moreover, there is only a single conflict in the order in which they evolved according to his hypothesis. This conflict arises from ambiguity as to when inversion first evolved in a volvocine ancestor; indeed, ASR yields alternative inferences depending upon which analytical assumptions are used. When we consider this conflict as an unordered 3-state problem (no inversion, "partial" inversion, or "complete" inversion) (Additional file 1: Figure S19-20), which assumes that no state is a prerequisite for another to evolve, the ancestor to the Goniaceae and Volvocaceae is shown not to have evolved inversion (Fig. 3, node C). However, when the analysis is performed as an "ordered" 3-state problem, where "partial" inversion is required for "complete" inversion to evolve, the common ancestor to the Goniaceae and Volvocaceae is shown to have partial inversion (Additional file 1: Figure S21). Lastly, when inversion is treated as a 2-state problem (e.g., no inversion, inversion), the common ancestor of the two clades is shown to have evolved inversion (Additional file 1: Figure S22-23). This last result conflicts with Yamashita and Nozaki's [146] explanation for how inversion in *Gonium* and the Volvocaceae are mechanistically distinct. Inversion should be treated as different states (e.g., "partial" and "complete"), as in our first example of an "unordered" 3-state problem, because it is unknown whether "partial" inversion is necessary for "complete" inversion to evolve. Thus, we report that only the ancestor to *Gonium* had evolved "partial" inversion (*Step 2*) by ~146 MYA (206-106 MYA).

The evolution of somatic cellular differentiation (*Step 9*) is the first step towards a complete reproductive division of labor where a specific cell line (the soma) evolved to house cells (germ-line) that transmit genetic information to subsequent generations. Somatic cell lines observed in *Astrephomene*, *Pleodorina*, and *Volvox* are distinct in that they assume specific cellular functions, notably motility, exhibit a finite lifespan, and do *not* pass on genetic information. Two other cell types are present in the volvocine algae: undifferentiated cells and fully-differentiated germ cells. Undifferentiated cells, the ancestral cell state in volvocine algae, participate in motility and can undergo mitotic division to produce the next generation. By contrast, fully-differentiated germ cells, observed only in *Volvox*, also undergo mitotic division to generate progeny, but these cells never significantly participate in motility. By comparing our ASR results to our phylogenetic results we can infer specific number of independent origins of sterile somatic cells. Thus, the gain of sterile somatic cells (*Step 9*) occurred five times in three different genera of the volvocine algae (Fig. 3). According to our results, true somatic differentiation arose in (*i*) *Pleodorina thompsonii* sometime after it diverged from the common ancestor it shared with *V. carteri* (169-99 MYA) and in the common ancestors of: (*ii*) *Astrephomene* (170-75 MYA), (*iii*) section *Volvox* (136-57 MYA), (*iv*) the *Volvox* + *Pleodorina japonica* clade (166-98 MYA), and (*v*) *P. starrii* + *V. gigas* (140-82 MYA) (Fig. 3 and Additional file 1: Figures S24 and 25). In addition to this trait being gained, it was also lost once in the *Eudorina* clade sister to *V. gigas* + *V. powersii*.

The advent of true germ-soma differentiation in the volvocine algae is characterized by evolution of specialized germ cells (*Step 10*) that do not significantly contribute to colony motility. In genera like *Astrephomene* and *Pleodorina* that have not made this step, colonies consist of somatic cells as well as biflagellate gonidia, cells that contribute to motility and can undergo mitotic division. Indeed, in these taxa gonidia greatly outnumber somatic cells. In *A. gubernaculifera* a 32 or 64-celled colony may have 30 or 60 gonidia [148], respectively, and 128-celled *P. californica* colonies can have >70 gonidia [149]. Gonidia in these instances contribute significantly to motility until cell division is initiated, which ultimately leads to hatched daughter colonies and subsequent death of the mother. In contrast, *Volvox* colonies are all composed of >500 somatic cells and far fewer flagellated or aflagellate gonidia. For example, at the end of *V. rousseletii*'s embryogenesis, a colony may consist of ~2000-14,000 biflagellate, nearly indistinguishable cells with only 1-16 gonidia [17, 150]. In *V. rousseletii*, gonidia contribute to motility for <20% of a colony's asexual life cycle before resorbing their flagella and undergoing mitotic division [17]. *Volvox carteri* colonies are typically composed of ~2000-6000 cells with ~8 gonidia [150]. Somatic and gonidial cells are clearly defined during *V. carteri*'s embryogenesis. Gonidial cells are aflagellate; thus, in this species gonidia never contribute to colony motility during its life cycle [17]. Therefore, *Volvox* spp. are the only extant volvocine algae to have evolved specialized germ cells (*Step 10*). In this group, specialized germ cells evolved by ~96 MYA (136-57 MYA) in section *Volvox*, ~46 MYA (68-28 MYA) in the ancestor to *V. gigas* and *V. powersii,* and ~129 MYA (166-98 MYA) in *V. aureus* and *V. carteri*'s most recent common ancestor (Fig. 3 and Additional file 1: Figures S24 and 25). This trait was lost once in *P. japonica* (Fig. 3 and Additional file 1: Figures S24 and 25).

The final two steps of the 12-Step Program are coupled, consisting of asymmetric cell division (*Step 11*) and a bifurcated cell division program (*Step 12*) (Additional file 1: Figures S26-29). Both traits evolved twice in the genus *Volvox*: (*i*) *V. africanus* sometime after ~106 MYA and (*ii*) ~98 MYA for the ancestor of *V. obversus* and *V. carteri*. During *V. carteri* embryogenesis, anterior embryo cells undergo asymmetric divisions producing both small and large cells. Size controls cell fate during embryogenesis. Cells <8μm may divide up to 6 times before they arrest, whereas larger cells undergo 1-2 asymmetric cell divisions. Smaller cells, and the remaining embryonic cells, eventually become soma, whereas the larger cells are destined to become specialized germ cells. Through these two coupled processes (*Steps 11* and *12*), complete division of labor is achieved between somatic and germ cell lines.

Ancestral state reconstruction shows that evolution did not progress linearly within the volvocine algae from a unicellular organism like *Chlamydomonas* to one that is multicellular and differentiated like *V. carteri.* Numerous trait gains and losses occurred in the history of this clade, exemplifying the non-linear nature of evolution. However, Kirk's 12-Step Program does not attempt to linearize the evolutionary history of this group, but rather to provide a roadmap of steps needed for true differentiated multicellularity to evolve in the volvocine algae. Our results indicate that the temporal sequence in which Kirk organized his 12 traits aligns well with the inferred evolutionary history of this group.

Ancestral-state reconstructions support previous findings that multicellularity drives evolution of sexual traits in the volvocine algae. The volvocine algae in their various forms have given rise to multicellular lineages that express a variety of sexual traits. For this group, asexual reproduction via mitosis is the typical means by which the next generation of unicells or colonies is produced. Sexual reproduction is facultative, and is precipitated by reduction in environmental nitrogen and, in some cases, a sex-inducing hormone [151]. Regardless, a diploid zygote is formed after fertilization, resulting in a spore. Meiosis only occurs in the volvocine algae when spores germinate to produce haploid progeny [151].

Unicellular forms, such as *Chlamydomonas* (*sensu* Pröschold et al. [152]), and in some multicellular forms (*Astrephomene*, *Basichlamys*, *Gonium*, *Pandorina*, *Tetrabaena*, *Volvulina*, and *Yamagishiella*), *plus* and *minus* type gametes are morphologically identical (isogamy). However, in many multicellular forms (*Colemanosphera*, *Eudorina*, *Platydorina*, and *Pleodorina*) the two gamete types differ in size, with the smaller of the two recognized as being male (anisogamy) [153]. In certain instances (*Volvox*), gametic differentiation may be further exaggerated, with the small male gamete being motile, and the large female gamete being sessile (oogamy). One longstanding question is: why does anisogamy evolve in the first place? Another is: what role does multicellularity play in gametic differentiation? Because the volvocine algae encompass both unicellular and multicellular forms, and because they exhibit the full range of known gamete types, they are especially well-suited to addressing these two fundamental questions.

Hanschen et al. [32] systematically investigated whether multicellularity is a prerequisite for the evolution of anisogamy and whether multicellular volvocines exhibit greater complexity in the form of more elaborate sexual traits. Their study, like ours, concludes that multicellularity is correlated with the acquisition of anisogamy, as all anisogamous volvocines are multicellular, and anisogamy appears to evolve from isogamous ancestors (Fig. 4) [32]. Using our updated phylogeny, with its newly inferred relationships, we conducted ASR for 7 sexual traits identified in previous studies [31, 32]. Except where noted (Fig. 4), we focused on the same strains used in the two prior studies of how sexual traits are distributed among the volvocine algae. Like those earlier studies, we excluded *Vitreochlamys*, *Pleodorina thompsonii*, and *Volvox ovalis* from certain ASR analyses due to a lack of information about how they reproduce sexually.

Unicellular ancestors of the volvocine algae, like *Chlamydomonas*, express the following sexual phenotypes: *(i)* isogamy, *(ii)* full meiotic hatching of zygospores, *(iii)* dioecy, *(iv)* outcrossing, *(v)* external fertilization of the egg, and no evidence of sexual dimorphisms such as (*vi*) "extrafertile" females or (*vii*) "dwarf" males. Prior to the evolution of anisogamy, the number of "gone cells" released from a zygospore decreases. (Note: a gone cell constitutes the haploid meiotic product arising from a diploid zygospore [154].) *Chlamydomonas reinhardtii*, *Tetrabaena*, and *Gonium* produce all four haploid gone cells per diploid zygospore [88]. According to our ASR results full meiotic hatching is lost either once or twice in the volvocine algae (Additional file 1: Figures S30 and 31) Our MrBayes ASR indicates a single loss, but this could be an artifact arising from low sampling of taxa that exhibit full meiotic hatching. In contrast, and consistent with Hanschen et al. [32], our Phytools ASR indicates two losses of full meiotic hatching in the following lineages: once in *Astrephomene* sometime after diverging from *Gonium* ~220 MYA (274-170 MYA), and again in the Volvocaceae ~228 MYA (279-175 MYA) (Fig. 4 and Additional file 1: Figures S30 and 31). Each loss represents a reduction in meiotic products from four haploid gone cells to a single haploid gone cell and three polar bodies. Among our identified sexual traits, no other gain or loss occurs until the evolution of anisogamy.
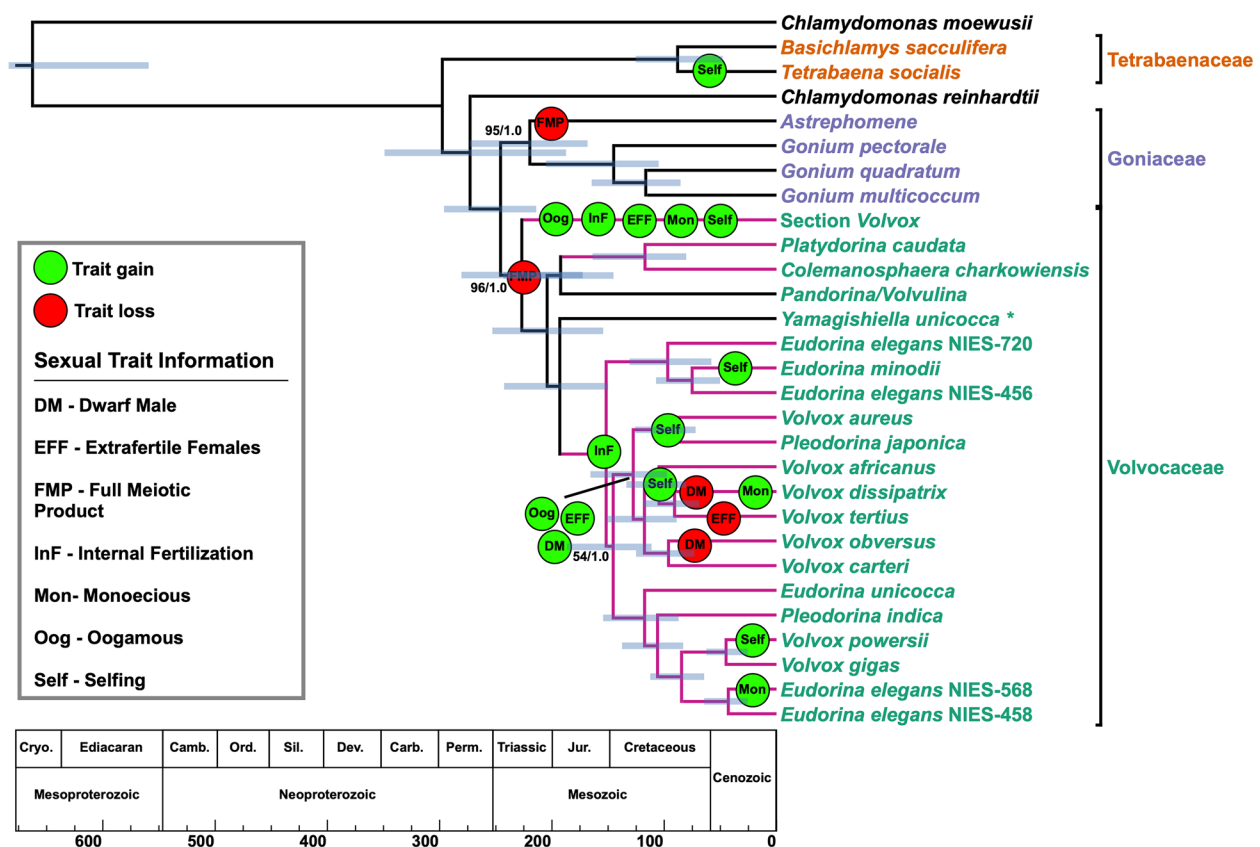
Fig. 4. Sexual trait gain and loss in the volvocine algae. Branching order is of a collapsed tree inferred under maximum-likelihood analysis from an aligned amino acid, concatenated dataset of 263 nuclear genes. Numbers on branches represent bootstrap and posterior probability values, respectively. Branch lengths, corresponding to time, were inferred under the CIR relaxed clock model using 16 most clock-like genes as determined by Sortadate. Blue bars correspond to the inferred 95% HPD interval for each node. Green bubbles correspond to a sexual trait gain, and red bubbles corresponds to loss of a trait. The figure table lists the 7 mapped sexual traits and their acronyms. Pink branches denote lineages where anisogamy evolved. Members of the multicellular volvocine algae (Tetrabaenaceae, Goniaceae, and Volvocaceae) are denoted in orange, purple, and green. Taxa in black font are unicellular volvocine algae

The most significant difference between our findings and those of Hanschen et al. [32] is the number of times we find anisogamy to have evolved independently. Their results inferred two independent origins of anisogamy from isogamous ancestors, whereas we conclude three, consistent with Lindsey et al. [11] (Fig. 4 and Additional file 1: Figures S32 and 33) [11, 32]. Section *Volvox*'s position as sister to all other Volvocaceans underlies this distinction. The three origins of anisogamy in our analysis occur in the ancestors of the following lineages: (*i*) section *Volvox* (136-57 MYA), (*ii*) *Platydorina* + *Colemanosphaera* (113-46 MYA) and (*iii*) the *Eudorina* + *Volvox* + *Pleodorina* (EVP) clade (196-117 MYA) (Fig. 4 and Additional file 1: Figure S32 and 33). It is noteworthy that we infer three independent origins of anisogamy, regardless of whether this trait is treated as a 2-state problem (e.g., isogamy and anisogamy) or a 3-state problem (e.g., isogamy, anisogamy, or oogamy) (Additional file 1: Figures S32-35).

Consistent with our mapping of developmental traits, anisogamy evolved in lineages that had already acquired a diverse suite of characters differing from a unicellular ancestor. In each case, the eventual anisogamous ancestor had evolved a spheroidal body plan with increased colony size, polarity, and in the case of section *Volvox*, may also have acquired true germ-soma differentiation. It is possible that in some anisogamous ancestors, the body plan was larger with distinct cell types compared to a *Pandorina*-like or *Volvulina*-like morphology, as hypothesized by Hanschen et al. [32] In section *Volvox* and the EVP clade, internal fertilization was gained around the same time (sometime after ~228 MYA (279-175 MYA) for section *Volvox* and ~152 MYA (196-117 MYA) for the EVP clade) as anisogamy, suggesting that anisogamy might be a prerequisite for this trait (Fig. 4 and Additional file 1: Figures S36 and 37).

After gaining anisogamy and internal fertilization, oogamous lineages appear rapidly to gain several other

sexual traits (Fig. 4). This suggests that oogamy, an extreme form of anisogamy, may drive further sexual evolution. By ~95 MYA (136-57 MYA), section *Volvox*'s ancestor had become monoecious (i.e., colonies could produce both sperm and ova) and capable of selfing. Necessarily, such an ancestor had lost dioecy (i.e., colonies able to produce only sperm or ova) as well as the capacity to outcross (Additional file 1: Figures S38-41). By the same time, this ancestor gained a sexual dimorphism in the form of "extrafertile" females (Fig. 4). During sexual reproduction, an "extrafertile" female *Volvox* colony exhibits a doubling in the number of eggs inside its colony, thus giving sexual reproduction the potential to numerically outperform an asexual reproductive cycle.

MrBayes and PhyTools ASR results indicate different ancestral states for when the "dwarf" male phenotype was gained. Results from MrBayes ASR (MBASR) indicate "dwarf" males were gained independently in the lineages leading to *V. africanus*, *V. tertius*, and in the common ancestor to *V. carteri* f. *nagariensis* and *V. carteri* f. weismannia (Additional file 1: Figure S44). Phytools ASR, however, indicate that the common ancestor to *V. carteri* and *V. africanus* had evolved the "dwarf" male phenotype with two losses occurring in the lineages leading to *V. obversus* and *V. dissipatrix* (Additional file 1: Figure S45). These differences arise due to model selection. MBASR results are replicated in Phytools if the "equal rates" model is exclusively used (Additional file 1: Figure S46). When an ANOVA test is performed in Phytools comparing the "equal rates" and "all-rates-different" models, the "all-rates-different" model has a markedly higher weight (0.70 vs 0.29) indicating that it is highly favored over the "equal rates" model. Therefore, in the common ancestor to *V. carteri* and *V. africanus*, two sexual dimorphisms, "extrafertile" females and "dwarf" males, were gained by ~118 MYA (152-90 MYA) (Fig. 4 and Additional file 1: Figures S42-43 and 45). The "dwarf" male phenotype is expressed as an exaggeratedly reduced male colony size when compared to female colonies. Within this same clade, *V. africanus* and *V. dissipatrix* independently evolved monoecious, selfing colonies, in addition to the previously mentioned sexual dimorphisms. Sexual trait gains occur in various other lineages of the volvocine algae, but none are as clear as the ones occurring in the two oogamous lineages of *Volvox*.

In summary, we conclude that multicellularity was a likely driver to the evolution of anisogamy. Once anisogamy is established, a variety of other sexual traits may arise, notably the evolution of oogamy, which in turn opens the door for more elaborate forms of sexual dimorphism such as extrafertile females and dwarf males. These derived sexual traits are positively associated with traits related to the degree of multicellular complexity, such as increased cell number, expanded extracellular matrix, organismal polarity, inversion, and cellular differentiation [32, 33]. Thus, our more robust phylogeny of the volvocine algae, based on not 5 but on 263 protein coding loci, confirms previous findings [31, 32] that multicellularity ultimately drives the evolution and elaboration of diverse sexual traits.

We would be remiss if anisogamous/oogamous unicellular chlamydomonads outside the volvocine algae were not briefly discussed. Using morphological and SSU rRNA data, Pröschold and colleagues have proposed placing oogamous *Chlamydomonas* strains in two other genera: *Oogamochlamys* and *Chloromonas* [152]. In their taxonomic revision, *Chlamydomonas reinhardtii* is proposed as the type species of the genus [152]. Pröschold et al. [152] contend that because these other *Chlamydomonas* species are not monophyletic with *C. reinhardtii* (i.e., do not fall within the Reinhardtinia clade), they should be reclassified from *Chlamydomonas* to other genera [152]. Thus, while we recognize that anisogamy and oogamy exist in unicellular genera outside the volvocine algae, discussion of these taxa is beyond the scope of our investigation.

Molecular clock analysis also suggests cryptic volvocine species in the Tetrabaenaceae. Much like the longstanding debate on what constitutes a biological species [155, 156], consensus is lacking for how we should recognize cryptic species [157]. Most simply, cryptic species are those that have been binned into a single taxon at the species rank. Discriminating cryptic species can enlarge estimates of biodiversity and provide insight into the mechanisms that underlie adaptation and speciation. Following suggestions and criticisms offered by Struck *et al.* [158], we provide here divergence times and molecular evidence that point to the existence of cryptic species within the genus *Tetrabaena*. This insight is reinforced by previously published biogeographic [82, 159], phenotypic [159], and genetic data [160] on this enigmatic genus.

Tetrabaena is a monotypic, colonial volvocine alga and one of two genera composing the multicellular Tetrabanaeceae. Several strains of this genus have been described as far back as 1841 [161]; all described *Tetrabaena* spp. have been of freshwater origin except one Antarctic strain, NIES-691. Following collection of the Antarctic *T. socialis* from King George Island [162], Nozaki and Ohtani [159] provided morphological data and temperature growth profiles that demonstrated phenotypic idiosyncrasies of this strain in comparison to the freshwater NIES-571. Phenotypic traits distinctive to NIES-691 included: (*i*) vegetative colonies measuring up to 50 μm in diameter (~18 μm larger than NIES-571), (*ii*) immobile colonies that adhere to surfaces at 20 °C, and (*iii*) failure to grow at 25 °C. In contrast to NIES-691, NIES-571 was

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 14 of 21

observed to produce smaller colonies and to grow normally at all tested temperatures ranging from 5-25 °C, with no phenotypic changes that produce immobile colonies or ones adhering to surface s[159]. These phenotypic differences, which may be local environmental adaptation, point towards the possibility of a cryptic species in *Tetrabaena*, arguing for the examination of detailed molecular data on these two strains.

In 1997, Mai and Coleman [160] produced complete Internal Transcribed Spacer 2 (ITS-2) sequences for 111 Volvocales. including *Tetrabaena socialis* strains NIES-571 and NIES-691. When these two *T. socialis* ITS-2 sequences are subjected to pairwise alignment, the percent identity between them is 92.47% (Additional file 2: Table S3). Similar to the percent identity (91.25%) between ITS-2 sequences of *V. gigas* UTEX 1895 and *V. powersii* UTEX 1864, two sister taxa recognized as separate species, when these are subject to pairwise alignment (Additional file 2: Table S3). When the same test is applied to ITS-2 sequences of *Volvox carteri f. nagariensis* strains (Poona and 72-52), the percent identity between them is 98.06% (Additional file 2: Table S3). Considered in relation to previously reported phenotypic data, these molecular differences lead us to hypothesize that NIES-571 and NIES-691 are cryptic species within the genus *Tetrabaena*.

This hypothesis is bolstered by our clock analysis and molecular data. The two strains appear to have diverged from one another ~51 MYA (82-25 MYA) (Fig. 3). This divergence time estimate stands in contrast to that estimated between strains of *Volvox carteri* f. *nagariensis* (NIES-865 and HK10) as well as to that estimated between strains of *Pleodorina starrii* (NIES-1362 and NIES-1363). In each of these two examples, strains are thought to have diverged from one another <1 MYA (Fig. 3). And in both cases, there is good reason to believe that each strain pair is a biological species *sensu* Mayr[163]. *V. carteri* NIES-865 is an isolate of *V. carteri* HK10[164, 165], and mating experiments [166] between *Pleodorina starrii* strains (NIES-1362 and NIES-1363) produce fertile offspring. Exploration of conserved regions from our 16 amino acid sequence alignments used in our molecular clock analyses further highlight differences in sequence percent identity. Pairwise percent identity between *V. carteri* and *P. starrii* strains was 100% and 99.65-100%, respectively, whereas the percent identity between each set of *T. socialis* sequences ranges between 88-99.27% indicating substantial evolutionary differences in the protein sequences (Additional file 2: Table S4).

Given the divergence times indicated by our dataset, and the phenotypic and molecular features known to distinguish these *T. socialis* strains, we propose recognizing *Tetrabaena socialis* NIES-571 as a cryptic species

to clarify *socialis* N-691's current status as the type species. Formal elevation of NIES-571 to species status will require additional data. For example, complete plastid sequences should be obtained to complement the pairwise alignments of nuclear genes provided here. Because chloroplast genomes are known to be highly conserved [167], marked differences between plastid sequences of these two *T. socialis* strains would strengthen the case for one to be recognized as a cryptic species. Also, NIES-571 and NIES-691 should be crossed. And, even if mating experiments result in viable F1 progeny, additional crosses should be undertaken to assess the fertility of F1 hybrids in relation to their parents and to one another.

Conclusions

Through the use of a more comprehensive molecular and taxonomic dataset, we have confirmed the major findings of Lindsey et al.'s phylogeny [11], and further supported them via ancestral-state reconstructions using fossil-calibrated molecular clock analyses. We find that the origin of the crown Tetrabaenaceae-Goniaceae-Volvocaceae (TGV) clade may be much older than previously thought, possibly emerging as early as the Carboniferous Period. We also find that multicellularity in the Goniaceae + Volvocaceae evolved sometime between the Carboniferous-Triassic Periods, and that multicellularity in Tetrabaenaceae may have evolved as recently as the Cretaceous Period. Given the wide range of time in our 95% HPD intervals for volvocine divergence times (e.g., second independent origin of multicellularity occurring between 297-190 MYA), we hesitate to speculate about specific geological and ecological changes that may have attended the evolution of multicellularity in this clade. However, even a conservative view of our divergence times shows that volvocine diversification occurred in the context of ecosystem reorganization in the wake of end-Permian mass extinction [168]. Armed with our time-tree and ancestral-state reconstructions, we reassessed Kirk's 12-Step Program to differentiated multicellularity using a new, more robust phylogeny of the volvocine algae; we determined that the essential elements of his hypothesis to be correct. Lastly, our data point to the possibility of a cryptic species in the multicellular Tetrabaenaceae, a possibility that could change its monotypic status, and provide an exciting opportunity to investigate the genomics of two diverged species that have adapted to very different environments.

In ancestral character state reconstructions, it is nearly always the case that multiple, nearly equally likely reconstructions exist, and this limits the confidence we can have in the specifics of the (marginally) most parsimonious. For example, in the clade at the bottom of Fig. 4, which includes *P. indica, P. starrii, V. powersii, V. gigas,*

and some *E. elegans* strains, our reconstruction shows a gain of sterile soma at the base of the clade and a loss in *E. elegans*. However, it would be only slightly less parsimonious to show independent gains in *P. indica* + *P. starrii* and in *V. powersii* + *V. gigas*, with no loss. There is, however, no way to reconstruct the history of this trait that doesn't involve several independent gains and/or losses.

The distinction between mortal somatic cells, which do not pass on genetic information to subsequent generations, and potentially immortal germ cells, which do, has been recognized as a fundamentally important evolutionary trait as far back as Weismann's germ-plasm theory [169]. The evolution of anisogamy is the first step in the differentiation of male and female sexes, setting the stage for sexual selection to produce all of the spectacular sexual dimorphisms we see throughout the animal kingdom [14, 170, 171] as well as the more modest dimorphisms in some species of *Volvox* (dwarf males and extrafertile females). Although nearly equally likely reconstructions preclude a high degree of confidence in the details of these traits' histories, this much is clear: given the long-recognized fundamental evolutionary significance of these traits, they are surprisingly evolutionarily labile, with multiple independent losses and/or gains, within the relatively recently-evolved multicellular volvocine algae.

Methods

Command-line arguments and software repository links for each program have been uploaded to a single file that may be downloaded from our Dryad repository (DOI: 10.5061/dryad.mcvdnck6b) [172].

Retrieval of genomic and RNA-Seq data for Archaeplastida phylogeny. A total of 164 taxa were sampled across Archaeplastida based upon publicly available genomic and RNA-Seq data. Specifically, longest primary transcript files were downloaded for 20 taxa from the Phytozome database, and protein files for 11 taxa were downloaded from Ensembl, NCBI, OrcAE, and Tokyo Institute of Technology. RNA-Seq data was downloaded for 133 taxa, and *de novo* transcriptomes were assembled for each taxon. Accession numbers and/or database information for each strain may be found in Additional file 2: Table S1.

Quality control of RNA-Seq reads. Quality of all RNA-Seqs was initially checked using FastQC v.0.11.8 with an additional FastQC assessment post-trimming. Quality trimming was conducted with Trimmomatic v.0.3 9[135]. Command line parameters and details for Trimmomatic in addition to adapter file information may be found in Additional file 3: Supplementary Methods.

De novo transcriptome assembly. Quality filtered, paired-end reads were used to assemble *de novo*

transcriptomes with SOAPdenovo-Trans v1.0.4 [173] using k-mer size of 25, GapCloser was used to close gaps in each *de novo* transcriptome, and CDHIT v4.8.1 [174] was used under default parameters to cluster redundant transcripts from each assembled transcriptome.

Annotation of de novo transcripts. Each assembled transcript had its longest open reading frame predicted using TransDecoder v5.5.0. Diamond [136] BLASTP and HMMER [137] protein domain identification searches were conducted using each .pep file generated in the previous step. Using the results of Diamond and HMMER, nucleotide CDS and amino acid PEP files were then generated by Transdecoder v5.5.0. Detailed information for Diamond BLASTP and HMMER searches may be found in Additional file 3: Supplementary Methods.

Orthologous gene identification infers 284 single-copy genes shared among red and green algae. A training set of 10 Archaeplastida taxa (1 rhodophyte, 5 streptophytes, and 4 chlorophytes) were chosen from the Phytozome database to infer a putative list of single-copy genes present across Rhodophyta, Streptophyta, and Chlorophyta (Additional file 2: Table S5). The primary transcript file for each taxon in our training set was used to determine a putative list of single-copy genes using Orthofinder v2.3.8 [175]. Orthofinder returned 284 orthologous groups where all species were present and all genes single-copy. These 284 orthologous genes across all species were aligned using MAFFT v7.487 [176]. HMMER was subsequently utilized to build profile Hidden Markov Model (HMM) files for each gene alignment. These profile HMMs were then used by Orthofisher [177] to mine our genomes and *de novo* transcriptomes for those single-copy genes. Putative organellar sequences from the inferred single-copy dataset were ruled out, and detailing about this process can be found in Additional file 3: Supplementary Methods.

Gene sequence alignments. Once the amino acid sequences were confirmed to be of nuclear origin and single-copy, they were aligned using MUSCLE v3.8.31 [178], resulting in 284 untrimmed gene alignments. Furthermore, each alignment was manually inspected and aligned in Aliview v1.26 [179]. Further details are found in Additional file 3: Supplementary Methods. Following manual quality control of alignments, ambiguously aligned regions from each alignment were trimmed using Trimal v1.4.1 [180] allowing for only conserved and reliably aligned regions to later undergo phylogenetic analysis. A total of 21 gene alignments from the original 284 alignments were discarded and not used in any downstream analyses. These 21 gene alignments were discarded due to poor sequence alignment and/or many missing sequences.

Parameters for Phylogenetic analyses. Each single-gene alignment had its best evolutionary model predicted by ProtTest v3.4.2 [181] under the Akaike Information Criterion (AIC). Each predicted evolutionary model and its associated alignment may be found in Additional file 2: Table S6. All single-gene alignments underwent Maximum-Likelihood (ML) analysis using IQtre e[182]. Phyutility v2.7.1 [183] was utilized to concatenate all gene alignment files for ML and BI analyses. IQtree was used to run a concatenated, gene-partitioned ML analysis with 1000 rapid bootstrap replicates [184]. The gene-partition strategy was based upon the best evolutionary model predicted by ProtTest. The Bayesian inference (BI) analysis was conducted with MrBayes 3.2.7a [142] under the same gene-partitioning strategy previously described. A total of 4 runs each with 4 Markov chains (3 heated and 1 cold). Trees were sampled every 1000 generations over 1,000,000 generations during which convergence (average standard deviation of split frequencies = <0.01) between runs was reached, and a burn-in of 25% was used (ngen=1000000 nruns=4 samplefreq=1000 nchains=4 starttree=random). Lastly, ASTRAL [185] was utilized to perform a coalescent-based analysis using the 263 single-gene trees produced by IQtree. Each single-gene tree had branches with low bootstrap support (<10) collapsed using Newick Utilities v1.6 [186].

Parameters for molecular clock analyses. For all clock analyses, Sortadate [138] was utilized to assemble concatenated datasets of 8 and 16 genes to offset high computational cost. For each of the smaller datasets, genes were ranked with an emphasis on their Sortadate bipartition score with the scores ranging from: 48-56% (8 genes) and 45-56% (16 genes) (Additional file 2: Table S7). Additional details concerning Sortadate may be found in Additional file 3: Supplementary Methods.

Phylobayes 4.1b [92] was used to perform relaxed molecular clock analyses on each dataset under the CIR (-cir), lognormal autocorrelated (-ln), uncorrelated gamma (-ugam), and white-noise (-wn) models with the LG protein substitution model applied (-lg), discrete gamma rate across sites (-dgam 4), and birth-death prior on divergence times (-bd). The red algae were specified as the outgroup taxa according to manual instructions (-r <outgroup.txt>). For our main analyses a total of 14 nodes were calibrated using fossil taxa with either a minimum bound or minimum + maximum bounds (Fig. 2) with the soft bounds flag (-sb) allowing date sampling to occur slightly outside of the specified bounds (5% on a pure minimum date, and 2.5% on lower and upper bounds). No secondary node calibrations were used in this study. The root age (red and green algal divergence) was set to a maximum bound of 2000 MY. Separate runs for our maximum-likelihood and coalescent-based

species tree topologies were conducted. A total of two MCMC chains were conducted per main run for a minimum of 10,000 generations, and the convergence of the two chains was determined in Phylobayes. A summarization of the chains was done after discarding the first 2000 generations (~20%) as burn-in.

Ancestral state reconstructions of discrete volvocine characters Two programs were used to estimate the probability of ancestral character states at internal nodes in the volvocine algae: the Phytools [139] R package and MrBayes [142] via the MBASR toolkit [141]. For each of the two programs, our tree topology was constrained to the maximum-likelihood, molecular-clock tree inferred using 16 protein-coding genes under the CIR model. For each ASR method, the following discrete character states were ancestrally reconstructed in the volvocine algae: (*i*) cellularity, (*ii*) spheroidal body plan, (*iii*) developmental traits identified by Kirk in his 12-Step Program [33], and (*iv*) 7 sexual traits from two previous studies [31, 32]. To estimate the probabilities of ancestral character states, stochastic character mapping ("simmap") was applied using Phytools. R scripts used to run each simmap analysis may be found at Dryad [172]. In MBASR, a Markov chain Monte Carlo (MCMC) simulation was conducted for 1,000,000 generations (MBASR n.samples=10,000) with a sample taken every 100 generations to estimate ancestral character state probabilities. Additionally, information for runs conducted in Phytools and MBASR may be found in Additional file 3: Supplementary Methods.

Abbreviations

ML	Maximum-likelihood
BI	Bayesian inference
CB	Coalescent-based
MLBS	Maximum-likelihood bootstrap support
BPP	Bayesian posterior probablility
CPP	Coalescent posterior probability
EVP	Eudorina + Volvox + Pleodorina
PVC	Pandorina + Volvulina + Colemanosphaera
LN	Autocorrelated lognormal
CIR	Cox-Ingersoll-Ross
UGAM	Uncorrelated gamma
WN	White-noise
MY	Million years
MYA	Million years ago
HPD	Highest posterior density
TGV	Tetrabaenanceae-goniaceae-volvocaceae
ITS-2	Internal transcribed spacer 2

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-024-01878-1.

Additional File 1: **Figures S1-S46**.

Additional File 2: **Tables S1-S7**.

Additional File 3: Supplementary Methods.

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 17 of 21

Data collected from plants
This study used previously published genomic and transcriptomic data from plant and algal organisms. We did not conduct experimental research or field studies on plants or algae during this project. Additionally, at no time did we collect or interact with plant or seed specimens (cultivated or wild). Therefore, no voucher specimens need to be deposited in a public collection or public herbarium.

Conflict of Interest Statement
The authors declare that they have no competing interests.

Authors' Contributions
Performed the experiments (CRL), carried out the analyses (CRL, MDH), wrote the manuscript (CRL, AHK, MDH, RFR). All authors have read and approved the final manuscript.

Availability of data and materials
No new raw genomic or transcriptomic data were generated for this study. All previously published sequence data were obtained through online data repositories such as: Phytozome, Ensembl, NCBI, OrcAE, and Tokyo Institute of Technology. Accession numbers and/or database information for sequence data may be found in Additional file S2: Table S1. A Dryad repository (https://doi.org/10.5061/dryad.mcvdnck6b) [172] has been established for the preservation of several file types. Within the Dryad database, all R scripts used to run our simmap analyses in Phytools will be found along with a MBASR template script. Also provided are the 263 gene alignment files that were concatenated in our maximum-likelihood and MrBayes phylogenetic analyses. We uploaded four concatenated files to Dryad: (*i*) 8-gene alignment file used in Phylobayes, (*ii*) 16-gene alignment file used in Phylobayes, (*iii*) 263-gene alignment file utilized in IQtree and our MrBayes phylogenetic analyses, and (*iv*) ASTRAL 263-gene tree file. We also uploaded our fossil calibration file and the following Phylobayes output files for each clock model analysis for the 16-gene dataset: chronogram, dates, and labels. Lastly, a command line PDF with software repository links was uploaded to Dryad for each program used in this study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA. [2]Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford St., Cambridge, MA 02138, USA. [3]Parker H. Petit Institute for Bioengineering and Biosciences, Atlanta, GA 30332, USA.

References

1. Maynard Smith J, Szathmáry E. The major transitions in evolution. Oxford University Press; 1995.
2. Schopf JW, Kudryavtsev AB, Agresti DG, Wdowiak TJ, Czaja AD. Laser–Raman imagery of Earth's earliest fossils. Nature. 2002;416:73–6.
3. Lamża Ł. Diversity of 'simple' multicellular eukaryotes: 45 independent cases and six types of multicellularity. Biol Rev. 2023;brv.13001.
4. Schirrmeister BE, de Vos JM, Antonelli A, Bagheri HC. Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. Proc Natl Acad Sci. 2013;110:1791–6.
5. Nickell S, Hegerl R, Baumeister W, Rachel R. Pyrodictium cannulae enter the periplasmic space but do not enter the cytoplasm, as revealed by cryo-electron tomography. J Struct Biol. 2003;141:34–42.
6. Ratcliff WC, Denison RF, Borrello M, Travisano M. Experimental evolution of multicellularity. Proc Natl Acad Sci. 2012;109:1595–600.
7. Olson BJ, Nedelcu AM. Co-option during the evolution of multicellular and developmental complexity in the volvocine green algae. Curr Opin Genet Dev. 2016;39:107–15.
8. Herron MD, Rashidi A, Shelton DE, Driscoll WW. Cellular differentiation and individuality in the 'minor' multicellular taxa: Differentiation and individuality. Biol Rev. 2013;88:844–61.
9. Umen J, Herron MD. Green Algal Models for Multicellularity. Annu Rev Genet. 2021;55:603–32.
10. Jiménez-Marín B, Olson BJSC. The Curious Case of Multicellularity in the Volvocine Algae. Front Genet. 2022;13:787665.
11. Lindsey C, Rosenzweig F, Herron MD. Phylotranscriptomics points to multiple independent origins of multicellularity and cellular differentiation in the volvocine algae. BMC Biol. 2021;19.
12. Ma X, Shi X, Wang Q, Zhao M, Zhang Z, Zhong B. A reinvestigation of multiple independent evolution and Triassic-Jurassic origin of multicellular Volvocine algae. Genome Biol Evol. 2023;evad142.
13. Hanschen ER, Davison DR, Grochau-Wright ZI, Michod RE. Evolution of Individuality: A Case Study in the Volvocine Green Algae. Philos Theory Pract Biol. 2017;9 https://doi.org/10.3998/ptb.6959004.0009.003.
14. Lehtonen J, Kokko H, Parker GA. What do isogamous organisms teach us about sex and the two sexes? Philos Trans R Soc B. 2016;371:20150532.
15. Bachtrog D, Mank JE, Peichel CL, et al. Sex determination: Why so many ways of doing it? PLoS Biol. 2014;12:e1001899.
16. Powers JH. New forms of Volvox. Trans Am Microsc Soc. 1907;27:123–49.
17. Kirk DL. Volvox: a search for the molecular and genetic origins of multicellularity and cellular differentiation. Cambridge: Cambridge University Press; 1998.
18. Herron MD, Hackett JD, Aylward FO, Michod RE. Triassic origin and early radiation of multicellular volvocine algae. Proc Natl Acad Sci. 2009;106:3254–8.
19. Heffel MG. Pandorina morum genome assembly, annotation, and analysis. Kansas State University, Doctoral Dissertation. 2020. https://krex.k-state.edu/server/api/core/bitstreams/99ddaa44-f40a-4ffa-8707-b44b4b79b39d/content.
20. Hanschen ER, Marriage TN, Ferris PJ, et al. The Gonium pectorale genome demonstrates co-option of cell cycle regulation during the evolution of multicellularity. Nat Commun. 2016;7:1–10.
21. Featherston J, Arakaki Y, Hanschen ER, Ferris PJ, Michod RE, Olson BJSC, Nozaki H, Durand PM. The 4-Celled Tetrabaena socialis Nuclear Genome Reveals the Essential Components for Genetic Control of Cell Number at the Origin of Multicellularity in the Volvocine Lineage. Mol Biol Evol. 2018;35:855–70.
22. Prochnik SE, Umen J, Nedelcu AM, et al. Genomic Analysis of Organismal Complexity in the Multicellular Green Alga Volvox carteri. Science. 2010;329:223–6.
23. Yamashita S, Yamamoto K, Matsuzaki R, et al. Genome sequencing of the multicellular alga Astrephomene provides insights into convergent evolution of germ-soma differentiation. Sci Rep. 2021;11:22231.
24. Hu Y, Xing W, Song H, Hu Z, Liu G. Comparison of colonial volvocine algae based on phylotranscriptomic analysis of gene family evolution and natural selection. Eur J Phycol. 2020;55:100–12.
25. Nozaki H, Yamada TK, Takahashi F, Matsuzaki R, Nakada T. New "missing link" genus of the colonial volvocine green algae gives insights into the evolution of oogamy. BMC Evol Biol. 2014;14:1–11.

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 18 of 21

26. Tang Q, Pang K, Yuan X, Xiao S. A one-billion-year-old multicellular chlorophyte. Nat Ecol Evol. 2020;4:543–9.

27. Gibson TM, Shih PM, Cumming VM, et al. Precise age of Bangiomorpha pubescens dates the origin of eukaryotic photosynthesis. Geology. 2018;46:135–8.

28. 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 https://doi.org/10.1371/journal.pbio.2000735.

29. Herron MD, Desnitskiy AG, Michod RE. Evolution of developmental programs in Volvox (Chlorophyta). J Phycol. 2010;46:316–24.

30. Grochau-Wright ZI, Hanschen ER, Ferris PJ, Hamaji T, Nozaki H, Olson BJSC, Michod RE. Genetic basis for soma is present in undifferentiated volvocine green algae. J Evol Biol. 2017;30:1205–18.

31. Hanschen ER, Herron MD, Wiens JJ, Nozaki H, Michod RE. Repeated evolution and reversibility of self-fertilization in the volvocine green algae*. Evolution. 2018;72:386–98.

32. Hanschen ER, Herron MD, Wiens JJ, Nozaki H, Michod RE. Multicellularity drives the evolution of sexual traits. Am Nat. 2018;192:E93–105.

33. Kirk DL. A twelve-step program for evolving multicellularity and a division of labor. Bioessays. 2005;27:299–310.

34. Bhattacharya D, Yoon HS, Hackett JD. Photosynthetic eukaryotes unite: endosymbiosis connects the dots. Bioessays. 2004;26:50–60.

35. Reyes-Prieto A, Weber APM, Bhattacharya D. The origin and establishment of the plastid in algae and plants. Annu Rev Genet. 2007;41:147–68.

36. Wodniok S, Brinkmann H, Glöckner G, Heidel AJ, Philippe H, Melkonian M, Becker B. Origin of land plants: Do conjugating green algae hold the key? BMC Evol Biol. 2011;11:104.

37. Finet C, Timme RE, Delwiche CF, Marlétaz F. Multigene phylogeny of the green lineage reveals the origin and diversification of land plants. Curr Biol. 2012;22:1456–7.

38. Turmel M, de Cambiaire J-C, Otis C, Lemieux C. Distinctive architecture of the chloroplast genome in the chlorodendrophycean green algae *Scherffelia dubia* and *Tetraselmis* sp. CCMP 881. PLoS One. 2016;11:e0148934.

39. Matsuzaki M, Misumi O, Shin-i T, et al. Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. Nature. 2004;428:653–7.

40. Nozaki H, Takano H, Misumi O, et al. A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga Cyanidioschyzon merolae. BMC Biol. 2007;5:28.

41. Liu S, Chiang Y, Yoon HS, Fu H. Comparative Genome Analysis Reveals Cyanidiococcus gen. nov., A New Extremophilic Red Algal Genus Sister to Cyanidioschyzon (Cyanidioschyzonaceae, Rhodophyta). J Phycol. 2020;56:1428–42.

42. Brawley SH, Blouin NA, Ficko-Blean E, et al. Insights into the red algae and eukaryotic evolution from the genome of Porphyra umbilicalis (Bangiophyceae, Rhodophyta). Proc Natl Acad Sci USA. 2017;114 https://doi.org/10.1073/pnas.1703088114.

43. One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. Nature. 2019;574:679–85.

44. Collén J, Porcel B, Carré W, et al. Genome structure and metabolic features in the red seaweed Chondrus crispus shed light on evolution of the Archaeplastida. Proc Natl Acad Sci USA. 2013;110:5247–52.

45. Janouškovec J, Liu S-L, Martone PT, Carré W, Leblanc C, Collén J, Keeling PJ. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. PLoS One. 2013;8:e59001.

46. Hori K, Maruyama F, Fujisawa T, et al. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. Nat Commun. 2014;5:3978.

47. University of Maryland College Park. Nitella mirabilis transcriptome. NCBI BioProject PRJNA158153. https://www.ncbi.nlm.nih.gov/bioproject/158153.

48. Nishiyama T, Sakayama H, De Vries J, et al. The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. Cell. 2018;174:448-464.e24.

49. De Vries J, Curtis BA, Gould SB, Archibald JM. Embryophyte stress signaling evolved in the algal progenitors of land plants. Proc Natl Acad Sci USA. 2018;115 https://doi.org/10.1073/pnas.1719230115.

50. University of Maryland College Park. Coleochaete orbicularis transcriptome. NCBI BioProject PRJNA734590. https://www.ncbi.nlm.nih.gov/bioproject/734590.

51. Bowman JL, Kohchi T, Yamato KT, et al. Insights into land plant evolution garnered from the marchantia polymorpha genome. Cell. 2017;171:287-304.e15.

52. Banks JA, Nishiyama T, Hasebe M, et al. The selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science. 2011;332:960–3.

53. Qi X, Kuo L-Y, Guo C, et al. A well-resolved fern nuclear phylogeny reveals the evolution history of numerous transcription factor families. Mol Phylogenet Evol. 2018;127:961–77.

54. Shalev TJ, Gamal El-Dien O, Yuen MMS, et al. The western redcedar genome reveals low genetic diversity in a self-compatible conifer. Genome Res. 2022;32 genome;gr.276358.121v2.

55. Jin W-T, Gernandt DS, Wehenkel C, Xia X-M, Wei X-X, Wang X-Q. Phylogenomic and ecological analyses reveal the spatiotemporal evolution of global pines. Proc Natl Acad Sci USA. 2021;118:e2022302118.

56. Amborella Genome Project, Albert VA, Barbazuk WB, et al. The *Amborella* genome and the evolution of flowering plants. Science. 2013;342:1241089.

57. Zhang L, Chen F, Zhang X, et al. The water lily genome and the early evolution of flowering plants. Nature. 2020;577:79–84.

58. Chaw S-M, Liu Y-C, Wu Y-W, et al. Stout camphor tree genome fills gaps in understanding of flowering plant genome evolution. Nat Plants. 2019;5:63–73.

59. Lamesch P, Berardini TZ, Li D, et al. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 2012;40:D1202–10.

60. Slotte T, Hazzouri KM, Ågren JA, et al. The Capsella rubella genome and the genomic consequences of rapid mating system evolution. Nat Genet. 2013;45:831–5.

61. Yang R, Jarvis DE, Chen H, et al. The reference genome of the halophytic plant *Eutrema salsugineum*. Front Plant Sci. 2013;4 https://doi.org/10.3389/fpls.2013.00046.

62. JGI. Brassica rapa (field mustard) genome. NCBI BioProject PRJNA41163. 2018. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA41163.

63. The International Brachypodium Initiative. Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature. 2010;463:763–8.

64. Ouyang S, Zhu W, Hamilton J, et al. The TIGR Rice Genome Annotation Resource: improvements and new features. Nucleic Acids Res. 2007;35:D883–7.

65. McCormick RF, Truong SK, Sreedasyam A, et al. The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant J. 2018;93:338–54.

66. JGI. Zea mays refGen_V4. Phytozome genome ID: 493. 2016. https://phytozome-next.jgi.doe.gov/info/Zmays_RefGen_V4.

67. Moreau H, Verhelst B, Couloux A, et al. Gene functionalities and genome structure in Bathycoccus prasinos reflect cellular specializations at the base of the green lineage. Genome Biol. 2012;13:R74.

68. Vaulot D, Lepère C, Toulza E, et al. Metagenomes of the picoalga *Bathycoccus* from the chile coastal upwelling. PLoS One. 2012;7:e39648.

69. Blanc-Mathieu R, Verhelst B, Derelle E, et al. An improved genome of the model marine alga Ostreococcus tauri unfolds by assessing Illumina de novo assemblies. BMC Genomics. 2014;15:1103.

70. Worden AZ, Lee J-H, Mock T, et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes micromonas. Science. 2009;324:268–72.

71. University of Maryland College Park. Tetraselmis striata transcriptome. NCBI BioProject PRJNA734590. https://www.ncbi.nlm.nih.gov/sra/?term=srs10979578.

72. Blanc G, Agarkova I, Grimwood J, et al. The genome of the polar eukaryotic microalga Coccomyxa subellipsoidea reveals traits of cold adaptation. Genome Biol. 2012;13:R39.

73. Blanc G, Duncan G, Agarkova I, et al. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. Plant Cell. 2010;22:2943–55.

74. De Clerck O, Kao S-M, Bogaert KA, et al. Insights into the evolution of multicellularity from the sea lettuce genome. Curr Biol. 2018;28:2921-2933.e5.

75. 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 (Ulvophyceae). Nat Commun. 2022;13:1610.

76. University of Maryland College Park. Oltmannsiellopsis unicellularis transcriptome. NCBI BioProject PRJNA734590. https://www.ncbi.nlm.nih.gov/sra/?term=srs10979569.

77. Arimoto A, Nishitsuji K, Narisoko H, Shoguchi E, Satoh N. Differential gene expression in fronds and stolons of the siphonous macroalga, Caulerpa lentillifera. Develop Growth Differ. 2019;61:475–84.

78. Ranjan A, Townsley BT, Ichihashi Y, Sinha NR, Chitwood DH. An intracellular transcriptomic atlas of the giant coenocyte *Caulerpa taxifolia*. PLoS Genet. 2015;11:e1004900.

79. Ünlü ES, Ünüvar ÖC, Aydın M. Identification of alternative oxidase encoding genes in Caulerpa cylindracea by de novo RNA-Seq assembly analysis. Mar Genomics. 2019;46:41–8.

80. Calixto Mancipe N, McLaughlin EM, Barney BM. Genomic analysis and characterization of Scenedesmus glucoliberatum PABB004: An unconventional sugar-secreting green alga. J Appl Microbiol. 2022;132:2004–19.

81. Roth MS, Cokus SJ, Gallaher SD, et al. Chromosome-level genome assembly and transcriptome of the green alga Chromochloris zofingiensis illuminates astaxanthin production. Proc Natl Acad Sci USA. 2017;114 https://doi.org/10.1073/pnas.1619928114.

82. Zhang Z, Qu C, Yao R, Nie Y, Xu C, Miao J, Zhong B. The parallel molecular adaptations to the antarctic cold environment in two psychrophilic green algae. Genome Biol Evol. 2019;11:1897–908.

83. Merchant SS, Prochnik SE, Vallon O, et al. The chlamydomonas genome reveals the evolution of key animal and plant functions. Science. 2007;318:245–50.

84. Herron MD, Michod RE. Evolution of complexity in the volvocine algae: transitions in individuality through Darwin's eye. Evolution. 2008;62:436–51.

85. Nozaki H, Misawa K, Kajita T, Kato M, Nohara S, Watanabe MM. Origin and evolution of the colonial volvocales (chlorophyceae) as inferred from multiple, chloroplast gene sequences. Mol Phylogenet Evol. 2000;17:256–68.

86. Pröschold T, Darienko T, Krienitz L, Coleman AW. Chlamydomonas schloesseri sp. nov. (Chlamydophyceae, Chlorophyta) revealed by morphology, autolysin cross experiments, and multiple gene analyses. Phytotaxa. 2018;362:21–38.

87. Nakada T, Tsuchida Y, Tomita M. Improved taxon sampling and multigene phylogeny of unicellular chlamydomonads closely related to the colonial volvocalean lineage Tetrabaenaceae-Goniaceae-Volvocaceae (Volvocales, Chlorophyceae). Mol Phylogenet Evol. 2019;130:1–8.

88. Nozaki H. Morphology and evolution of sexual reproduction in the Volvocaceae (Chlorophyta). J Plant Res. 1996;109:353–61.

89. Nozaki H, Itoh M. Phylogenetic relationships within the colonial volvocales (chlorophyta) inferred from cladistic analysis based on morphological data. J Phycol. 1994;30:353–65.

90. Nozaki H, Itoh M, Watanabe MM, Kuroiwa T. Ultrastructure of the vegetative colonies and systematic position of Basichlamys (Volvocales, Chlorophyta). Eur J Phycol. 1996;31:67–72.

91. Nozaki H, Itoh M, Sano R, Uchida H, Watanabe MM, Kuroiwa T. Phylogenetic relationships within the colonial volvocales (chlorophyta) inferred from rbcl gene sequence data. J Phycol. 1995;31:970–9.

92. Lartillot N, Lepage T, Blanquart S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics. 2009;25:2286–8.

93. Thorne JL, Kishino H, Painter IS. Estimating the rate of evolution of the rate of molecular evolution. Mol Biol Evol. 1998;15:1647–57.

94. Lepage T, Bryant D, Philippe H, Lartillot N. A general comparison of relaxed molecular clock models. Mol Biol Evol. 2007;24:2669–80.

95. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006;4:e88.

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

97. 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. American J Botany. 2004;91:214–27.

98. Zhou C, Li XH, Xiao S, et al. A new SIMS zircon U-Pb from the Ediacaran Doushantuo Formation: age constraint on the Weng'an biota. Geol Mag. 2017;154:1193–201.

99. Colbath GK, Grenfell HR. Review of biological affinities of Paleozoic acid-resistant, organic-walled eukaryotic algal microfossils (including "acritarchs"). Rev Palaeobot Palynol. 1995;86:287–314.

100. Strother PK, Foster C. A fossil record of land plant origins from charophyte algae. Science. 2021;373:792–6.

101. Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, et al. The timescale of early land plant evolution. Proc Natl Acad Sci USA. 2018;115 https://doi.org/10.1073/pnas.1719588115.

102. Kozlowski R, Kazmierczak C. On two Ordovician calcareous algae. Acta Palaentological Polonica. 1968;13.

103. Chai S, Aria C, Hua H. A stem group Codium alga from the latest Ediacaran of South China provides taxonomic insight into the early diversification of the plant kingdom. BMC Biol. 2022;20:199.

104. Parfrey LW, Lahr DJG, Knoll AH, Katz LA. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci USA. 2011;108:13624–9.

105. Eme L, Sharpe SC, Brown MW, Roger AJ. On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. Cold Spring Harb Perspect Biol. 2014;6:a016139–a016139.

106. Walker JW, Walker AG. Ultrastructure of lower cretaceous angiosperm pollen and the origin and early evolution of flowering plants. Ann Mo Bot Gard. 1984;71:464.

107. Gillespie WH, Rothwell GW, Scheckler SE. The earliest seeds. Nature. 1981;293:462–4.

108. Jennings JR. A review of some fossil plant compressions associated with Mississippian and Pennsylvanian coal deposits in the central Appalachians, Illinois Basin, and elsewhere in the United States. Int J Coal Geol. 1986;6:303–25.

109. Penny JHJ. An Early Cretaceous angiosperm pollen assemblage from Egypt. Spec Pap Palaeontol. 1986;35:121–34.

110. Friis EM, Crane PR, Raunsgaard K. Early flowers and angiosperm evolution. Cambridge: Cambridge University Press; 2011.

111. Morris SC, Robison RA More soft-bodied animals and algae from the middle Cambrian of Utah and British Columbia. 48.

112. Baschnagel RA. New Fossil Algae from the Middle Devonian of New York. Trans Am Microsc Soc. 1966;85:297.

113. Graham LE, Cook ME, Wilcox LW, Graham J, Taylor W, Wellman CH, Lewis L. Resistance of filamentous chlorophycean, ulvophycean, and xanthophycean algae to acetolysis: testing proterozoic and paleozoic microfossil attributions. Int J Plant Sci. 2013;174:947–57.

114. Poinar G, Brown AE. A green algae (Chaetophorales: Chaetophoraceae) in Burmese amber. Hist Biol. 2021;33:323–7.

115. Felsenstein J The Newick tree format. In: https://evolution.genetics.washington.edu/phylip/newicktree.html. https://evolution.genetics.washington.edu/phylip/newicktree.html.

116. de Vries J, Stanton A, Archibald JM, Gould SB. Streptophyte terrestrialization in light of plastid evolution. Trends Plant Sci. 2016;21:467–76.

117. Gitzendanner MA, Soltis PS, Wong GK-S, Ruhfel BR, Soltis DE. Plastid phylogenomic analysis of green plants: A billion years of evolutionary history. Am J Bot. 2018;105:291–301.

118. Forest F. Calibrating the tree of life: fossils, molecules and evolutionary timescales. Ann Bot. 2009;104:789–94.

119. Ho SYW, Tong KJ, Foster CSP, Ritchie AM, Lo N, Crisp MD. Biogeographic calibrations for the molecular clock. Biol Lett. 2015;11:20150194.

120. Graur D, Martin W. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends Genet. 2004;20:80–6.

121. Marshall CR. Using the fossil record to evaluate timetree timescales. Front Genet. 2019;10:1049.

Lindsey *et al. BMC Biology*   (2024) 22:79

Page 20 of 21

122. Yang Z, Rannala B. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol Biol Evol. 2006;23:212–26.

123. Bekker A, Holland HD, Wang P-L, Rumble D, Stein HJ, Hannah JL, Coetzee LL, Beukes NJ. Dating the rise of atmospheric oxygen. Nature. 2004;427:117–20.

124. Edwards D, Davies KL, Axe L. A vascular conducting strand in the early land plant Cooksonia. Nature. 1992;357:683–5.

125. Wickett NJ, Mirarab S, Nguyen N, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc Natl Acad Sci USA. 2014;111 https://doi.org/10.1073/pnas.1323926111.

126. 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:evab101.

127. Lang D, Weiche B, Timmerhaus G, Richardt S, Riaño-Pachón DM, Corrêa LGG, Reski R, Mueller-Roeber B, Rensing SA. Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: a timeline of loss, gain, expansion, and correlation with complexity. Genome Biol Evol. 2010;2:488–503.

128. Sánchez-Baracaldo P, Raven JA, Pisani D, Knoll AH. Early photosynthetic eukaryotes inhabited low-salinity habitats. Proc Natl Acad Sci USA. 2017;114 https://doi.org/10.1073/pnas.1620089114.

129. 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;12:1879.

130. 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;2:1556–62.

131. Lutzoni F, Nowak MD, Alfaro ME, et al. Contemporaneous radiations of fungi and plants linked to symbiosis. Nat Commun. 2018;9:5451.

132. 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.

133. Becker B. Snow ball earth and the split of Streptophyta and Chlorophyta. Trends Plant Sci. 2013;18:180–3.

134. Del Cortona A, Jackson CJ, Bucchini F, et al. Neoproterozoic origin and multiple transitions to macroscopic growth in green seaweeds. Proc Natl Acad Sci USA. 2020;117:2551–9.

135. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

136. Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021;18:366–8.

137. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 2013;41:e121–e121.

138. Smith SA, Brown JW, Walker JF. So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. PLoS One. 2018;13:e0197433.

139. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things): phytools: R package. Methods Ecol Evol. 2012;3:217–23.

140. Revell LJ. phytools 2.0: An updated R ecosystem for phylogenetic comparative methods (and other things). 2023; https://doi.org/10.1101/2023.03.08.531791.

141. Heritage S. MBASR: Workflow-simplified ancestral state reconstruction of discrete traits with MrBayes in the R environment. 2021; https://doi.org/10.1101/2021.01.10.426107.

142. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst Biol. 2012;61:539–42.

143. Arakaki Y, Kawai-Toyooka H, Hamamura Y, Higashiyama T, Noga A, Hirono M, Olson BJSC, Nozaki H. The simplest integrated multicellular organism unveiled. PLoS One. 2013;8:e81641.

144. Arakaki Y, Fujiwara T, Kawai-Toyooka H, Kawafune K, Featherston J, Durand PM, Miyagishima S, Nozaki H. Evolution of cytokinesis-related protein localization during the emergence of multicellularity in volvocine green algae. BMC Evol Biol. 2017;17:243.

145. Yamashita S, Arakaki Y, Kawai-Toyooka H, Noga A, Hirono M, Nozaki H. Alternative evolution of a spheroidal colony in volvocine algae:

146. developmental analysis of embryogenesis in Astrephomene (Volvocales, Chlorophyta). BMC Evol Biol. 2016;16:243.

147. Yamashita S, Nozaki H. Embryogenesis of flattened colonies implies the innovation required for the evolution of spheroidal colonies in volvocine green algae. BMC Evol Biol. 2019;19:120.

148. Cole DG, Reedy MV. Algal morphogenesis: how volvox turns itself inside-out. Curr Biol. 2003;13:R770–2.

149. Pocock MA. Two multicellular motile green algae, volvulina playfair and astrephomene, a new genus. Trans R Soc S Afr. 1954;34:103–27.

150. Kikuchi K. Cellular differentiation in Pleodorina californica. Cytologia. 1978;43:153–60.

151. Smith GM. A comparative study of the species of volvox. Trans Am Microsc Soc. 1944;63:265–310.

152. Umen JG. Evolution of sex and mating loci: An expanded view from Volvocine algae. Curr Opin Microbiol. 2011;14:634–41.

153. Pröschold T, Marin B, Schlösser UG, Melkonian M. Molecular phylogeny and taxonomic revision of chlamydomonas (chlorophyta). i. emendation of chlamydomonas ehrenberg and chloromonas gobi, and description of oogamochlamys gen. nov. and lobochlamys gen. nov. Protist. 2001;152:36.

154. Kodric-Brown A, Brown JH. Anisogamy, sexual selection, and the evolution and maintenance of sex. Evol Ecol. 1987;1:95–105.

155. Togashi T, Horinouchi Y, Parker GA A comparative test of the gamete dynamics theory for the evolution of anisogamy in Bryopsidales green algae.

156. Zink RM, Mckitrick MC. The debate over species concepts and its implications for ornithology. Auk. 1995;112.

157. Häuser CL. The debate about the biological species concept - a review. J Zool Syst Evol Res. 2009;25:241–57.

158. Korshunova T, Picton B, Furfaro G, Mariottini P, Pontes M, Prkić J, Fletcher K, Malmberg K, Lundin K, Martynov A. Multilevel fine-scale diversity challenges the 'cryptic species' concept. Sci Rep. 2019;9:6732.

159. Struck TH, Feder JL, Bendiksby M, et al. Finding evolutionary processes hidden in cryptic species. Trends Ecol Evol. 2018;33:153–63.

160. Nozaki H, Ohtani S. Gonium sociale (Volvocales, Chlorophyta) from Antarctica. Jpn J Phycol. 1992;40:267–71.

161. Mai JC, Coleman AW. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J Mol Evol. 1997;44:258–71.

162. Dujardin F. Historie naturelle des zoophytes. In: Text + Atlas, Librairie Encyclop. de Roret. Paris; 1841.

163. Ohtani S, Nakatsubo T. Japan-China collaboration research program on terrestrial biology at great wall station in king george island, in the summer of 1990/91. Antarctic Record (Nankyoku Shiryō). 36:109–55.

164. Mayr E. What is a species, and what is not? Philos Sci. 1996;63:262–77.

165. UTEX. Algae Details: UTEX. 1885. https://utex.org/products/utex-lb-1885.

166. NIES Microbial Culture Collection Strain Data: NIES-865. https://mcc.nies.go.jp/strainList.do?condition=&strainId=1130.

167. Herron MD, Ghimire S, Vinikoor CR, Michod RE. Fitness trade-offs and developmental constraints in the evolution of soma: an experimental study in a volvocine alga. Evol Ecol Res. 2014;16:203–21.

168. Palmer JD, Stein DB. Conservation of chloroplast genome structure among vascular plants. Curr Genet. 1986;10:823–33.

169. Knoll AH, Bambach RK, Payne JL, Pruss S, Fischer WW. Paleophysiology and end-Permian mass extinction. Earth Planet Sci Lett. 2007;256:295–313.

170. Weismann A. The Germ-Plasm: a Theory of Heredity (English Translation by W. N. Parker & H. Rönnfeldt). New York: Charles Scribner's Sons; 1893.

171. Parker GA, Baker RR, Smith VGF. The origin and evolution of gamete dimorphism and the male-female phenomenon. J Theor Biol. 1972;36:529–53.

172. Maynard-Smith J. Evolution and the Theory of Games: In situations characterized by conflict of interest, the best strategy to adopt depends on what others are doing. Am Sci. 1976;64:41–5.

173. Lindsey CR, Knoll AH, Herron MD, Rosenzweig F. Fossil-calibrated molecular clock data enable reconstruction of steps leading to differentiated multicellularity and anisogamy in the Volvocine algae. Dryad. https://doi.org/10.5061/dryad.mcvdnck6b.

174. Xie Y, Wu G, Tang J, et al. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics. 2014;30:1660–6.

174. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28:3150–2.
175. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019;20:238.
176. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
177. Steenwyk JL, Rokas A. Orthofisher: a broadly applicable tool for automated gene identification and retrieval. G3 Genes|Genomes|Genetics. 2021;11:jkab250.
178. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
179. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30:3276–8.
180. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.
181. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics. 2011;27:1164–5.
182. 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;32:268–74.
183. Smith SA, Dunn CW. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics. 2008;24:715–6.
184. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018;35:518–22.
185. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 2018;19:153.
186. Junier T, Zdobnov EM. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics. 2010;26:1669–70.

Publisher's Note