Genome sequence and cell biological toolbox of the highly regenerative, coenocytic green feather alga *Bryopsis*

Kanta K. Ochiai[1#], Daiki Hanawa[2#], Harumi A. Ogawa[1], Hiroyuki Tanaka[2], Kazuma Uesaka[3], Tomoya Edzuka[1], Maki Shirae-Kurabayashi[1], Atsushi Toyoda[4,5], Takehiko Itoh[2*], Gohta Goshima[1,6*]

1. Sugashima Marine Biological Laboratory, Graduate School of Science, Nagoya University, Toba 517-0004, Japan
2. School of Life Science and Technology, Tokyo Institute of Technology, Meguro-ku, Tokyo 152-8550, Japan
3. Centre for Gene Research, Nagoya University, Nagoya 464-8602, Japan
4. Comparative Genomics Laboratory, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan
5. Advanced Genomics Center, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan
6. Department of Biological Science, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan

#These authors contributed equally to this work.

*To whom correspondence should be addressed.
Email: ggoshima@gmail.com; takehiko@bio.titech.ac.jp

Author contributions: M.S-K. and G.G. conceived the project; K.K.O., M.S-K. and G.G. designed the research; K.K.O., H.A.O., T.E. and M.S-K. performed the experiments; A. T. performed sequencing; K.K.O., D.H., H.T., K.U., and T.I. analysed the sequence data; K.K.O. and G.G. wrote the paper. All authors contributed to the writing of the methodology.

Key words: Marine macroalgae, coenocyte, regeneration, lectin, kinesin, *Bryopsis*, chloroplast transport

Abstract

Green feather algae (Bryopsidales) undergo a unique life cycle in which a single cell repeatedly executes nuclear division without cytokinesis, resulting in the development of a thallus (> 100 mm) with characteristic morphology called coenocyte. *Bryopsis* is a representative coenocytic alga that has exceptionally high regeneration ability: extruded cytoplasm aggregates rapidly in seawater, leading to the formation of protoplasts. However, the genetic basis of the unique cell biology of *Bryopsis* remains poorly understood. Here, we present a high-quality assembly and annotation of the nuclear genome of *Bryopsis* sp. (90.7 Mbp, 27 contigs, N50 = 6.7 Mbp, 14,034 protein-coding genes). Comparative genomic analyses indicate that the genes encoding BPL-1/Bryohealin, the aggregation-promoting lectin, are heavily duplicated in *Bryopsis*, whereas homologous genes are absent in other Ulvophycean algae, suggesting the basis of regeneration capability of *Bryopsis*. *Bryopsis* sp. possesses >30 kinesins but only a single myosin, which differs from other green algae that have multiple types of myosin genes. Consistent with this biased motor toolkit, we observed that the bidirectional motility of chloroplasts in the cytoplasm was dependent on microtubules but not actin in *Bryopsis* sp. Unexpectedly, most genes required for cytokinesis in plants are present in *Bryopsis*, including those in the SNARE or kinesin superfamily. Nevertheless, a kinesin crucial for cytokinesis initiation in plants (NACK/Kinesin-7II) is hardly expressed in the coenocytic part of the thallus, possibly underlying the lack of cytokinesis in this portion. The present genome sequence lays the foundation for experimental biology in coenocytic macroalgae.

Significance statement

The exceptionally coenocytic body and remarkable regeneration ability of *Bryopsis* have attracted biologists for years. However, molecular biological tools remain underdeveloped, partly due to the lack of genome information. Here, we report high-quality assembly and annotation of the genome, providing a crucial resource for experimental biology and genomics studies of *Bryopsis*. Furthermore, comparative genomic analysis reveals a unique gene repertoire that possibly underlies the highly regenerative coenocytic body.

Introduction

Eukaryotic cells are typically characterised by a single nucleus at the centre of the cytoplasm. However, some exceptions exist. For example, red blood cells are anucleated. Multinucleated cells have also been observed in a variety of species. In animals, the syncytium in *Drosophila* embryos and muscle cells in mammals have been extensively studied in cell and developmental biology, for example for the mechanisms of nuclear positioning and synchronised nuclear division (Kwon and Scholey, 2004; Padilla et al., 2022). In flowering plants, seed endosperm undergoes repeated mitotic nuclear divisions

without cytokinesis after double fertilisation, forming a large multinucleated cell called 'coenocyte' (Ali et al., 2023). Many species of marine macroalgae (seaweeds) possess multinucleated cells in their body (Graham et al., 2008). An extreme situation is seen in green feather algae; the thalli of *Caulerpa* or *Bryopsis* develop and reach over 10 cm in length with characteristic side branches, but strikingly, there are no cell walls to separate the numerous nuclei (Mine et al., 2008). This coenocytic feature raises many evolutionary and cellular biology questions, such as how the characteristic features evolve specifically in this algal lineage or how intracellular components are organised in the extremely large cytoplasm (Mine et al., 2008; Umen and Herron, 2021). Non-uniform distribution of transcripts might partly contribute to cytoplasmic organisation in coenocytes (Ranjan et al., 2015). However, the underlying mechanism remains poorly understood, partly because of the lack of an experimental model system in which genetic and molecular biological tools can be instantly applied. As the first step, it is critical to understand the genome sequences and gene repertoires of these species.

Among green feather algae, *Bryopsis* has garnered special attention for its remarkable regenerative capabilities in laboratory settings: cytoplasm extruded from mature thalli is rapidly clustered and transformed into protoplasts, followed by thallus development under the laboratory culture condition (Ikeuchi et al., 2016; Kim et al., 2001; Pak et al., 1991; Tatewaki and Nagata, 1970). This amazing regeneration ability, undergoing 'life without a cell membrane' (Kim et al., 2001), might be critical for this single-celled organism when they are physically damaged, for example by predators (Zan et al., 2019). Regarding the factors required for regeneration, Kim and colleagues found that the aggregation of the extruded cytoplasm is facilitated by the F-type domain-containing lectin termed Bryohealin (also called BPL-1) in *B. plumosa* (Kim et al., 2006). The BPL-1-like protein similarly facilitates aggregation in *Bryopsis hypnoides* (Niu et al., 2009). Aggregation is inhibited by N-acetyl-D-glucosamine and N-acetyl-D-galactosamine, which possess high affinity to BPL-1 (Kim et al., 2006; Niu et al., 2009; Yoon et al., 2008). Three other types of lectins, BPL-2 (Han et al., 2010a), BPL-3 (Han et al., 2010b), and BPL-4 (Han et al., 2012), have been also identified in *Bryopsis*, which bind to the above two sugars (BPL-3/4) or D-mannose (BPL-2). Since extremely high regeneration ability is a unique feature of *Bryopsis*, an interesting scenario would be that some of these lectins uniquely evolved in *Bryopsis*.

The taxon Chlorophyta, to which most green algae belong, exhibits remarkably varied body plans (Del Cortona et al., 2020; Gulbrandsen et al., 2021; Hou et al., 2022; Leebens-Mack et al., 2019). Green microalgae, such as the model species *Chlamydomonas reinhardtii*, are generally unicellular with a single nucleus, whereas Dasycladales, including the classical cell biology model organism *Acetabularia*, is unicellular with a single nucleus but with extremely large cytoplasm (up to 10 cm). Ulvales species have canonical multicellular bodies that are made of mono-nucleated cells separated by cell walls, while Cladophorales is multicellular with multiple nuclei

per cell. Several genome sequences of green algae are available, including those of the coenocytes *Caulerpa lentillifera* and *Ostreobium quekettii* (Arimoto et al., 2019; Hanschen and Starkenburg, 2020; Iha et al., 2021). However, genomic information is lacking for the family Bryopsidaceae, which includes the genus *Bryopsis*. Moreover, the gene repertoire that possibly characterises coenocytic cells has not been extensively investigated yet. In this study, we present the first and high-quality genome sequences of *Bryopsis* species (registered as *Bryopsis* sp. KO-2023). We then report the cell biological toolbox of *Bryopsis* and other green algae.

Results and Discussion

Characterisation of *Bryopsis* species isolated on Sugashima Island, Japan

We isolated two *Bryopsis*-like specimens from an outdoor tank at Sugashima Marine Biological Laboratory (Fig. 1A). Sequencing of the rDNA ITS locus showed > 99.5% identity in 437 base pairs (bp) with that of a *Bryopsis* species registered in the database (line name: HIRO:HIRO-MY 77087). DNA staining showed that multiple nuclei were distributed in the cytoplasm of the main axis, confirming the coenocytic feature (Fig. 1A, middle). High regeneration ability was also confirmed. When the cytoplasm was squeezed out, the extrusion quickly aggregated and transformed into a membrane-encircling protoplast, followed by tip growth (Fig. 1B, Movie 1). Furthermore, this process was suppressed by N-acetyl-D-glucosamine (Fig. S1A) (Kim et al., 2006; Niu et al., 2009).

Next, we tested whether the obtained lines underwent a previously reported life cycle (Tatewaki, 1973). The morphology of the gametes suggested that one line was male and the other was female. Under conditions similar to those used in previous studies, we successfully observed gamete production from both lines, mating of the gametes to generate a sporophyte (diploid), and zoospore generation (Fig. 1A).

We also observed the microtubules and actin filaments using confocal microscopy after immunostaining. They were observed only near the thallus surface, that is, in the cortical cytoplasm, and ran along the main axis of the thallus (Fig. S1B, C). They overlapped largely, but not entirely. The microtubules were not visible after treatment with oryzalin, a microtubule-destabilising drug widely used in land plants. Colocalised actin filaments were also diminished, whereas other short actin bundles remained (Fig. S1D). In contrast, the commonly used actin inhibitor, latrunculin A, completely destroyed actin filaments, whereas microtubule bundles remained intact (Fig. S1E). These observations are largely consistent with those of previous studies using different drugs and epifluorescence microscopy (Menzel and Schliwa, 1986a; Menzel and Schliwa, 1986b).

Based on these observations, we concluded that the collected lines were male and female *Bryopsis*.

Genome sequences and annotation – nucleus

We extracted RNA and DNA separately from haploid thalli (female) and performed sequencing. A draft nuclear genome was assembled based on the short and long reads. The genome comprised 27 contigs (90.7 Mbp, N50 length 6.7 Mbp) (Table 1). The average coverage was 45× (short reads) and 322× (long reads). The GC content was 45.9%, similar to that of *O. quekettii* (52.4%) and *C. lentillifera* (40.4%) (Table S1).

Several contigs had a common repeat sequence (CCCTAAA) at the end (Fig. 2A, red bars at the end of contigs). This sequence was identical to the telomeric repeat sequences of *Arabidopsis thaliana* (Richards and Ausubel, 1988), suggesting that they represent the chromosomal end. This repeat was identified at both ends of the five contigs, suggesting that complete sequences of the five chromosomes were obtained in our analysis. In the other eight contigs, the repeat was observed at one end. Provided that this repeat indeed represents telomeric sequences, *Bryopsis* sp. haploid would possess nine or more chromosomes.

A total of 14,034 protein-coding genes were predicted in 27 contigs (Table 1). BUSCO analysis (protein mode) using the chlorophyta lineage dataset indicated that 92.6% of the single-copy orthologues were recovered, which was higher than those of *O. quekettii* (55.0%) and *C. lentillifera* (67.0%) (Table S1).

These analyses suggest that the nuclear genome of *Bryopsis* sp. was assembled and annotated with high quality compared to many other algal genomes (Hanschen and Starkenburg, 2020).

Genome sequences and annotation – chloroplast and mitochondrion

The chloroplast genome was assembled into a single circular sequence. The number and identity of protein-coding genes, rRNA, and tRNA, as well as the overall genome size were comparable to those of the reported sequences derived from *B. plumosa* and *B. hypnoides* (Leliaert and Lopez-Bautista, 2015; Lu et al., 2011) (Table 2). Detailed information on the genome, including unique features identified in our line, is provided in the **Supplementary Document**.

The mitochondrial genome was assembled into a single circular sequence (Table 2). Our sequence substantially diverged from the reported '*Bryopsis plumosa*' sequence (Han et al., 2020). However, our own analysis of the reported sequences indicated that the specimen belonged to the order Ulvales, and not Bryopsidales (Fig. S2). We think that ours represent the first full mitochondrial DNA sequences of *Bryopsis*. The detailed description the genome feature is provided in the **Supplementary Document**.

Overview of the *Bryopsis* sp. nuclear genome

The availability of high-quality genome allowed us to conduct a high-level comparative genomic study of *Bryopsis*. As comparison, we selected two land plant

species and 20 green algal species (5 macroalgae and 15 microalgae), which covered several classes in Chlorophyta (Fig. 2B, Table S2). The genomes of most species have been annotated in high quality, except for *O. quekettii* (Bryopsidales), whose BUSCO value (genome mode) is less than 70% (Table S1).

First, the comparison of the sequences of 10 single copy genes indicated that *Bryopsis* sp. was indeed phylogenetically classified into the order Bryopsidales and was closer to *C. lentillifera* than *O. quekettii* (Fig. 2B) (Del Cortona et al., 2020; Gulbrandsen et al., 2021; Hou et al., 2022; Leebens-Mack et al., 2019). Second, the repeat sequences were surveyed, as they would reflect the phylogeny (Dodsworth et al., 2014). In all three Bryopsidales species, Ty1/Copia-type long terminal repeat (LTR) retrotransposons were scarcely detected (<0.01%), in contrast to their prevalence in *Ulva mutabilis*, *C. reinhardtii*, and land plant (Table S1). The LINEs were also infrequently detected in Bryopsidales. These results are consistent with the phylogenetic tree derived from gene sequences. Third, we provided functional annotation based on KEGG (Kyoto Encyclopedia of Genes and Genomes) and investigated which unigenes are over- or under-represented in *Bryopsis* (Table S3). *Bryopsis* sp. had >10% more unigenes than the average numbers of green algae in several categories, including signal transduction, transport and catabolism, and cell motility (Table S3). This analysis, however, could not be applied to other Bryopsidales species, as their relatively poor gene annotation would result in underestimation of the unigene numbers. We next analysed total numbers of the genes in each category, which would be less sensitive to genome quality. This analysis showed that the genes in the signalling pathway including SnRK2 kinase were expanded in Bryopsidales (Fig. S3, Table S4). This pathway is involved in stress response in plants (Chen et al., 2021). How this expansion contributes to coenocytic life cycle remains to be determined.

Overall, the global survey suggests that *Bryopsis* in essence possesses a similar set of genetic pathways to other green algal species.

Massive duplication of genes encoding Bryohealin, a lectin required for cytoplasmic aggregation, specifically in *Bryopsis*

Next, we aimed to identify the specific genes (or gene families) that might characterise *Bryopsis*.

The best-known feature of *Bryopsis* is its amazing regeneration ability, which appears to be specific to this genus. We therefore focused on lectin, which facilitates cytoplasmic aggregation during regeneration (Kim et al., 2006; Niu et al., 2009). We searched for *BPL* lectin genes in the *Bryopsis* sp. genome and identified 12 genes highly homologous to *BPL-1* (named *BPL-1A – BPL-1L*) (Fig. 3). BPL-1 is characterised by a conserved 'F-type domain', which is widely observed in the genome of animals but not of land plants. Interestingly, the F-type domain was hardly found in other green algae genome we surveyed, and could not be identified also in *C. lentillifera* or *O. quekettii*,

which belongs to Bryopsidales; we found them only in Volvocaceae among 26 green algal species surveyed in this study (Fig. 3A, Table S4). Thus, this type of lectin was lost in the majority of the green plant lineage, but dramatically expanded in *Bryopsis*.

BPL-2 lectin protein was also found only in *Bryopsis* (Fig. S4, Table S4). BPL-3 and BPL-4 possess the H-type domain. Our survey identified three and two homologues in the genome of *Bryopsis* sp., respectively. Unlike BPL-1 (F-type domain-containing), the H-type domain was found in the genome of *C. lentillifera* (11 genes). However, we could not identify this type of lectin in other green algae (Fig. S4, Table S4).

We searched for other lectin families, including R-type, L-type, and B-type lectins that are found in land plants, and C-type lectin and galectin that have been extensively studied in animals (Varki et al., 2022). However, we could not identify any of them. The only lectin we found was calnexin/chitinase, which is commonly present in eukaryotes.

Thus, our analysis revealed an intriguing correlation in which key lectin genes that facilitate cytoplasmic aggregation are expanded in *Bryopsis*. Lectin gene duplication might endow *Bryopsis* with its exceptional regeneration ability.

No peculiarity in gene superfamily involved in membrane trafficking, including those essential for plant cytokinesis, in *Bryopsis*

Conceivably, the development of an extremely large cell is accompanied by a sophisticated organisation of the cytoplasm. Genes involved in membrane trafficking, which is required for cellular organisation and cytokinesis, are possibly increased or decreased in Bryopsidales.

Conserved gene families regulating membrane trafficking include the Rab GTPase, which is crucial for vesicle trafficking, and SNARE, which is required for the final step in vesicular trafficking, namely membrane fusion (Lipka et al., 2007). Previous study suggested that the increase in the number of SNARE genes parallels the rise of multicellularity among the green plants (Viridiplantae) and also Opisthokonta, based on the genome-wide survey of model species, such as *A. thaliana*, *P. patens*, *C. reinhardtii*, *Ostreococcus tauri*, *Saccharomyces cerevisiae*, and *Homo sapiens* (Sanderfoot, 2007). Similarly, the number of Rab GTPase is dramatically increased in land plants and animals compared to unicellular yeast, leading to the notion that multicellular organisms have more complex systems of internal membranous organelles than unicellular organisms (Saito and Ueda, 2009). Notably, land plants harbour a large number of Rabs and SNAREs that diverge in a manner unique to plant lineage (Saito and Ueda, 2009).

We searched for genes encoding Rab GTPase and SNARE based on BLAST and confirmed their massive increases in land plants compared to *Chlamydomonas* (Table S4). However, further survey in coenocytic Bryopsidales (*Bryopsis* sp., *C. lentillifera*, *O. quekettii*) and multicellular *Ulva*, and *Chara* (closest relative of land plants) indicated that the gene number was comparable to *Chlamydomonas*, regardless of the body form.

Among SNARE genes, *KNOLLE* is specifically required for the final step of

cytokinesis, namely vesicle fusion to the cell plate; the loss of KNOLLE proteins produces multinucleated cells in land plant cells (Lauber et al., 1997; Saito and Ueda, 2009). However, this type of SNARE was present in Bryopsidales (Table S4). These results suggest that the lack of cytokinesis in *Bryopsis*'s main axis cannot be attributed to the lack of vesicle trafficking machinery.

Cytoskeletal motor toolbox

Cytoskeleton and the associated motor proteins, which are categorised into 'cell motility' in KEGG database, are also key elements to cellular organisation. Microtubules and actin filaments serve as tracks for motor proteins (kinesin/dynein and myosin, respectively) to carry various cargo such as organelles. Although α/β-tubulin and G-actin, the building blocks of microtubules and actin filaments, respectively, are highly conserved molecules, different organisms have remarkably different motor repertoires (Reddy and Day, 2001; Vale, 2003). The motor repertoire reflects the cellular dynamics and lifecycle of a species. For example, the development and function of sperm flagella requires the dynein motor as the force generator and driver of intraflagellar transport, and the loss of flagellated sperm during plant evolution coincides with the loss of dynein genes (Lucas and Geisler, 2022). Long-range transport in filamentous fungi is driven by fast and processive motor Kinesin-3, which is lost in short budding yeast (Siddiqui and Straube, 2017). Spatial distribution of mRNA encoding motor proteins may also be indicative of spatially regulated cellular activity (Andresen et al., 2021b).

We analysed cytoskeletal motor proteins based on the conserved motor domains of myosin, dynein heavy chain (DHC), and kinesin. The targeted genome sequences were of two land plant and nine green algal species (Table S2). In addition, we obtained the raw data on RNA-seq from the database for three species from Dasycladales (*Acetabularia acetabulum*, *Polyphysa clavata*, *Chlorocladus australasicus*), and two from Cladophorales (*Chlorocladiella pisiformis* and *Chlorocladiella medogensis*) (Andresen et al., 2021b; Hou et al., 2022). We assembled those sequences and annotated the genes (BUSCO values in Table S5). Dasycladales has a unique life cycle, in which a giant cytoplasm develops without nuclear division. Cladophorales is multicellular but each cell has multiple nuclei; cytokinesis is not coupled with nuclear division (Del Cortona et al., 2020; Shirae-Kurabayashi et al., 2022). For some motors, BLAST search was conducted for those species.

Myosin

Three classes of myosin have been identified in green plants. Myosin-XI drives cytoplasmic streaming and organelle/vesicle transport in *Arabidopsis* and moss (Tamura et al., 2013; Vidali et al., 2010). Closely-related Myosin-XIII is also likely involved in intracellular transport as well as cell growth in green algae, based on localisation study in *Acetabularia* (Andresen et al., 2021b; Vugrek et al., 2003). Cytoplasmic streaming is

dependent on actin filaments in the extremely large cytoplasm of *Acetabularia* (Nagai and Fukui, 1981). Myosin-VIII regulates microtubule-actin crosslinking and is required for cell tip growth, branching, and cytokinesis in moss (Wu and Bezanilla, 2014; Wu and Bezanilla, 2018; Wu et al., 2011). We anticipated that myosin genes would be conserved and the numbers possibly increased in organisms with giant cytoplasm.

This was indeed the case for Dasycladales: we identified at least five Myosin-XI/XIII in all three species examined. In surprising contrast, we identified only one myosin gene (Myosin-XI) in *Bryopsis* sp. (Fig. 4A, S5, Table S4). Other Bryopsidales species had two Myosin-XI genes, but no Myosin-VIII or -XIII. This contrasted with Cladophorales, which had multiple Myosin-XI and Myosin-XIII genes, or *U. mutabilis* and *C. reinhardtii*, where Myosin-VIII was present (Fig. 4A, B, S5, Table S4).

The lack of Myosin VIII in Bryopsidales and Cladophorales might be consistent with the lack of nuclear division-coupled cytokinesis in these organisms. In contrast, the underdevelopment of Myosin-XI/XIII suggests that actomyosin system is unexpectedly not prevalent in the intracellular transport of Bryopsidales.

Dynein

Dynein is the major minus-end-directed (or 'retrograde') transporter in many species, except for seed plants, which lack dynein genes. Our analysis identified 13 dynein heavy chain (*DHC*) genes in *Bryopsis* sp. (Table S4). Each belongs to one of the 16 subfamilies of *C. reinhardtii* DHC (Hom et al., 2011), which consists of either the inner arm, outer arm, or intraflagellar transport (IFT) dynein complex. This was an expected finding, as flagella were present in the gametes and zoospores of *Bryopsis* sp. (Fig. 1A). We analysed the expression level of *DHC* genes based on RNA-seq. We observed that the expression of each *DHC* gene was extremely low in the main axis or rhizoid and elevated in the side branch where flagella were later developed (Table S6, $p < 0.05$ for 8 out of 13 genes, Likelihood ratio test). Similar *DHC* repertoire was identified in other green algal species (some genes were not identifiable either because they are absent or genome assembly is incomplete).

In the Opisthokonta lineage, 'cytoplasmic dynein' was evolved and acts as the major retrograde transporter in the cytoplasm of animal and fungal cells. However, we could not see the development of new types of dynein (i.e. non-flagellar dynein) in any green algal species, including *Bryopsis* sp.

Kinesin

The kinesin superfamily has been further classified into 14 subfamilies (Lawrence et al., 2004; Shen et al., 2012). We identified a total of 34 kinesin genes in *Bryopsis* (Fig. S6.1–6.3, Table S4). Several notable features are as follows:

<u>*Kinesin-GA.*</u> The phylogenetic tree indicated that 20 genes belong to the canonical kinesin subfamily. Their functions can be deduced from the rich research history on

kinesins in animal and plant models. However, 14 kinesins form clades that are apparently green algae-specific and do not contain plant kinesins (termed GA1–10 clades). GAs represent 40% of the total kinesins of *Bryopsis* sp.; the function of each kinesin-GA is unknown. We suggest the addition of these new subfamilies to the kinesin superfamily.

Kinesin-14. Land plants duplicated Kinesin-14 genes and utilise them as retrograde transporters. In *P. patens*, Kinesin-14II (KCH) is responsible for nuclear migration, whereas Kinesin-14VI (KCBP) transports the chloroplasts and others (Yamada and Goshima, 2018; Yamada et al., 2017; Yoshida et al., 2019). We identified in *Bryopsis* sp. two Kinesin-14II and three Kinesin-14VI genes, which may act as transporters (Fig. 4A, Table S4); the expression level of Kinesin-14VI is high (Table S6). Three or more Kinesin-14VI genes were found in Bryopsidales and Cladophorales, whereas Dasycladales and *Ulva* have one or two. The increase in kinesin-14VI genes and their high expression are consistent with the notion that Bryopsidales heavily utilises a microtubule-based system for cargo transport.

Kinesin-ARK. Animals use Kinesin-1 (also called 'conventional kinesin') as the versatile plus-end-directed (or 'anterograde') transporter, whereas ARK kinesin has recently been identified as the plant counterpart (Kanda et al., 2023; Yoshida et al., 2023). Some algal species possess a kinesin whose motor domain is similar to ARK but lacks their characteristic tail (here termed Kinesin-ARK). These are candidate anterograde transporters. However, the orthologous genes are missing in Bryopsisdales. Instead, they encode an algae-specific kinesin (kinesin-GA1) that is phylogenetically close to Kinesin-ARK (Fig. 4A, C). This kinesin subfamily possibly participates in anterograde transport; however, our RNA-seq analysis suggested that the expression level of GA1 was extremely low throughout the haploid thallus (Table S6). Therefore, it remains unclear which genes drive anterograde transport in *Bryopsis*. Intriguingly, an algae-specific Kinesin-GA9 gene (GMH32198.1) showed the highest expression level among cytoskeletal motors throughout the thallus, comparable to a sum of three Kinesin-14VIs (Table S6: total reads of this GA9 and 14VI were 535 and 498 [Deseq2]). We speculate that this novel kinesin subfamily plays an important role in *Bryopsis*, possibly as anterograde transporters.

Kinesin-12. Kinesin-12 genes are expanded in plants; six and 18 genes have been identified in the genomes of *A. thaliana* and *P. patens*, respectively (Shen et al., 2012). The majority of plant Kinesin-12 genes studied thus far are involved in cytokinesis. For example, plant Kinesin-12II (PAKRP) is localised in the midzone of phragmoplasts (a microtubule-based apparatus assembled in late mitosis) and is required for cytokinesis (Lee et al., 2007). Kinesin-12I (POK) is essential for the directed expansion of phragmoplasts and for division plane orientation (Livanos and Muller, 2019). In our survey, Kinesin-12II was found only in *Chara braunii* and land plants. This coincides with the development of phragmoplasts in plant evolution (Buschmann and Zachgo, 2016). However, multiple other Kinesin-12 genes, including POK-like kinesin and

unclassified ones, were present in coenocytic *Bryopsis* sp. or *C. lentillifera* (Fig. 4A). They were highly expressed throughout the haploid thallus (Table S6). The result suggests that Kinesin-12I has a hitherto unknown, non-cytokinetic function in cells.

Kinesin-7. Mutants of Kinesin-7II (also known as NACK) fail to form the cell plate, resulting in multinucleate cells in tobacco and *Arabidopsis* (Nishihama et al., 2002; Tanaka et al., 2004). Upon sister chromatid separation in mitotic anaphase, Kinesin-7II recruits MAP kinase to the phragmoplast, by which conserved microtubule-binding protein MAP65 is phosphorylated (Sasabe and Machida, 2012). MAP65 then recruits proteins involved in vesicle trafficking for cell plate formation (Steiner et al., 2016). Thus, this kinesin acts at cytokinesis initiation. In this context, the presence of kinesin-7II in *Bryopsis* and *C. lentillifera* was unexpected (Fig. 4A). However, gamete formation in the side branch involves cellularisation in *Bryopsis*. RNA-seq analysis indicated that kinesin-7II is hardly expressed in the main axis (1.5 ± 0.49 reads, ±SD, n = 4, normalised by DESeq2) or rhizoid (0.0 ± 0.0) but is expressed at higher levels in the side branch (7.7 ± 2.6). Thus, it is tempting to speculate that the lack of cell separation in the cytoplasm in the *Bryopsis* is partly attributed to the reduced presence of this kinesin protein.

Chloroplast motility depends on cytoplasmic microtubules, but not actin filaments

Cytoplasmic streaming in the giant cytoplasm of *Acetabularia* or in the internodal cell of *Chara* is inhibited by actin filament disassembly (Nagai and Fukui, 1981; Nagai and Kamiya, 1977). Consistent with this, multiple myosin-XIs, one of which is the fastest cytoskeletal motor (Haraguchi et al., 2022), are encoded by *C. braunii* (Fig. 4A, Table S4). Similarly, the addition of an actin polymerisation inhibitor suppressed chloroplast motility in the *Bryopsis* thallus (Menzel and Schliwa, 1986b). However, this observation was hard to reconcile with the genomics data where only one myosin gene was identified in *Bryopsis*. Therefore, we empirically revisited the contribution of microtubules and actin in intracellular transport (Fig. 5).

We focused on chloroplasts because they are autofluorescent and can be traced unambiguously using live confocal imaging. A previous study indicated that motility is dependent on both actin filaments and microtubules (Menzel and Schliwa, 1986b). We observed that chloroplasts moved along the long axis at 339 ± 131 nm/s (± SD, n =50). The movement was bidirectional and a directional switch was occasionally observed (Fig. 5B, red arrowhead; 5C, arrow; Movie 2). Motility was dependent on microtubules; oryzalin treatment almost completely abolished motility (Fig. 5C, D). Surprisingly, motility was not affected by latrunculin A treatment, although the concentration and incubation time were identical to those used when actin disappearance was confirmed by immunofluorescence microscopy (Fig. 5C, D, S1E). We presumed that cytochalasin D, which was used in a previous study to disrupt the actin cytoskeleton, has an off-target effect in *Bryopsis*. The presence of only one myosin in *Bryopsis* sp. is consistent with the notion that bidirectional transport is not driven by actomyosin. We conclude that

chloroplast motility is dependent on microtubules, but not on actin filaments. The bidirectional nature of motility suggests the involvement of both retrograde and anterograde transporters. Multiplicated kinesin-14VI genes are strong candidates responsible for retrograde motility.

Conclusions

This study provides the first information on the nuclear genome of the family Bryopsidaceae. Small contig numbers (27) and the detection of probable telomere sequences at both ends of the five contigs suggested a high-level assembly. These sequences allowed comparative genomic analyses, as illustrated here for several gene families. In addition, specialised chromosomal DNA sequences such as centromeres may be analysable. Male and female lines have been cultured in the laboratory for a few years and could, therefore, be excellent targets for developing tools for genetics and the cell and developmental biology of *Bryopsis*.

Materials and methods

Bryopsis isolation and culture

Two *Bryopsis*-like macroalgal thalli were collected on 7th November 2019 from an outside tank at the Sugashima Marine Biological Laboratory. In addition to having relevant morphology and life cycle, they were confirmed to be *Bryopsis* by PCR, using primers designed for the rDNA ITS region (Shirae-Kurabayashi et al., 2022). Daily cultivation of haploid thalli was conducted at 15 °C (90 µmol $m^{-2}s^{-1}$, light: 16 h, dark: 8 h) in ocean surface water (salt concentration 2.8–3.4%), which was filtered using a 0.22-µm Millipore Stericup, autoclaved, and supplied with Daigo's IMK medium (252 mg/L, Shiotani M.S.). Male and female gametes were obtained by culturing severed haploid thalli for 1–2 weeks at 15 °C (90 µmol $m^{-2}s^{-1}$, light: 16 h, dark: 8 h). They were mixed and cultured under the same conditions for ~1 week. Once sporophyte (diploid) germination was detected, the culture condition was changed (25 °C, 20 µmol $m^{-2}s^{-1}$, light: 10 h, dark: 14 h). After six months, the cells darkened. Culture conditions were changed (15 °C, 90 µmol $m^{-2}s^{-1}$, light: 16 h, dark: 8 h). Zoospores (haploids) were released under these conditions, followed by germination in ~1 week.

Protoplast formation from extruded cytoplasm

The thallus was cut with a scalpel, and sandwiched and crushed with two slide glasses. The extruded cytoplasm was slowly dripped into autoclaved seawater in the presence or absence of N-acetyl-D-glucosamine (40 mM) or the control glucose (40 mM).

RNA sequencing (RNA-seq)

For genome assembly and gene annotation *The Bryopsis* sample (female line, *Bryopsis*

sp. KO-2023) was crushed in liquid nitrogen, and the total RNA was purified using the RNeasy Plant Mini Kit (#74904; Qiagen, Hilden, Germany) with DNase treatment, according to the manufacturer's instructions. The RNA yield was quantified using a NanoVue microplate reader (GE Healthcare, Chicago, IL, USA). The sample volume was adjusted to 2 µg/100 µL for subsequent RNA-seq analysis. RNA-seq analysis was performed at the core facility of Nagoya University following the protocol described by (Matsumura et al., 2022). Briefly, 1 µg total RNA was used for mRNA purification with NEBNext Oligo d(T)$_{25}$ (NEBNext poly(A) mRNA Magnetic Isolation Module; New England Biolabs, Ipswich, MA, USA), followed by first-strand cDNA synthesis with the NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs) and NEBNext Multiplex Oligo for Illumina (New England Biolabs) according to the manufacturer's protocols. The amount of cDNA was determined using an Agilent 4150 TapeStation System (Agilent, Santa Clara, CA, USA). The cDNA libraries were sequenced as paired-end reads of 81 nucleotides using an Illumina NextSeq 550 (Illumina, San Diego, CA, USA).

Spatial dissection Fragments of < 1 mm from the tip of the main axis of *Bryopsis* sp. were cut and cultured in autoclaved seawater supplemented with Daigo's IMK medium for 10 –14 days at 15 °C (90 µmol $m^{-2}s^{-1}$, light: 16 h, dark: 8 h). The thalli that developed side branches were cut into three parts; 'side branch', 'main axis' (central stalk), and 'rhizoid'. After removing water, each sample was separately crushed with mortar and pestle that had been prechilled at -80 °C, and the total RNA was purified using the RNeasy Plant Mini Kit. This manipulation was independently performed four times on different days. RIN values for all samples were greater than 8.0. The samples were sequenced with Illumina NovaSeq6000 platform, which produced 150 bp paired-end reads. The amount of reads for each gene was calculated using RSEM v1.2.28 (Li and Dewey, 2011) with STAR v2.7.10b (Dobin et al., 2012) for mapping. Normalisation was performed using TPM and DESeq2 (Love et al., 2014).

Genome sequencing

Whole-genome shotgun sequencing was performed using the PacBio and Illumina sequencing platforms. Genomic DNA from *Bryopsis* sp. KO-2023 (female) was isolated using a CTAB/Genomic-tip Kit (QIAGEN). A SMRTbell library for continuous long-read (CLR) sequencing was prepared using a SMRTbell Express Template Prep Kit 2.0 (Pacific Bioscience, CA, USA) according to the manufacturer's instructions. The CLR library was size-selected using the BluePippin system (Sage Science, Beverly, MA, USA) with a lower cutoff of 30 kb. One SMRT Cell 8M was sequenced on the PacBio Sequel II system with Binding Kit 2.0 and Sequencing Kit 2.0 (20 h collection times). In addition, genomic DNA was fragmented to an average size of 500 bp using an M220 Focused-ultrasonicator M220 (Covaris Inc., Woburn, MA. USA). A paired-end library with insert

sizes ranging from 450 to 550 bp was constructed using the TruSeq DNA PCR-Free Library Prep kit (Illumina) and was size-selected on an agarose gel using a Zymoclean Large Fragment DNA Recovery Kit (Zymo Research, Irvine, CA. USA). The final library was sequenced using a 2 × 150 bp paired-end protocol on the NovaSeq 6000 system (Illumina).

Genome assembly

Chloroplast De novo assembly of the chloroplast genome was performed using a combination of 150 bp × 2 short reads and Get-organelle v 1.7.6.1 (Jin et al., 2020) with the options -k 21, 45, 65, 85, 105, -P 1000000, and -R 50. Two complete *Bryopsis* chloroplast sequences (NC_026795.1 and NC_013359.1) were used as seeds. This provided two closed circular sequences of identical length (91,672 nt). The two sequences were nearly identical except for the central region (~11 kb). One sequence was discarded because structural errors were found near the central region when it was aligned with long reads. The other sequences showed no structural errors across the entire sequence length. The error check was repeated at different starting positions. Finally, the downstream of *psbA* was set at +1 position.

Mitochondrion Highly fragmented contigs with a total length of ~150 kb were obtained using Get-organelle v 1.7.6.1 assembly (Jin et al., 2020) with the seed references of green algal species (NC_045361.1, KU161104.1, and NC_001638.1) (Repetti et al., 2020; Vahrenholz et al., 1993; Zhou et al., 2016). These putative mitochondrial sequences had a sequencing depth ~200 times higher than that of the nuclear genome. The high copy number of the mitochondrial genome enabled assembly based on random selection of a small portion of PacBio long reads (≥20 kb). One percent of the long reads was sufficient for the assembly of the mitochondrial genome. Flye (Kolmogorov et al., 2019; Lin et al., 2016) with '--pacbio-raw' option produced one circular sequence (356,161 bp) that had global synteny with other algal mitochondrial sequences. To check if there was mis-assembly in this sequence, full long and short reads were aligned using minimap2 (Li, 2021) with the 'map-pb' and 'sr' presets, respectively. This revealed six indel errors at the homopolymer sites but did not identify any large sequence gaps or structural errors. Small indels were corrected using bwa (mapping) and Pilon (Walker et al., 2014). To confirm the completeness of the mitochondrial genome assembly, the +1 position was changed by 20,000 bp and the long reads were aligned using minimap2. No sequence gaps were found during this operation, indicating that no structural errors existed in the mitochondrial assembly. Finally, the +1 position was reset downstream of *rrnL3b*.

Nuclear genome The assembly of long-read data was used to determine the nuclear genome. However, the genome sequences of symbiotic bacteria, commonly detected in marine macroalgae, inevitably contaminate *Bryopsis* genome sequences. Therefore, a

provisional genome assembly was first performed, in which the obtained genome sequences were clustered into groups which were thought to originate from the same species. Based on the sequence characteristics and mapping results of the RNA-seq data, grouped sequences considered to be derived from *Bryopsis* were identified. Sequences were extracted from clustered groups.

Illumina reads were used for K-mer analysis and genome size estimation. The 21-mer frequencies were calculated using Jellyfish v2.3.0 (Marcais and Kingsford, 2011), and the genome size was estimated using GenomeScope 2.0 (Ranallo-Benavidez et al., 2020). The estimated genome size was used as the input parameter for *de novo* pre-assembly. Pre-*de novo* assembly of the nuclear genome was performed based on the PacBio reads using Canu v2.1.1 (Koren et al., 2017) with the following options: genomeSize = 500M, corOutCoverage = 200, and 'batOptions = -dg 3 -db 3 -dr 1 -ca 500 -cp 50'. Pre-assembled contigs were polished using long and short reads. They were polished through three rounds of Arrow v2.3.3, and three rounds of Pilon v1.23 (Walker et al., 2014). In these steps, PacBio reads were mapped using pbmm2 v1.3.0 (https://github.com/PacificBiosciences/pbmm2), and trimmed Illumina reads were mapped using BWA v0.7.17 (Li, 2013). Then, binning was performed using MetaBAT2 v2.15 (Kang et al., 2019) to group contigs derived from the same species, and each cluster was named 'bin'. As input for MetaBat2, read coverage information was calculated from the Illumina read mapping results against polished pre-assembled contigs using BWA v0.7.17.

Raw RNA-seq data were trimmed and filtered using Platanus_trim v1.0.7. *De novo* transcriptome assembly was performed based on the trimmed RNA-seq reads using Trinity v2.8.5 (Grabherr et al., 2011). Transcriptome assembly contigs were splice-mapped to polished, pre-assembled genomic contigs using GMAP v.2018-08-25 (Wu and Watanabe, 2005). The bin containing the most-mapped transcriptome assembly contigs was designated as the main nuclear bin. In addition, other bins and contigs derived from *Bryopsis* were manually selected based on the overall information, such as the transcriptome assembly contig mapping rate, GC rate, and Illumina read coverage.

PacBio and Illumina reads derived from *Bryopsis* were extracted for the final *de novo* assembly. PacBio reads were extracted from Canu intermediate files used in the pre-*de novo* assembly. Illumina reads were extracted by mapping the trimmed Illumina reads to contigs derived from *Bryopsis* using BWA v0.7.17. The extracted trimmed Illumina reads were used for K-mer analysis and genome size estimation, as described above. The estimated genome size was used as an input parameter for the final *de novo* assembly. Final *de novo* assembly of the nuclear genome was performed based on the PacBio reads derived from *Bryopsis* using Canu v2.2 with the following options: genomeSize = 100M, corOutCoverage = 200, and 'batOptions= -dg 3 -db 3 -dr 1 -ca 500 -cp 50'. The final assembled contigs were polished using long and short reads. The final assembly contigs were polished through three rounds of Arrow v2.3.3 and three rounds of NextPolish

v1.4.0 (Hu et al., 2019). Next, the arrow-identified variants were filtered via Merfin v1.0 (Formenti et al., 2022) using the trimmed Illumina reads derived from *Bryopsis*. In the long-read-based polish, PacBio reads derived from *Bryopsis* were mapped using pbmm2 v1.3.0. Haplotigs were then removed using Purge_dups v1.2.3 (Guan et al., 2020) to reduce sequence redundancy and increase assembly continuity.

These analyses yielded the assembly and selection of 49 contigs. Finally, to verify the origin of each contig, BLASTx searches were conducted for a portion of the sequence of each contig. The sequences derived from 22 contigs were highly homologous to bacterial and fungal sequences, whereas those of the other 27 contigs were not. Thus, 27 contigs were considered derived from *Bryopsis*.

Gene annotation

Chloroplast ncRNAs were annotated using the GeSeq web server. 'DNA search identity' was set at 85. Four reference sequences (NC_013359.1, NC_026795.1, NC_037363.1, and NC_ 030629.1) were used as '3rd Party References.' The CDS was manually annotated using a combination of GeSeq annotation, protein alignment with *B. plumosa* (NC_026795.1), and RNA-seq alignment. This collaborative annotation was further curated using a homology-based approach against the proteomes of closely related species to verify the completeness of each CDS. In total, 83 predicted protein-coding genes, three rRNAs, and 26 tRNAs were identified.

Mitochondrion Annotation of the mitochondrial genome using GeSeq predicted virtually no protein-coding genes. This suggests that no closely related protein-coding genes were annotated. To overcome this limitation, open reading frames (ORFs) were searched using the NCBI ORF finder (https://www. ncbi.nlm.nih.gov/orffinder/). The predicted ORFs of all six frames were manually aligned with the mitochondrial proteins of *Ostreobium quekettii* (Repetti et al., 2020) and the putative CDS coding frame of the RNA-seq was constructed with TransDecorder. To verify the obtained CDS, promising coding frames were manually searched for homology to proteins of closely related species using BLASTx. This procedure identified 40 protein-coding genes with complete CDS sequences. In addition, a tBLASTn search using publicly available green algal mitochondrial protein sequences identified 14 small genes encoded in the introns of already annotated genes. ncRNAs were annotated using the Geseq web server. tRNAs were identified with Geseq, where the following '3rd Party References' were used: NC_045361.1, NC_001638.1, NC_028538.1, NC_035722.1, NC_029701.1, NC_035809.1, NC_28081.1, NC_040163.1, and NC_041082.1. This resulted in 17 annotated tRNAs. rRNAs were searched against the mitochondrial genome using BLASTn with the following queries: NC_045361.1, NC_001638.1, NC_028538.1, NC_035722.1, NC_029701.1, NC_035809.1, NC_28081.1, NC_040163.1, and NC_041082.1. Candidate genes were manually compared with the RNA-seq alignment

data. This procedure identified three rRNA genes in the mitochondrial genome. Intron length is defined as the length of the region between exons within a gene (protein-coding or non-coding). When other genes were present within the introns of a host gene, the length of the internal gene was not excluded from the intron length of the host gene. Domains of genes present in introns were searched using NCBI's Conserved Domains database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) with default settings.

Nucleus Protein-coding genes were predicted by combining the results of RNA-seq-, homology-, and *ab initio*-based prediction methods. RNA-seq-based prediction utilises both assembly-first and mapping-first methods. For the assembly-first method, RNA-seq data were assembled using Trinity v2.12.0 (Grabherr et al., 2011) and Oases v2.0.9 (Schulz et al., 2012). The redundant assembled RNA contigs were removed using CD-HIT v4.8.1 (Fu et al., 2012), and then splice-mapped to the genome sequences using GMAP v2018-07-04 (Wu and Watanabe, 2005). For the mapping-first method, RNA-seq data were mapped to genome scaffolds using HISAT2 v2.2.1 (Kim et al., 2019), and gene sets were predicted with StringTie v2.2.0 (Pertea et al., 2016) from mapped results. The ORF regions were estimated using TransDecorder v5.5.0 (https://github.com/TransDecoder/TransDecoder) from both the assembly-first and mapping-first method results. Regarding homology-based prediction, amino acid sequences of *O. quekettii* (NCBI accession No: GCA_905146915.1), *C. reinhardtii* (NCBI accession No: GCF_000002595.2), *Volvox carteri* (NCBI accession No: GCF_000143455.1), and *Monoraphidium neglectum* (NCBI accession No: GCF_000611645.1), were splice-mapped to genome scaffolds using Spaln v2.3.3f (Gotoh, 2008), and gene sets were predicted. For *ab initio* prediction, training sets were first selected from the RNA-seq-based prediction results. Then, AUGUSTUS v3.3.3 (Stanke and Waack, 2003) was trained using this set. The SNAP v2006-07-28 (Korf, 2004) was used in this study. All predicted genes were combined using an in-house merging tool. However, the ORF of some genes did not start with ATG (methionine), which was manually fixed. In some cases, the start codon was manually identified, and the amino acid sequences were corrected. In other cases (~700), the ORF assignment was rejected as the start codon and transcript could not be identified. Finally, 14,034 genes encoding proteins were identified.

De novo transcriptome assembly and annotation

De novo transcriptome assembly and gene annotation were conducted based on the published RNA-seq raw data, following the methods described in (Andresen et al., 2021b) and (Hou et al., 2022) for the following species: [Dasycladales] *Acetabularia acetabulum*, *Chlorocladus australasicus* and *Polyphysa clavata*; [Cladophorales] *Chlorocladiella pisiformis* and *Chlorocladiella medogensis* (Supplementary Data). The raw sequence data

were obtained from the European Nucleotide Archive under the accession No. PRJEB40460 and PRJNA726747.

Genome information used in this study

The genomes primarily used in each analysis were *Bryopsis* sp. KO-2023 (this study), *C. lentillifera* (Arimoto et al., 2019), *O. quekettii* (Iha et al., 2021), *U. mutabilis* (De Clerck et al., 2018), *C. reinhardtii* (Merchant et al., 2007), *Dunaliella salina* (Polle et al., 2017), *Pleodorina starrii* (Takahashi et al., 2023), *V. carteri* (Prochnik et al., 2010), *Raphidocelis subcapitata* (Suzuki et al., 2018), *Monoraphidium neglectum* (Bogen et al., 2013), *Auxenochlorella protothecoides* (Gao et al., 2014), *Coccomyxa subellipsoidea* C-169 (Blanc et al., 2012), *Chlorella vulgaris* (Cecchin et al., 2019), Pedinophyceae sp. YPF-701 (Repetti et al., 2022), *Chloropicon primus* (GCA_023205875.1), *Micromonas pusilla* (Worden et al., 2009), *O. tauri* (Blanc-Mathieu et al., 2014), and *Bathycoccus prasinos* (Yau et al., 2020) for Chlorophyta and *Klebsormidium nitens* (Hori et al., 2014), *C. braunii* (Nishiyama et al., 2018), *P. patens* (Lang et al., 2018) and *A. thaliana* (Lin et al., 1999; Mayer et al., 1999; Salanoubat et al., 2000; Tabata et al., 2000; Theologis et al., 2000) for Streptophyta (Table S2). Note that the available *A. acetabulum* genome sequences were not amenable to comparative genomics due to low quality (BUSCO <11%) (Andresen et al., 2021a).

Circular visualization of the genome assembly (Circos plot)

The genomic features of the 27 contigs were plotted in a circular genome plot using shinyCircus V2.0 hosted in a local server (Wang et al., 2023). GC content was calculated as the ratio of the average of AT and GC per 10,000 bp. For repetitive sequences plot, all types of repeats were used from the result of repeatmasker (see below). All information used for the circus-plot is available (https://github.com/KantaOchiai/Bryopsis_sp._KO-2023_genome_sequence_Information).

Comparative genomics analysis

Repetitive sequences Repetitive sequences were identified using a combination of *de novo* and homology-based methods. First, Repeat sequences were *de novo* searched using RepeatModeler v2.0.1 (http://www.repeatmasker.org/RepeatModeler/) with "--LTRstruct". Then, identified repetitive sequences, including transposable elements, were counted using RepeatMasker v4.1.1 (http://www.repeatmasker.org/) based on the repeat model created by RepeatModeler (Table S1).

Evaluation of assembly quality BUSCO metrics were used to assess the integrity of the genome assembly and the completeness of the gene prediction (Waterhouse et al., 2017). BUSCO v5.5.0 was run with genome or protein mode on 18 published genomes of Chlorophyta, including *Bryopsis* sp. with Chlorophyta dataset (chlorophyta_odb10), and

four published genomes of Streptophyta with the Viridiplantae (viridiplantae_odb10) or Brassicales (brassicales_odb10) dataset (Table S1). The transcriptome mode was applied for transcriptomes of two Cladophorales and three Dasycladales with Chlorophyta dataset (chlorophyta_odb10) (Table S5).

Functional annotation with KEGG database Functional annotation was performed based on KEGG (Kyoto Encyclopedia of Genes and Genomes) using GhostKoala (Kanehisa et al., 2016). The unigenes of each pathway in each genome were counted with KEGG mapper (https://www.genome.jp/kegg/mapper/) (Table S3). Subsequently, 'MAPK signaling pathway-plants' in the 'Signal transduction' category was analysed with BLASTp searches using the representative *A. thaliana* proteins as queries, as extremely high number of genes were identified in this category for Bryopsidales including *Bryopsis* sp. (accession No: PYR/PYL/RCARs (NP_180174.1, O49686.1, NP_563626.1), PP2C_GroupA (P49598.1), HOS15 (Q9FN19.1), RBOH (O48538.1, Q9FIJ0.1), KAT1 (Q39128.1), QUAC1 (O49696.1), SLAC1 (Q9LD83.1), ABFs/ABI (Q9M7Q3.1, Q9SJN0.1, Q9M7Q5.1), SOD (AEE74978.1, AEE85010.1), CAT1 (Q96528.3)). SnRK2 annotated with KEGG was confirmed by KEGG BLASTp web server (Fig. S3, Table S4).

Phylogenetic inference

Chlorophyta species 10 highly conserved single-copy OGs were selected from 63 single copy ortholog genes (OGs) obtained using Orthofinder v2.3.14 (Emms and Kelly, 2019) in 18 published genomes of Chlorophyta including *Bryopsis* sp. and three Streptophyta (Table S2). 10 single-copy OGs list is available in Supplementary Data. Each OG sequences were aligned using MAFFT v7.505 (Katoh and Standley, 2013) with FFT-NE-2 strategy. All gaps were removed using MEGAX (Kumar et al., 2018), and the individual OGs were combined to obtain a sequence of 2,713 amino acids (Supplementary Data). Finally, ML tree was inferred using IQ-TREE v1.6.12 (Nguyen et al., 2015) with LG+F+R4 selected as the best-fit model and branch support estimated with ultrafast 1,000 bootstrap.

Mitochondrial genome Seven mitochondrial housekeeping genes, including *nad1*, *nad2*, *nad4*, *nad5*, *nad6*, *cob*, *cox1*) were retrieved from 17 species, including *Bryopsis* sp. and registered *B. plumosa* (MN853874.1) (Fig. S2). The same procedure as for chloroplasts was used for the subsequent analysis.

Lectin BLASTp/tBLASTn searches were conducted for published *Bryopsis* BPL-1, -2, -3, and -4 proteins. For all possible hit sequences (Supplementary Data), the presence of characteristic domains of each BPL protein was confirmed with the NCBI conserved domain search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). BLASTp/tBLASTn searches were also conducted against *Bryopsis* sp. for R-, L-, B- and

C-type lectins (accession No: P06750.1, PWZ39448.1, AAL09432.1, Q9FVA1.1, Q9FV99.1, Q9NNX6.1), malectin (accession No: AEE78805.1), calnexin (accession No: KAB1259615.1), calreticulin (accession No: CAA55890.1), chitinase (accession No: AEC10291.1), and galectin (accession No: KAJ0248405.1) as queries. Amino acid sequences of each gene were aligned by MAFFT v7.505 with FFT-NE-2 strategy. All gaps were removed using MEGAX, and sequences of 116 amino acids (BPL-1), 132 amino acids (BPL-2), and 102 amino acids (BPL-3/4) were obtained (Supplementary Data). ML tree was drawn using IQ-TREE v1.6.12 with WAG+G4 (BPL-1, -2) or LG+G4 (BPL-3/4) selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap.

Rab GTPase and SNARE Genes were searched with BLASTp/tBLASTn using the representative *A. thaliana* proteins as queries (Rab GTPase accession No: NP_568678.1, SNARE: (Lipka et al., 2007)).

Myosin Genes were searched with BLASTp/tBLASTn in nine genomes of Chlorophyta, including *Bryopsis* sp., three genomes of Streptophyta, and five transcriptomes of Cladophorales and Dasycladales (Fig. 4A, Table S2), using the following queries: Myosin-VIII (accession No: F4JIU4.1), Myosin-XI (accession No: F4HXP9.1, GMH40817.1), and Myosin-XIII (accession No: AAB53061.1, AAB53062.1). All hit sequences with the e-value $\leq e^{-10}$ were subjected to the NCBI conserved domain search, and the sequences in which conserved motor domains could not be identified were removed from the list (Supplementary Data). Some myosin proteins, for which long amino acid sequences could be retrieved, were shown as schematic diagrams (Fig. 4B) and/or subjected to phylogenetic tree construction (Fig. S5). For tree construction, the amino acid sequences were aligned by MAFFT v7.505 with FFT-NE-2 strategy and all gaps were removed using MEGAX, and a sequence of 184 amino acids was obtained (Supplementary Data). ML tree was drawn using IQ-TREE v1.6.12 with LG+I+G4 selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap (Fig. S5).

Dynein heavy chain (DHC) Genes were searched with BLASTp/tBLASTn using previously reported *C. reinhardtii* DHC1–16 proteins (Hom et al., 2011) as queries (Table S4).

Kinesin Genes were searched with BLASTp/tBLASTn using the amino acid sequences of 1–350 aa of the human kinesin heavy chain (KIF5B/kinesin-1: accession No: P33176.1) and *Arabidopsis thaliana* KIN4C (accession No: F4K0J3.2) as queries. Additional BLASTp/tBLASTn searches were conducted for several kinesins: Kinesin-ARK and Kinesin-GA1 (ARK-like) in 10 genomes of Chlorophyta, seven genomes of

Streptophyta, and five transcriptomes of Cladophorales and Dasycladales; Kinesin-7II, Kinesin-12, Kinesin-14II, and Kinesin-14VI in five transcriptomes of Cladophorales and Dasycladales (Table S2). All hit sequences with the e-value ≤ e^{-10} were subjected to the NCBI conserved domain search, and the sequences in which conserved motor domains could not be identified were removed from the list (Supplementary Data). The kinesin amino acid sequences in nine published genomes of Chlorophyta, including *Bryopsis* sp., and three Streptophyta were aligned by MAFFT v7.505 with FFT-NE-2 strategy and all gaps were removed using MEGAX, followed by ML tree construction (IQ-TREE v1.6.12 with LG+I+G4 and branch support was estimated with 1000 ultrafast bootstrap) (Fig. S6).

Immunostaining

A three-week-old thallus after cytoplasm extrusion was fixed with 4% paraformaldehyde in modified PHEM buffer (Sobue et al., 1988) (60 mM Pipes, 25 mM Hepes, 0.5 M NaCl, 10 mM EGTA, 2 mM $MgCl_2$; pH 6.9) for 1 h at 25 °C, followed by permeabilisation with 1% Triton X-100 in PBS for 1 h at 25°C. After washing twice with PBST (0.1% Triton X-100 in PBS), the specimen was incubated with blocking solution (1% BSA in PBST) for 1 h at 25 °C, followed by addition of primary antibodies at 4 °C overnight with rotation (mouse anti-β-actin [Proteintech, 66009-1-Ig], 1:1000, and rat anti-α-tubulin [YOL1/34, MCA78G, Bio-Rad], 1:1000). The specimen was washed three times with PBST and incubated with secondary antibodies (anti-mouse, Jackson ImmunoResearch, 715-545-151, 1:1000, and anti-rat, Jackson ImmunoResearch, 712-165-153, 1:1000) and DAPI (final 1 µg/ml) overnight at 4 °C with rotation. After washing twice with PBST, the specimen was mounted on a glass slide with a mounting medium (Fluoromount™; Diagnostic BioSystems).

Microscopy

Bryopsis sp. thalli were imaged using a Nikon SMZ800N stereo microscope, Plan Apo 1x/WF lens, and NY1S-EA camera (SONY). The gametes and zoospores were imaged using an ECLIPSE E200 microscope (Nikon) and NY1-EA2. Fluorescent images of DNA (DAPI), chloroplasts, microtubules, and actin were acquired using a Nikon Ti2 inverted microscope equipped with a CSU-10 spinning-disc confocal scanner unit (Yokogawa), a Zyla CMOS camera (Andor), and four laser lines (637, 561, 488, and 405 nm). 40× 0.95 NA lens or a 100× 1.40 NA lens was used to image live or fixed cells, respectively. To obtain the chloroplast motility rate, a 35-mm glass-bottom dish was prepared, on which a piece of kitchen garbage net (~10 × 20 mm) was attached with double-sided tape. After cytoplasmic extrusion, a 3-week-old thallus and a coverslip were laid over the net, followed by the addition of 1-mL of autoclaved seawater. This net prevented thallus movement during imaging. Autofluorescent chloroplasts were imaged every 10 s using a spinning-disc confocal microscope and a 40× 0.95 NA lens. At 2 min during imaging of untreated specimen, oryzalin (10 µM), latrunculin A (10 µM), or control DMSO was

834 added (3 mL volume each). The unidirectional motility rate of randomly selected
835 chloroplasts 5–6 min after drug addition was manually measured after obtaining
836 kymograph images using Fiji.
837
838 **Data availability**
839 The genome sequence of *Bryopsis* sp. is available at the DNA Data Bank of Japan
840 (DDBJ) under project PRJDB15746
841 (https://ddbj.nig.ac.jp/resource/bioproject/PRJDB15746) and sample accession
842 SAMD00599708 (https://ddbj.nig.ac.jp/resource/biosample/SAMD00599708) with
843 accession numbers BSYQ01000001.1–BSYQ01000027.1 (nuclear genome), LC768901
844 (chloroplast), and LC768902 (mitochondria). The raw sequence data for NextSeq 550,
845 NovaSeq 6000, and Sequel II are available under accession numbers DRA016305,
846 DRA016314, and DRA016315, respectively. The assembled genome and annotation are
847 also available from NCBI with GenBank accession ID: GCA_030272585.1. The IDs of
848 the genes used for the phylogenetic tree construction are shown in the figures.
849 Gene and protein sequences used for phylogenetic tree construction and comparative
850 genomic analyses are summarised in Supplementary data
851 (https://github.com/KantaOchiai/Bryopsis_sp._KO-
852 2023_genome_sequence_Information).
853
854 ## Acknowledgements
855 We are grateful to the staff of the Comparative Genomics Laboratory at NIG for
856 supporting genome sequencing. This work was funded by the Japan Society for the
857 Promotion of Science KAKENHI (16H06279 (PAGS) for whole-genome sequencing and
858 22K19308, 22H04717, and 22H02644 for experimental biology). The authors declare no
859 conflict of interest.

860 **Table 1. Information of the nuclear genome of *Bryopsis* sp. KO-2023**

861

		***Bryopsis* sp. KO-2023**
Accession number of genome		BSYQ01000001.1–BSYQ01000027.1
Assembly statistics	Genome size (Mbp)	90.7
	Assembly level	Contig
	Number of chromosomes/scaffolds/contigs	27
	scaffolds/contigs N50 (Mbp)	6.7
	GC percent (%)	46
	Predicted protein coding genes	14,034
BUSCO_protein	Dataset	Cholophyta
	Complete (%)	95.5
	Complete and single-copy (%)	92.6
	Complete and duplicated (%)	2.9
	Fragmented (%)	0.2
	Missing (%)	4.3
	Total BUSCO groups searched	1,519
Repeat	Total repeat (%)	30.02
	Retroelements	11.87
	SINEs	0.00
	LINEs:	0.76
	LTR elements:	11.10
	Ty1/Copia	0.01
	Gypsy/DIRS1	10.94
	Retroviral	0.03
	DNA transposons:	0.57
	Unclassified:	15.70
	Total interspersed repeats:	28.15

862

863 **Table 2. Comparison of the chloroplast and mitochondrial genome**
864

	***Bryopsis* sp.**	*Caulerpa lentillifera*	*Ostreobium quekettii*	*Ulva* sp.	*Chlamydomonas reinhardtii*	*Physcomitrium patens*	*Arabidopsis thaliana*
Genome	Chloroplast	Chloroplast	Chloroplast	Chloroplast	Chloroplast	Chloroplast	Chloroplast
Accession number of genome	LC768901.1	NC_039377.1	NC_030629.1	KP720616.1	NC_005353.1	NC_005087.2	NC_000932.1
Genome size (Kbp)	91.7	119.4	82.0	100.0	203.8	122.8	154.4
GC percent (%)	30.4	32.6	31.9	25.3	34.5	28.5	36.3
Predicted protein coding genes*	83	91	78	79	65	85	79
rRNA genes*	3	3	3	3	5	3	4
tRNA genes*	26	28	31	28	29	32	30
Coding DNA (%) **	85.4	86.0	84.0	81.8	49.9	72.3	72.0
Large inverted repeat (>5 kb)	absent	absent	absent	absent	present	present	present
Genome	Mitochondrion	Mitochondrion	Mitochondrion	Mitochondrion	Mitochondrion	Mitochondrion	Mitochondrion
Accession number of genome	LC768902.1	KX761577.1	NC_045361.1	KP720617.1	NC_001638.1	NC_007945.1	NC_037304.1
Genome size (Kbp)	356.2	209	241.7	73.5	15.8	105.3	367.8
GC percent (%)	54.4	50.9	48.3	32.4	45.2	40.6	44.8
Predicted protein coding genes*	54	76	54	50	8	42	122
rRNA genes*	3	3	3	2	14	3	3
tRNA genes*	17	20	28	25	3	24	22
Intron number	72	29	47	10	0	26	18
Intronic DNA (%) **	54.1	43.4	39.3	21.7	0	28.4	8.14

865 * Duplicate genes were counted as single genes.
866 ** Total gene length, which includes introns, was divided by the entire genome length.
867

Supplementary document

Overview of the chloroplast genome

In this study, the chloroplast genome was assembled into a single closed sequence of 91,672 base pairs (bp). This length was close to the size of previously reported chloroplast genomes of *Bryopsis plumosa* (106,859 bp) (Leliaert and Lopez-Bautista, 2015) and *Bryopsis hypnoides* (153,429 bp) (Lu et al., 2011). No long reverse repeat sequences were identified, consistent with other green algae of the order Bryopsidales of the family Ulvophyceae, and genus *Ulva* (Turmel and Lemieux, 2018; Turmel et al., 2017). The GC content was 30.4%, which was similar to the reported chloroplast genomes of *B. plumosa* (30.8%) (Leliaert and Lopez-Bautista, 2015) and *B. hypnoides* (33.1%) (Lu et al., 2011). The coding DNA sequences occupied 85.4% of the chloroplast genome, which was much higher than that of the mitochondria (66.1%) (Table 2, Table S7, S9). Drastic expansion of introns, which was evident in the mitochondrial genome, was not observed in either *Bryopsis* lines.

GeSeq-based annotation revealed that the chloroplast genome contained 83 protein-coding genes, 79 of which were identical to the previously annotated *bona fide* or hypothetical protein-coding genes of *B. plumosa* (NC_026795.1) and were conserved in other green algae (Table S8). The remaining four protein-coding genes included two open reading frames (ORFs) found within the introns of *psaA* and *psbB*, one previously reported ORF, and one novel ORF. The two ORFs in the introns showed high homology with the previously reported *orf1* and *orf2* of *B. plumosa* (NC_026795.1). ORF480 (i.e. 480 a.a.) in the intron of *psaA* encodes a protein that has a reverse transcriptase-like superfamily and RVT_N superfamily domains, suggesting that it functions as a reverse transcriptase. In contrast, ORF300 in the intron of *psbB* did not contain any characteristic domains, suggesting that it might not represent a protein.

One of the two isolated ORFs, termed ORF92, is a 281 bp reading frame (i.e. 92 a.a.) found in a ~2.5 kb flanking region between *chlN* and *trnL*. RNA-seq analysis indicated that this gene was transcribed *in vivo*. However, the translated sequences showed no homology to known proteins in the database. Thus, this might be specifically encoded in the chloroplast genome of our line. The other orphan ORF, termed ORF431, showed weak sequence identity with GIIM superfamily proteins (group II intron, maturase-specific domain) according to a domain search (CD-search). ORFs with the GIIM superfamily domain were also present in other orders of Bryopsidales, except *O. quekettii,* suggesting that they are widely conserved in Bryopsidales. A portion of the amino acid sequence also showed weak homology with reverse transcriptases of the order Bryopsidales, suggesting that it may function as a reverse transcriptase.

In addition to protein-coding genes, 26 tRNAs and 3 rRNAs were annotated, consistent with a previous report on *B. plumosa* (NC_026795.1). The anticodons of all the 26 tRNA genes were identical (Table S8).

The chloroplast genome of our line was ~15 kb shorter than the registered genome (NC_026795.1). This was largely because our line had smaller intergenic regions and fewer introns. For example, the intergenic region between *trnG* (*ucc*) and *rrnF* in our line was 1,362 bp, which was much shorter than that of the other line (13,011 bp). NC_026795.1 had an intron and an intronic ORF in the *rrnL* gene, while neither was present in our line.

Overview of the mitochondrial genome

The mitochondrial genome of our *Bryopsis* sp. line was assembled as a single closed sequence of 356,152 bp, which was much longer than the hitherto-reported longest sequence in green algae (*O. quekettii*: 241,739 bp) (Repetti et al., 2020). There is one report on the mitochondrial genome of *B. plumosa* (Han et al., 2020). However, our sequences were substantially different from registered sequences. Our own survey of the sequences reported by Han et al. strongly suggested that their specimen belong to Ulvales, and not Bryopsis (Fig. S2).

We compared the obtained sequences with those of other green algae (Table S9). The size of the genome (356,152 bp) was much larger than that of any other mitochondrial genome of green algae (second longest was that of *O. quekettii* at 241,739 bp (Repetti et al., 2020)). This was partly attributed to an increase of introns: we identified 72 introns in 17 genes, which was more than in *O. quekettii* (47 introns in 18 genes) or *C. lentillifera* (29 introns in 13 genes). In extreme cases, 17 introns and 18 exons were present in *cox1*, whereas only 11, 5, and 4 introns were found in *cox1* of *O. quekettii*, *C. lentillifera*, and *Ulva* sp., respectively (Melton et al., 2015; Repetti et al., 2020; Zheng et al., 2018). In total, introns occupied 54.1% of the genome, which was higher than that in *O. quekettii* (39.3%) or *C. lentillifera* (43.4%).

Manual annotation revealed 54 protein-coding genes, 17 tRNAs, and 3 rRNAs. The rRNA numbers were similar to those of most other green algae (Table S9). tRNAs corresponding to 15 amino acids were identified, whereas those corresponding to Ala, Cys, Glu, Lys, and Asn were not.

Of the 54 protein-coding genes, seven were not unambiguously assigned as real ORFs because the encoded amino acid sequences did not show homology to proteins with known functions. In contrast, 47 genes encoded proteins that have conserved domains, many of which are required for mitochondrial function, such as NADH:ubiquinone oxidoreductase (complex I; *nad* genes) or ATP synthase (complex V; *atp* genes) (Table S10). The number of *nad* and *atp* genes encoded in the mitochondrial genome varies among green algae; our *Bryopsis* line often had more than the average number. For example, *nad10* and *tatC* have been found in the mitochondrial genome but not in many other green algae species. However, there was also a reverse case: the mitochondrial genome of *O. quekettii*, but not ours, had *atp4* gene (Table S10).

Manual annotation revealed 72 introns in 17 genes. Introns were more prevalent than

those in *O. quekettii* (47 introns in 18 genes) or *C. lentillifera* (29 introns in 13 genes) (Table S9). The number of introns was particularly high in *nad5*, *cob*, *cox1* and *atp1*. In extreme cases, 17 introns and 18 exons were identified in *cox1*, whereas only 11, 5, and 4 introns were found in *cox1* of *O. quekettii*, *C. lentillifera*, and *Ulva* sp., respectively (Fig. S7) (Melton et al., 2015; Repetti et al., 2020; Zheng et al., 2018). The mean intron length was 2,676 bp, which was comparable to that of the two Bryopsodales *O. quekettii* (2,022 bp) and *C. lentillifera* (3,126 bp) (Fig. S8). Introns accounted for 54.1% of the mitochondrial genome, which was higher than that in *O. quekettii* (39.3%) and *C. lentillifera* (43.4%) (Table S9).

Interestingly, 14 protein-coding genes were found in the introns of other genes. A tBLASTn search for published green algal mitochondrial proteins (https://ftp.ncbi.nlm.nih.gov/refseq/release/mitochondrion/) identified three ORFs showing homology to the putative LAGLIDADG endonuclease, ten ORFs showing homology to the putative group II intron reverse transcriptase/maturase, and one ORF encoding a putative protein in the introns of *cox1*, *atp1*, and *rnl* (six in *cox1*, five in *atp1*, and three in *rnl*). The introns of *cox1* contain one gene encoding a LAGLIDADG endonuclease and five genes encoding putative group II intron reverse transcriptases/maturases. The encoded LAGLIDADG endonuclease is likely functional because it possesses LAGLIDADG domains at the N- and C-termini that are required for endonuclease activity (Hausner, 2012; Lambowitz and Belfort, 1993). Four ORFs of the putative group II intron reverse transcriptase/maturase contained one or more RT_G2 introns or RT_like superfamily domains, and three of them possessed the Intron_maturas2 superfamily domain, suggesting that these reverse transcriptases are functional (Table S11).

O. quekettii also has endonuclease-like protein ORFs and a putative group II intron reverse transcriptase/maturase on the introns of *cox1*, *atp1*, *rns,* and *rnl*. Thus, the mitochondrial genome size of green algae belonging to the order Bryopsidales, including *Bryopsis*, may have increased in accordance with the increased number and size of introns compared with the mitochondrial genomes of other green algae.

The alignment of *nad2*, *nad7*, *nad5*, *nad9* genes with several green algae, including *O. quekettii* and land plants (*A. thaliana* and *P. patens*), suggested that UGA encodes Trp rather than a termination codon (Fig. S9). This is consistent with other green algae, including *O. queketti*, *Pedinomonas minor*, and *Pycnococcus provasolii* (Noutahi et al., 2019; Repetti et al., 2020).

References

Ali, M.F., J.M. Shin, U. Fatema, D. Kurihara, F. Berger, L. Yuan, and T. Kawashima. 2023. Cellular dynamics of coenocytic endosperm development in Arabidopsis thaliana. *Nat Plants*. 9:330-342.

Andresen, I.J., R.J.S. Orr, A.K. Krabberød, K. Shalchian-Tabrizi, and J. Bråte. 2021a. Genome sequencing and de novo assembly of the giant unicellular alga Acetabularia acetabulum using droplet MDA. *Scientific Reports*. 11:12820.

Andresen, I.J., R.J.S. Orr, K. Shalchian-Tabrizi, and J. Bråte. 2021b. Compartmentalization of mRNAs in the giant, unicellular green alga Acetabularia acetabulum. *Algal Research*. 59:102440.

Arimoto, A., K. Nishitsuji, Y. Higa, N. Arakaki, K. Hisata, C. Shinzato, N. Satoh, and E. Shoguchi. 2019. A siphonous macroalgal genome suggests convergent functions of homeobox genes in algae and land plants. *DNA Res*. 26:183-192.

Blanc, G., I. Agarkova, J. Grimwood, A. Kuo, A. Brueggeman, D.D. Dunigan, J. Gurnon, I. Ladunga, E. Lindquist, S. Lucas, J. Pangilinan, T. Proschold, A. Salamov, J. Schmutz, D. Weeks, T. Yamada, A. Lomsadze, M. Borodovsky, J.M. Claverie, I.V. Grigoriev, and J.L. Van Etten. 2012. The genome of the polar eukaryotic microalga Coccomyxa subellipsoidea reveals traits of cold adaptation. *Genome Biol*. 13:R39.

Blanc-Mathieu, R., B. Verhelst, E. Derelle, S. Rombauts, F.Y. Bouget, I. Carre, A. Chateau, A. Eyre-Walker, N. Grimsley, H. Moreau, B. Piegu, E. Rivals, W. Schackwitz, Y. Van de Peer, and G. Piganeau. 2014. An improved genome of the model marine alga Ostreococcus tauri unfolds by assessing Illumina de novo assemblies. *BMC Genomics*. 15:1103.

Bogen, C., A. Al-Dilaimi, A. Albersmeier, J. Wichmann, M. Grundmann, O. Rupp, K.J. Lauersen, O. Blifernez-Klassen, J. Kalinowski, A. Goesmann, J.H. Mussgnug, and O. Kruse. 2013. Reconstruction of the lipid metabolism for the microalga Monoraphidium neglectum from its genome sequence reveals characteristics suitable for biofuel production. *BMC Genomics*. 14:926.

Buschmann, H., and S. Zachgo. 2016. The Evolution of Cell Division: From Streptophyte Algae to Land Plants. *Trends Plant Sci*. 21:872-883.

Cecchin, M., L. Marcolungo, M. Rossato, L. Girolomoni, E. Cosentino, S. Cuine, Y. Li-Beisson, M. Delledonne, and M. Ballottari. 2019. Chlorella vulgaris genome assembly and annotation reveals the molecular basis for metabolic acclimation to high light conditions. *Plant J*. 100:1289-1305.

Chen, X., Y. Ding, Y. Yang, C. Song, B. Wang, S. Yang, Y. Guo, and Z. Gong. 2021. Protein kinases in plant responses to drought, salt, and cold stress. *Journal of Integrative Plant Biology*. 63:53-78.

De Clerck, O., S.M. Kao, K.A. Bogaert, J. Blomme, F. Foflonker, M. Kwantes, E. Vancaester, L. Vanderstraeten, E. Aydogdu, J. Boesger, G. Califano, B. Charrier, R. Clewes, A. Del Cortona, S. D'Hondt, N. Fernandez-Pozo, C.M. Gachon, M. Hanikenne, L. Lattermann, F. Leliaert, X. Liu, C.A. Maggs, Z.A. Popper, J.A. Raven, M. Van Bel, P.K.I. Wilhelmsson, D. Bhattacharya, J.C. Coates, S.A. Rensing, D. Van Der Straeten, A. Vardi, L. Sterck, K. Vandepoele, Y. Van de Peer, T. Wichard, and J.H. Bothwell. 2018. Insights into the Evolution of Multicellularity from the Sea Lettuce Genome. *Curr Biol*. 28:2921-2933 e2925.

1031 Del Cortona, A., C.J. Jackson, F. Bucchini, M. Van Bel, S. D'Hondt, P. Skaloud, C.F.
1032 Delwiche, A.H. Knoll, J.A. Raven, H. Verbruggen, K. Vandepoele, O. De Clerck,
1033 and F. Leliaert. 2020. Neoproterozoic origin and multiple transitions to
1034 macroscopic growth in green seaweeds. *Proc Natl Acad Sci U S A*. 117:2551-2559.
1035 Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M.
1036 Chaisson, and T.R. Gingeras. 2012. STAR: ultrafast universal RNA-seq aligner.
1037 *Bioinformatics*. 29:15-21.
1038 Dodsworth, S., M.W. Chase, L.J. Kelly, I.J. Leitch, J. Macas, P. Novák, M. Piednoël, H.
1039 Weiss-Schneeweiss, and A.R. Leitch. 2014. Genomic Repeat Abundances
1040 Contain Phylogenetic Signal. *Systematic Biology*. 64:112-126.
1041 Emms, D.M., and S. Kelly. 2019. OrthoFinder: phylogenetic orthology inference for
1042 comparative genomics. *Genome Biology*. 20:238.
1043 Formenti, G., A. Rhie, B.P. Walenz, F. Thibaud-Nissen, K. Shafin, S. Koren, E.W. Myers,
1044 E.D. Jarvis, and A.M. Phillippy. 2022. Merfin: improved variant filtering,
1045 assembly evaluation and polishing via k-mer validation. *Nature Methods*. 19:696-
1046 704.
1047 Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li. 2012. CD-HIT: accelerated for clustering the
1048 next-generation sequencing data. *Bioinformatics*. 28:3150-3152.
1049 Gao, C., Y. Wang, Y. Shen, D. Yan, X. He, J. Dai, and Q. Wu. 2014. Oil accumulation
1050 mechanisms of the oleaginous microalga Chlorella protothecoides revealed
1051 through its genome, transcriptomes, and proteomes. *BMC Genomics*. 15:582.
1052 Gotoh, O. 2008. Direct mapping and alignment of protein sequences onto genomic
1053 sequence. *Bioinformatics*. 24:2438-2444.
1054 Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis,
1055 L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke,
1056 N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman,
1057 and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data
1058 without a reference genome. *Nature Biotechnology*. 29:644-652.
1059 Graham, J.E., L.W. Wilcox, and L.E. Graham. 2008. Algae (2nd Edition). Benjamin
1060 Cummings.
1061 Guan, D., S.A. McCarthy, J. Wood, K. Howe, Y. Wang, and R. Durbin. 2020. Identifying
1062 and removing haplotypic duplication in primary genome assemblies.
1063 *Bioinformatics*. 36:2896-2898.
1064 Gulbrandsen, Ø.S., I.J. Andresen, A.K. Krabberød, J. Bråte, and K. Shalchian-Tabrizi.
1065 2021. Phylogenomic analysis restructures the ulvophyceae. *Journal of Phycology*.
1066 57:1223-1233.
1067 Han, H., Y. Li, S. Wei, Z. Wang, and X. Zhang. 2020. The complete mitochondrial
1068 genome sequence of Bryopsis plumosa. *Mitochondrial DNA B Resour*. 5:1067-
1069 1068.
1070 Han, J.-W., K.-S. Yoon, M.-G. Jung, K.-H. Chah, and G.-H. Kim. 2012. Molecular
1071 characterization of a lectin, BPL-4, from the marine green alga Bryopsis plumosa
1072 (Chlorophyta). *Algae*. 27:55-62.
1073 Han, J.W., M.G. Jung, M.J. Kim, K.S. Yoon, K.P. Lee, and G.H. Kim. 2010a. Purification
1074 and characterization of a D-mannose specific lectin from the green marine alga,
1075 Bryopsis plumosa. *Phycological Research*. 58:143-150.
1076 Han, J.W., K.S. Yoon, T.A. Klochkova, M.-S. Hwang, and G.H. Kim. 2010b. Purification
1077 and characterization of a lectin, BPL-3, from the marine green alga Bryopsis

plumosa. *Journal of Applied Phycology*. 23:745-753.

Hanschen, E.R., and S.R. Starkenburg. 2020. The state of algal genome quality and diversity. *Algal Research*. 50:101968.

Haraguchi, T., M. Tamanaha, K. Suzuki, K. Yoshimura, T. Imi, M. Tominaga, H. Sakayama, T. Nishiyama, T. Murata, and K. Ito. 2022. Discovery of ultrafast myosin, its amino acid sequence, and structural features. *Proc Natl Acad Sci U S A*. 119.

Hausner, G. 2012. Introns, Mobile Elements, and Plasmids. *In* Organelle Genetics: Evolution of Organelle Genomes and Gene Expression. C.E. Bullerwell, editor. Springer Berlin Heidelberg, Berlin, Heidelberg. 329-357.

Hom, E.F., G.B. Witman, E.H. Harris, S.K. Dutcher, R. Kamiya, D.R. Mitchell, G.J. Pazour, M.E. Porter, W.S. Sale, M. Wirschell, T. Yagi, and S.M. King. 2011. A unified taxonomy for ciliary dyneins. *Cytoskeleton (Hoboken)*. 68:555-565.

Hori, K., F. Maruyama, T. Fujisawa, T. Togashi, N. Yamamoto, M. Seo, S. Sato, T. Yamada, H. Mori, N. Tajima, T. Moriyama, M. Ikeuchi, M. Watanabe, H. Wada, K. Kobayashi, M. Saito, T. Masuda, Y. Sasaki-Sekimoto, K. Mashiguchi, K. Awai, M. Shimojima, S. Masuda, M. Iwai, T. Nobusawa, T. Narise, S. Kondo, H. Saito, R. Sato, M. Murakawa, Y. Ihara, Y. Oshima-Yamada, K. Ohtaka, M. Satoh, K. Sonobe, M. Ishii, R. Ohtani, M. Kanamori-Sato, R. Honoki, D. Miyazaki, H. Mochizuki, J. Umetsu, K. Higashi, D. Shibata, Y. Kamiya, N. Sato, Y. Nakamura, S. Tabata, S. Ida, K. Kurokawa, and H. Ohta. 2014. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nat Commun*. 5:3978.

Hou, Z., X. Ma, X. Shi, X. Li, L. Yang, S. Xiao, O. De Clerck, F. Leliaert, and B. Zhong. 2022. Phylotranscriptomic insights into a Mesoproterozoic–Neoproterozoic origin and early radiation of green seaweeds (Ulvophyceae). *Nature Communications*. 13:1610.

Hu, J., J. Fan, Z. Sun, and S. Liu. 2019. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics*. 36:2253-2255.

Iha, C., K.E. Dougan, J.A. Varela, V. Avila, C.J. Jackson, K.A. Bogaert, Y. Chen, L.M. Judd, R. Wick, K.E. Holt, M.M. Pasella, F. Ricci, S.I. Repetti, M. Medina, V.R. Marcelino, C.X. Chan, and H. Verbruggen. 2021. Genomic adaptations to an endolithic lifestyle in the coral-associated alga Ostreobium. *Curr Biol*. 31:1393-1402 e1395.

Ikeuchi, M., Y. Ogawa, A. Iwase, and K. Sugimoto. 2016. Plant regeneration: cellular origins and molecular mechanisms. *Development*. 143:1442-1451.

Jin, J.J., W.B. Yu, J.B. Yang, Y. Song, C.W. dePamphilis, T.S. Yi, and D.Z. Li. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 21:241.

Kanda, A., K. Otani, T. Ueda, T. Takahashi, and H. Motose. 2023. Plant specific armadillo repeat kinesin directs organelle transport and microtubule convergence to promote tip growth. *bioRxiv*:2022.2007.2008.499237.

Kanehisa, M., Y. Sato, and K. Morishima. 2016. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology*. 428:726-731.

Kang, D.D., F. Li, E. Kirton, A. Thomas, R. Egan, H. An, and Z. Wang. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction

from metagenome assemblies. *PeerJ*. 7:e7359.

Katoh, K., and D.M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30:772-780.

Kim, D., J.M. Paggi, C. Park, C. Bennett, and S.L. Salzberg. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*. 37:907-915.

Kim, G.H., T.A. Klochkova, K.-S. Yoon, Y.-S. Song, and K.P. Lee. 2006. Purification and Characterization of a Lectin, Bryohealin, Involved in the Protoplast Formation of a Marine Green Alga Bryopsis Plumosa (Chlorophyta) 1. *Journal of Phycology*. 0:060609080250009-???

Kim, G.H., T.A. Klotchkova, and Y.M. Kang. 2001. Life without a cell membrane: regeneration of protoplasts from disintegrated cells of the marine green alga Bryopsis plumosa. *Journal of Cell Science*. 114:2009-2014.

Kolmogorov, M., J. Yuan, Y. Lin, and P.A. Pevzner. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 37:540-546.

Koren, S., B.P. Walenz, K. Berlin, J.R. Miller, N.H. Bergman, and A.M. Phillippy. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 27:722-736.

Korf, I. 2004. Gene finding in novel genomes. *BMC Bioinformatics*. 5:59.

Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*. 35:1547-1549.

Kwon, M., and J.M. Scholey. 2004. Spindle mechanics and dynamics during mitosis in Drosophila. *Trends Cell Biol*. 14:194-205.

Lambowitz, A.M., and M. Belfort. 1993. Introns as mobile genetic elements. *Annu Rev Biochem*. 62:587-622.

Lang, D., K.K. Ullrich, F. Murat, J. Fuchs, J. Jenkins, F.B. Haas, M. Piednoel, H. Gundlach, M. Van Bel, R. Meyberg, C. Vives, J. Morata, A. Symeonidi, M. Hiss, W. Muchero, Y. Kamisugi, O. Saleh, G. Blanc, E.L. Decker, N. van Gessel, J. Grimwood, R.D. Hayes, S.W. Graham, L.E. Gunter, S.F. McDaniel, S.N.W. Hoernstein, A. Larsson, F.W. Li, P.F. Perroud, J. Phillips, P. Ranjan, D.S. Rokshar, C.J. Rothfels, L. Schneider, S. Shu, D.W. Stevenson, F. Thummler, M. Tillich, J.C. Villarreal Aguilar, T. Widiez, G.K. Wong, A. Wymore, Y. Zhang, A.D. Zimmer, R.S. Quatrano, K.F.X. Mayer, D. Goodstein, J.M. Casacuberta, K. Vandepoele, R. Reski, A.C. Cuming, G.A. Tuskan, F. Maumus, J. Salse, J. Schmutz, and S.A. Rensing. 2018. The Physcomitrella patens chromosome-scale assembly reveals moss genome structure and evolution. *Plant J*. 93:515-533.

Lauber, M.H., I. Waizenegger, T. Steinmann, H. Schwarz, U. Mayer, I. Hwang, W. Lukowitz, and G. Jurgens. 1997. The Arabidopsis KNOLLE protein is a cytokinesis-specific syntaxin. *J Cell Biol*. 139:1485-1493.

Lawrence, C.J., R.K. Dawe, K.R. Christie, D.W. Cleveland, S.C. Dawson, S.A. Endow, L.S. Goldstein, H.V. Goodson, N. Hirokawa, J. Howard, R.L. Malmberg, J.R. McIntosh, H. Miki, T.J. Mitchison, Y. Okada, A.S. Reddy, W.M. Saxton, M. Schliwa, J.M. Scholey, R.D. Vale, C.E. Walczak, and L. Wordeman. 2004. A standardized kinesin nomenclature. *J Cell Biol*. 167:19-22.

Lee, Y.R., Y. Li, and B. Liu. 2007. Two Arabidopsis phragmoplast-associated kinesins

play a critical role in cytokinesis during male gametogenesis. *Plant Cell*. 19:2595-2605.

Leebens-Mack, J.H., M.S. Barker, E.J. Carpenter, M.K. Deyholos, M.A. Gitzendanner, S.W. Graham, I. Grosse, Z. Li, M. Melkonian, S. Mirarab, M. Porsch, M. Quint, S.A. Rensing, D.E. Soltis, P.S. Soltis, D.W. Stevenson, K.K. Ullrich, N.J. Wickett, L. DeGironimo, P.P. Edger, I.E. Jordon-Thaden, S. Joya, T. Liu, B. Melkonian, N.W. Miles, L. Pokorny, C. Quigley, P. Thomas, J.C. Villarreal, M.M. Augustin, M.D. Barrett, R.S. Baucom, D.J. Beerling, R.M. Benstein, E. Biffin, S.F. Brockington, D.O. Burge, J.N. Burris, K.P. Burris, V. Burtet-Sarramegna, A.L. Caicedo, S.B. Cannon, Z. Çebi, Y. Chang, C. Chater, J.M. Cheeseman, T. Chen, N.D. Clarke, H. Clayton, S. Covshoff, B.J. Crandall-Stotler, H. Cross, C.W. dePamphilis, J.P. Der, R. Determann, R.C. Dickson, V.S. Di Stilio, S. Ellis, E. Fast, N. Feja, K.J. Field, D.A. Filatov, P.M. Finnegan, S.K. Floyd, B. Fogliani, N. García, G. Gâteblé, G.T. Godden, F. Goh, S. Greiner, A. Harkess, J.M. Heaney, K.E. Helliwell, K. Heyduk, J.M. Hibberd, R.G.J. Hodel, P.M. Hollingsworth, M.T.J. Johnson, R. Jost, B. Joyce, M.V. Kapralov, E. Kazamia, E.A. Kellogg, M.A. Koch, M. Von Konrat, K. Könyves, T.M. Kutchan, V. Lam, A. Larsson, A.R. Leitch, R. Lentz, F.-W. Li, A.J. Lowe, M. Ludwig, P.S. Manos, E. Mavrodiev, M.K. McCormick, M. McKain, T. McLellan, J.R. McNeal, et al. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*. 574:679-685.

Leliaert, F., and J.M. Lopez-Bautista. 2015. The chloroplast genomes of Bryopsis plumosa and Tydemania expeditiones (Bryopsidales, Chlorophyta): compact genomes and genes of bacterial origin. *BMC Genomics*. 16:204.

Li, B., and C.N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 12:323.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. doi: 10.48550/arXiv.1303.3997.

Li, H. 2021. New strategies to improve minimap2 alignment accuracy. *Bioinformatics*. 37:4572-4574.

Lin, X., S. Kaul, S. Rounsley, T.P. Shea, M.I. Benito, C.D. Town, C.Y. Fujii, T. Mason, C.L. Bowman, M. Barnstead, T.V. Feldblyum, C.R. Buell, K.A. Ketchum, J. Lee, C.M. Ronning, H.L. Koo, K.S. Moffat, L.A. Cronin, M. Shen, G. Pai, S. Van Aken, L. Umayam, L.J. Tallon, J.E. Gill, M.D. Adams, A.J. Carrera, T.H. Creasy, H.M. Goodman, C.R. Somerville, G.P. Copenhaver, D. Preuss, W.C. Nierman, O. White, J.A. Eisen, S.L. Salzberg, C.M. Fraser, and J.C. Venter. 1999. Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. *Nature*. 402:761-768.

Lin, Y., J. Yuan, M. Kolmogorov, M.W. Shen, M. Chaisson, and P.A. Pevzner. 2016. Assembly of long error-prone reads using de Bruijn graphs. *Proc Natl Acad Sci U S A*. 113:E8396-E8405.

Lipka, V., C. Kwon, and R. Panstruga. 2007. SNARE-Ware: The Role of SNARE-Domain Proteins in Plant Biology. *Annual Review of Cell and Developmental Biology*. 23:147-174.

Livanos, P., and S. Muller. 2019. Division Plane Establishment and Cytokinesis. *Annu Rev Plant Biol*. 70:239-267.

Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and

dispersion for RNA-seq data with DESeq2. *Genome Biology*. 15:550.

Lu, F., W. Xu, C. Tian, G. Wang, J. Niu, G. Pan, and S. Hu. 2011. The Bryopsis hypnoides plastid genome: multimeric forms and complete nucleotide sequence. *PLoS One*. 6:e14663.

Lucas, J., and M. Geisler. 2022. Sequential loss of dynein sequences precedes complete loss in land plants. *Plant Physiology*. 189:1237-1240.

Marcais, G., and C. Kingsford. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*. 27:764-770.

Matsumura, M., M. Nomoto, T. Itaya, Y. Aratani, M. Iwamoto, T. Matsuura, Y. Hayashi, T. Mori, M.J. Skelly, Y.Y. Yamamoto, T. Kinoshita, I.C. Mori, T. Suzuki, S. Betsuyaku, S.H. Spoel, M. Toyota, and Y. Tada. 2022. Mechanosensory trichome cells evoke a mechanical stimuli-induced immune response in Arabidopsis thaliana. *Nat Commun*. 13:1216.

Mayer, K., C. Schuller, R. Wambutt, G. Murphy, G. Volckaert, T. Pohl, A. Dusterhoft, W. Stiekema, K.D. Entian, N. Terryn, B. Harris, W. Ansorge, P. Brandt, L. Grivell, M. Rieger, M. Weichselgartner, V. de Simone, B. Obermaier, R. Mache, M. Muller, M. Kreis, M. Delseny, P. Puigdomenech, M. Watson, T. Schmidtheini, B. Reichert, D. Portatelle, M. Perez-Alonso, M. Boutry, I. Bancroft, P. Vos, J. Hoheisel, W. Zimmermann, H. Wedler, P. Ridley, S.A. Langham, B. McCullagh, L. Bilham, J. Robben, J. Van der Schueren, B. Grymonprez, Y.J. Chuang, F. Vandenbussche, M. Braeken, I. Weltjens, M. Voet, I. Bastiaens, R. Aert, E. Defoor, T. Weitzenegger, G. Bothe, U. Ramsperger, H. Hilbert, M. Braun, E. Holzer, A. Brandt, S. Peters, M. van Staveren, W. Dirske, P. Mooijman, R. Klein Lankhorst, M. Rose, J. Hauf, P. Kotter, S. Berneiser, S. Hempel, M. Feldpausch, S. Lamberth, H. Van den Daele, A. De Keyser, C. Buysshaert, J. Gielen, R. Villarroel, R. De Clercq, M. Van Montagu, J. Rogers, A. Cronin, M. Quail, S. Bray-Allen, L. Clark, J. Doggett, S. Hall, M. Kay, N. Lennard, K. McLay, R. Mayes, A. Pettett, M.A. Rajandream, M. Lyne, V. Benes, S. Rechmann, D. Borkova, H. Blocker, M. Scharfe, M. Grimm, T.H. Lohnert, S. Dose, M. de Haan, A. Maarse, M. Schafer, et al. 1999. Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana. *Nature*. 402:769-777.

Melton, J.T., 3rd, F. Leliaert, A. Tronholm, and J.M. Lopez-Bautista. 2015. The complete chloroplast and mitochondrial genomes of the green macroalga Ulva sp. UNA00071828 (Ulvophyceae, Chlorophyta). *PLoS One*. 10:e0121020.

Menzel, D., and M. Schliwa. 1986a. Motility in the siphonous green alga Bryopsis. I. Spatial organization of the cytoskeleton and organelle movements. *Eur J Cell Biol*. 40:275-285.

Menzel, D., and M. Schliwa. 1986b. Motility in the siphonous green alga Bryopsis. II. Chloroplast movement requires organized arrays of both microtubules and actin filaments. *Eur J Cell Biol*. 40:286-295.

Merchant, S.S., S.E. Prochnik, O. Vallon, E.H. Harris, S.J. Karpowicz, G.B. Witman, A. Terry, A. Salamov, L.K. Fritz-Laylin, L. Marechal-Drouard, W.F. Marshall, L.H. Qu, D.R. Nelson, A.A. Sanderfoot, M.H. Spalding, V.V. Kapitonov, Q. Ren, P. Ferris, E. Lindquist, H. Shapiro, S.M. Lucas, J. Grimwood, J. Schmutz, P. Cardol, H. Cerutti, G. Chanfreau, C.L. Chen, V. Cognat, M.T. Croft, R. Dent, S. Dutcher, E. Fernandez, H. Fukuzawa, D. Gonzalez-Ballester, D. Gonzalez-Halphen, A. Hallmann, M. Hanikenne, M. Hippler, W. Inwood, K. Jabbari, M. Kalanon, R.

Kuras, P.A. Lefebvre, S.D. Lemaire, A.V. Lobanov, M. Lohr, A. Manuell, I. Meier, L. Mets, M. Mittag, T. Mittelmeier, J.V. Moroney, J. Moseley, C. Napoli, A.M. Nedelcu, K. Niyogi, S.V. Novoselov, I.T. Paulsen, G. Pazour, S. Purton, J.P. Ral, D.M. Riano-Pachon, W. Riekhof, L. Rymarquis, M. Schroda, D. Stern, J. Umen, R. Willows, N. Wilson, S.L. Zimmer, J. Allmer, J. Balk, K. Bisova, C.J. Chen, M. Elias, K. Gendler, C. Hauser, M.R. Lamb, H. Ledford, J.C. Long, J. Minagawa, M.D. Page, J. Pan, W. Pootakham, S. Roje, A. Rose, E. Stahlberg, A.M. Terauchi, P. Yang, S. Ball, C. Bowler, C.L. Dieckmann, V.N. Gladyshev, P. Green, R. Jorgensen, S. Mayfield, B. Mueller-Roeber, S. Rajamani, R.T. Sayre, P. Brokstein, et al. 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science*. 318:245-250.

Mine, I., D. Menzel, and K. Okuda. 2008. Morphogenesis in giant-celled algae. *Int Rev Cell Mol Biol*. 266:37-83.

Nagai, R., and S. Fukui. 1981. Differential treatment ofAcetabularia with cytochalasin B and N-Ethylmaleimide with special reference to their effects on cytoplasmic streaming. *Protoplasma*. 109:79-89.

Nagai, R., and N. Kamiya. 1977. Differential treatment of Chara cells with cytochalasin B with special reference to its effect on cytoplasmic streaming. *Experimental Cell Research*. 108:231-237.

Nguyen, L.T., H.A. Schmidt, A. von Haeseler, and B.Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32:268-274.

Nishihama, R., T. Soyano, M. Ishikawa, S. Araki, H. Tanaka, T. Asada, K. Irie, M. Ito, M. Terada, H. Banno, Y. Yamazaki, and Y. Machida. 2002. Expansion of the cell plate in plant cytokinesis requires a kinesin-like protein/MAPKKK complex. *Cell*. 109:87-99.

Nishiyama, T., H. Sakayama, J. de Vries, H. Buschmann, D. Saint-Marcoux, K.K. Ullrich, F.B. Haas, L. Vanderstraeten, D. Becker, D. Lang, S. Vosolsobe, S. Rombauts, P.K.I. Wilhelmsson, P. Janitza, R. Kern, A. Heyl, F. Rumpler, L. Villalobos, J.M. Clay, R. Skokan, A. Toyoda, Y. Suzuki, H. Kagoshima, E. Schijlen, N. Tajeshwar, B. Catarino, A.J. Hetherington, A. Saltykova, C. Bonnot, H. Breuninger, A. Symeonidi, G.V. Radhakrishnan, F. Van Nieuwerburgh, D. Deforce, C. Chang, K.G. Karol, R. Hedrich, P. Ulvskov, G. Glockner, C.F. Delwiche, J. Petrasek, Y. Van de Peer, J. Friml, M. Beilby, L. Dolan, Y. Kohara, S. Sugano, A. Fujiyama, P.M. Delaux, M. Quint, G. Theissen, M. Hagemann, J. Harholt, C. Dunand, S. Zachgo, J. Langdale, F. Maumus, D. Van Der Straeten, S.B. Gould, and S.A. Rensing. 2018. The Chara Genome: Secondary Complexity and Implications for Plant Terrestrialization. *Cell*. 174:448-464 e424.

Niu, J., G. Wang, F. Lü, B. Zhou, and G. Peng. 2009. Characterization of a new lectin involved in the protoplast regeneration of Bryopsis hypnoides. *Chinese Journal of Oceanology and Limnology*. 27:502-512.

Noutahi, E., V. Calderon, M. Blanchette, N. El-Mabrouk, and B.F. Lang. 2019. Rapid Genetic Code Evolution in Green Algal Mitochondrial Genomes. *Molecular Biology and Evolution*. 36:766-783.

Padilla, J.R., L.M. Ferreira, and E.S. Folker. 2022. Nuclear movement in multinucleated cells. *Development*. 149.

Pak, J.Y., C. Solorzano, M. Arai, and T. Nitta. 1991. Two distinct steps for spontaneous

generation of subprotoplasts from a disintegrated bryopsis cell. *Plant Physiol*. 96:819-825.

Pertea, M., D. Kim, G.M. Pertea, J.T. Leek, and S.L. Salzberg. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc*. 11:1650-1667.

Polle, J.E.W., K. Barry, J. Cushman, J. Schmutz, D. Tran, L.T. Hathwaik, W.C. Yim, J. Jenkins, Z. McKie-Krisberg, S. Prochnik, E. Lindquist, R.B. Dockter, C. Adam, H. Molina, J. Bunkenborg, E. Jin, M. Buchheim, and J. Magnuson. 2017. Draft Nuclear Genome Sequence of the Halophilic and Beta-Carotene-Accumulating Green Alga Dunaliella salina Strain CCAP19/18. *Genome Announc*. 5.

Prochnik, S.E., J. Umen, A.M. Nedelcu, A. Hallmann, S.M. Miller, I. Nishii, P. Ferris, A. Kuo, T. Mitros, L.K. Fritz-Laylin, U. Hellsten, J. Chapman, O. Simakov, S.A. Rensing, A. Terry, J. Pangilinan, V. Kapitonov, J. Jurka, A. Salamov, H. Shapiro, J. Schmutz, J. Grimwood, E. Lindquist, S. Lucas, I.V. Grigoriev, R. Schmitt, D. Kirk, and D.S. Rokhsar. 2010. Genomic analysis of organismal complexity in the multicellular green alga Volvox carteri. *Science*. 329:223-226.

Ranallo-Benavidez, T.R., K.S. Jaron, and M.C. Schatz. 2020. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun*. 11:1432.

Ranjan, A., B.T. Townsley, Y. Ichihashi, N.R. Sinha, and D.H. Chitwood. 2015. An intracellular transcriptomic atlas of the giant coenocyte Caulerpa taxifolia. *PLoS Genet*. 11:e1004900.

Reddy, A.S., and I.S. Day. 2001. Kinesins in the Arabidopsis genome: a comparative analysis among eukaryotes. *BMC Genomics*. 2:2.

Repetti, S.I., C. Iha, K. Uthanumallian, C.J. Jackson, Y. Chen, C.X. Chan, and H. Verbruggen. 2022. Nuclear genome of a pedinophyte pinpoints genomic innovation and streamlining in the green algae. *New Phytol*. 233:2144-2154.

Repetti, S.I., C.J. Jackson, L.M. Judd, R.R. Wick, K.E. Holt, and H. Verbruggen. 2020. The inflated mitochondrial genomes of siphonous green algae reflect processes driving expansion of noncoding DNA and proliferation of introns. *PeerJ*. 8:e8273.

Richards, E.J., and F.M. Ausubel. 1988. Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. *Cell*. 53:127-136.

Saito, C., and T. Ueda. 2009. Chapter 4 Functions of RAB and SNARE Proteins in Plant Life. *In* International Review of Cell and Molecular Biology. Vol. 274. Academic Press. 183-233.

Salanoubat, M., K. Lemcke, M. Rieger, W. Ansorge, M. Unseld, B. Fartmann, G. Valle, H. Blocker, M. Perez-Alonso, B. Obermaier, M. Delseny, M. Boutry, L.A. Grivell, R. Mache, P. Puigdomenech, V. De Simone, N. Choisne, F. Artiguenave, C. Robert, P. Brottier, P. Wincker, L. Cattolico, J. Weissenbach, W. Saurin, F. Quetier, M. Schafer, S. Muller-Auer, C. Gabel, M. Fuchs, V. Benes, E. Wurmbach, H. Drzonek, H. Erfle, N. Jordan, S. Bangert, R. Wiedelmann, H. Kranz, H. Voss, R. Holland, P. Brandt, G. Nyakatura, A. Vezzi, M. D'Angelo, A. Pallavicini, S. Toppo, B. Simionati, A. Conrad, K. Hornischer, G. Kauer, T.H. Lohnert, G. Nordsiek, J. Reichelt, M. Scharfe, O. Schon, M. Bargues, J. Terol, J. Climent, P. Navarro, C. Collado, A. Perez-Perez, B. Ottenwalder, D. Duchemin, R. Cooke, M. Laudie, C. Berger-Llauro, B. Purnelle, D. Masuy, M. de Haan, A.C. Maarse, J.P. Alcaraz, A. Cottet, E. Casacuberta, A. Monfort, A. Argiriou, M. flores, R.

Liguori, D. Vitale, G. Mannhaupt, D. Haase, H. Schoof, S. Rudd, P. Zaccaria, H.W. Mewes, K.F. Mayer, S. Kaul, C.D. Town, H.L. Koo, L.J. Tallon, J. Jenkins, T. Rooney, M. Rizzo, A. Walts, T. Utterback, C.Y. Fujii, T.P. Shea, T.H. Creasy, B. Haas, R. Maiti, D. Wu, J. Peterson, et al. 2000. Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana. *Nature*. 408:820-822.

Sanderfoot, A. 2007. Increases in the Number of SNARE Genes Parallels the Rise of Multicellularity among the Green Plants. *Plant Physiology*. 144:6-17.

Sasabe, M., and Y. Machida. 2012. Regulation of organization and function of microtubules by the mitogen-activated protein kinase cascade during plant cytokinesis. *Cytoskeleton (Hoboken)*. 69:913-918.

Schulz, M.H., D.R. Zerbino, M. Vingron, and E. Birney. 2012. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics*. 28:1086-1092.

Shen, Z., A.R. Collatos, J.P. Bibeau, F. Furt, and L. Vidali. 2012. Phylogenetic analysis of the Kinesin superfamily from physcomitrella. *Front Plant Sci*. 3:230.

Shirae-Kurabayashi, M., T. Edzuka, M. Suzuki, and G. Goshima. 2022. Cell tip growth underlies injury response of marine macroalgae. *PLoS One*. 17:e0264827.

Siddiqui, N., and A. Straube. 2017. Intracellular cargo transport by kinesin-3 motors. *Biochemistry (Moscow)*. 82:803-815.

Sobue, K., Y. Fujio, and K. Kanda. 1988. Tumor promoter induces reorganization of actin filaments and calspectin (fodrin or nonerythroid spectrin) in 3T3 cells. *Proceedings of the National Academy of Sciences*. 85:482-486.

Stanke, M., and S. Waack. 2003. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*. 19 Suppl 2:ii215-225.

Steiner, A., K. Rybak, M. Altmann, H.E. McFarlane, S. Klaeger, N. Nguyen, E. Facher, A. Ivakov, G. Wanner, B. Kuster, S. Persson, P. Braun, M.T. Hauser, and F.F. Assaad. 2016. Cell cycle-regulated PLEIADE/AtMAP65-3 links membrane and microtubule dynamics during plant cytokinesis. *Plant J*. 88:531-541.

Suzuki, S., H. Yamaguchi, N. Nakajima, and M. Kawachi. 2018. Raphidocelis subcapitata (=Pseudokirchneriella subcapitata) provides an insight into genome evolution and environmental adaptations in the Sphaeropleales. *Sci Rep*. 8:8058.

Tabata, S., T. Kaneko, Y. Nakamura, H. Kotani, T. Kato, E. Asamizu, N. Miyajima, S. Sasamoto, T. Kimura, T. Hosouchi, K. Kawashima, M. Kohara, M. Matsumoto, A. Matsuno, A. Muraki, S. Nakayama, N. Nakazaki, K. Naruo, S. Okumura, S. Shinpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, S. Sato, M. de la Bastide, E. Huang, L. Spiegel, L. Gnoj, A. O'Shaughnessy, R. Preston, K. Habermann, J. Murray, D. Johnson, T. Rohlfing, J. Nelson, T. Stoneking, K. Pepin, J. Spieth, M. Sekhon, J. Armstrong, M. Becker, E. Belter, H. Cordum, M. Cordes, L. Courtney, W. Courtney, M. Dante, H. Du, J. Edwards, J. Fryman, B. Haakensen, E. Lamar, P. Latreille, S. Leonard, R. Meyer, E. Mulvaney, P. Ozersky, A. Riley, C. Strowmatt, C. Wagner-McPherson, A. Wollam, M. Yoakum, M. Bell, N. Dedhia, L. Parnell, R. Shah, M. Rodriguez, L.H. See, D. Vil, J. Baker, K. Kirchoff, K. Toth, L. King, A. Bahret, B. Miller, M. Marra, R. Martienssen, W.R. McCombie, R.K. Wilson, G. Murphy, I. Bancroft, G. Volckaert, R. Wambutt, A. Dusterhoft, W. Stiekema, T. Pohl, K.D. Entian, N. Terryn, N. Hartley, E. Bent, S. Johnson, S.A. Langham, B. McCullagh, J. Robben, B. Grymonprez, W. Zimmermann, U. Ramsperger, H. Wedler, et al. 2000. Sequence and analysis of

chromosome 5 of the plant Arabidopsis thaliana. *Nature*. 408:823-826.

Takahashi, K., S. Suzuki, H. Kawai-Toyooka, K. Yamamoto, T. Hamaji, R. Ootsuki, H. Yamaguchi, M. Kawachi, T. Higashiyama, and H. Nozaki. 2023. Reorganization of the ancestral sex-determining regions during the evolution of trioecy in Pleodorina starrii. *Commun Biol*. 6:590.

Tamura, K., K. Iwabuchi, Y. Fukao, M. Kondo, K. Okamoto, H. Ueda, M. Nishimura, and I. Hara-Nishimura. 2013. Myosin XI-i links the nuclear membrane to the cytoskeleton to control nuclear movement and shape in Arabidopsis. *Curr Biol*. 23:1776-1781.

Tanaka, H., M. Ishikawa, S. Kitamura, Y. Takahashi, T. Soyano, C. Machida, and Y. Machida. 2004. The AtNACK1/HINKEL and STUD/TETRASPORE/AtNACK2 genes, which encode functionally redundant kinesins, are essential for cytokinesis in Arabidopsis. *Genes Cells*. 9:1199-1211.

Tatewaki, M. 1973. Life cycle of Bryopsis（緑藻ハネモ及びオオハネモの生活史）. *Japanese Journal of Phycology (Japanese)*. 21:125-129.

Tatewaki, M., and K. Nagata. 1970. Surviving Protoplasts in Vitro and Their Development in Bryopsis. *Journal of Phycology*. 6:401-403.

Theologis, A., J.R. Ecker, C.J. Palm, N.A. Federspiel, S. Kaul, O. White, J. Alonso, H. Altafi, R. Araujo, C.L. Bowman, S.Y. Brooks, E. Buehler, A. Chan, Q. Chao, H. Chen, R.F. Cheuk, C.W. Chin, M.K. Chung, L. Conn, A.B. Conway, A.R. Conway, T.H. Creasy, K. Dewar, P. Dunn, P. Etgu, T.V. Feldblyum, J. Feng, B. Fong, C.Y. Fujii, J.E. Gill, A.D. Goldsmith, B. Haas, N.F. Hansen, B. Hughes, L. Huizar, J.L. Hunter, J. Jenkins, C. Johnson-Hopson, S. Khan, E. Khaykin, C.J. Kim, H.L. Koo, I. Kremenetskaia, D.B. Kurtz, A. Kwan, B. Lam, S. Langin-Hooper, A. Lee, J.M. Lee, C.A. Lenz, J.H. Li, Y. Li, X. Lin, S.X. Liu, Z.A. Liu, J.S. Luros, R. Maiti, A. Marziali, J. Militscher, M. Miranda, M. Nguyen, W.C. Nierman, B.I. Osborne, G. Pai, J. Peterson, P.K. Pham, M. Rizzo, T. Rooney, D. Rowley, H. Sakano, S.L. Salzberg, J.R. Schwartz, P. Shinn, A.M. Southwick, H. Sun, L.J. Tallon, G. Tambunga, M.J. Toriumi, C.D. Town, T. Utterback, S. Van Aken, M. Vaysberg, V.S. Vysotskaia, M. Walker, D. Wu, G. Yu, C.M. Fraser, J.C. Venter, and R.W. Davis. 2000. Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana. *Nature*. 408:816-820.

Turmel, M., and C. Lemieux. 2018. Chapter Six - Evolution of the Plastid Genome in Green Algae. *In* Advances in Botanical Research. Vol. 85. S.-M. Chaw and R.K. Jansen, editors. Academic Press. 157-193.

Turmel, M., C. Otis, and C. Lemieux. 2017. Divergent copies of the large inverted repeat in the chloroplast genomes of ulvophycean green algae. *Scientific Reports*. 7:994.

Umen, J., and M.D. Herron. 2021. Green Algal Models for Multicellularity. *Annual Review of Genetics*. 55:603-632.

Vahrenholz, C., G. Riemen, E. Pratje, B. Dujon, and G. Michaelis. 1993. Mitochondrial DNA of Chlamydomonas reinhardtii: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. *Curr Genet*. 24:241-247.

Vale, R.D. 2003. The molecular motor toolbox for intracellular transport. *Cell*. 112:467-480.

Varki, A., R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N.H. Packer, J.H. Prestegard, R.L. Schnaar, and P.H. Seeberger. 2022. *In* Essentials of Glycobiology, 4th Edition. Cold Spring Harbor Laboratory Press

1454 Copyright © 2022 by the Consortium of Glycobiology Editors, La Jolla, California.
1455 Published by Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New
1456 York. All rights reserved., Cold Spring Harbor (NY).
1457 Vidali, L., G.M. Burkart, R.C. Augustine, E. Kerdavid, E. Tuzel, and M. Bezanilla. 2010.
1458 Myosin XI is essential for tip growth in Physcomitrella patens. *Plant Cell*.
1459 22:1868-1882.
1460 Vugrek, O., H. Sawitzky, and D. Menzel. 2003. Class XIII myosins from the green alga
1461 Acetabularia: driving force in organelle transport and tip growth? *J Muscle Res*
1462 *Cell Motil*. 24:87-97.
1463 Walker, B.J., T. Abeel, T. Shea, M. Priest, A. Abouelliel, S. Sakthikumar, C.A. Cuomo,
1464 Q. Zeng, J. Wortman, S.K. Young, and A.M. Earl. 2014. Pilon: an integrated tool
1465 for comprehensive microbial variant detection and genome assembly
1466 improvement. *PLoS One*. 9:e112963.
1467 Wang, Y., L. Jia, G. Tian, Y. Dong, X. Zhang, Z. Zhou, X. Luo, Y. Li, and W. Yao. 2023.
1468 shinyCircos-V2.0: Leveraging the creation of Circos plot with enhanced usability
1469 and advanced features. *iMeta*. 2:e109.
1470 Waterhouse, R.M., M. Seppey, F.A. Simão, M. Manni, P. Ioannidis, G. Klioutchnikov,
1471 E.V. Kriventseva, and E.M. Zdobnov. 2017. BUSCO Applications from Quality
1472 Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and*
1473 *Evolution*. 35:543-548.
1474 Worden, A.Z., J.H. Lee, T. Mock, P. Rouze, M.P. Simmons, A.L. Aerts, A.E. Allen, M.L.
1475 Cuvelier, E. Derelle, M.V. Everett, E. Foulon, J. Grimwood, H. Gundlach, B.
1476 Henrissat, C. Napoli, S.M. McDonald, M.S. Parker, S. Rombauts, A. Salamov, P.
1477 Von Dassow, J.H. Badger, P.M. Coutinho, E. Demir, I. Dubchak, C. Gentemann,
1478 W. Eikrem, J.E. Gready, U. John, W. Lanier, E.A. Lindquist, S. Lucas, K.F.
1479 Mayer, H. Moreau, F. Not, R. Otillar, O. Panaud, J. Pangilinan, I. Paulsen, B.
1480 Piegu, A. Poliakov, S. Robbens, J. Schmutz, E. Toulza, T. Wyss, A. Zelensky, K.
1481 Zhou, E.V. Armbrust, D. Bhattacharya, U.W. Goodenough, Y. Van de Peer, and
1482 I.V. Grigoriev. 2009. Green evolution and dynamic adaptations revealed by
1483 genomes of the marine picoeukaryotes Micromonas. *Science*. 324:268-272.
1484 Wu, S.Z., and M. Bezanilla. 2014. Myosin VIII associates with microtubule ends and
1485 together with actin plays a role in guiding plant cell division. *Elife*. 3.
1486 Wu, S.Z., and M. Bezanilla. 2018. Actin and microtubule cross talk mediates persistent
1487 polarized growth. *J Cell Biol*. 217:3531-3544.
1488 Wu, S.Z., J.A. Ritchie, A.H. Pan, R.S. Quatrano, and M. Bezanilla. 2011. Myosin VIII
1489 regulates protonemal patterning and developmental timing in the moss
1490 Physcomitrella patens. *Mol Plant*. 4:909-921.
1491 Wu, T.D., and C.K. Watanabe. 2005. GMAP: a genomic mapping and alignment program
1492 for mRNA and EST sequences. *Bioinformatics*. 21:1859-1875.
1493 Yamada, M., and G. Goshima. 2018. The KCH Kinesin Drives Nuclear Transport and
1494 Cytoskeletal Coalescence to Promote Tip Cell Growth in Physcomitrella patens.
1495 *Plant Cell*. 30:1496-1510.
1496 Yamada, M., Y. Tanaka-Takiguchi, M. Hayashi, M. Nishina, and G. Goshima. 2017.
1497 Multiple kinesin-14 family members drive microtubule minus end-directed
1498 transport in plant cells. *J Cell Biol*. 216:1705-1714.
1499 Yau, S., M. Krasovec, L.F. Benites, S. Rombauts, M. Groussin, E. Vancaester, J.M. Aury,
1500 E. Derelle, Y. Desdevises, M.L. Escande, N. Grimsley, J. Guy, H. Moreau, S.

Sanchez-Brosseau, Y. Van de Peer, K. Vandepoele, S. Gourbiere, and G. Piganeau. 2020. Virus-host coexistence in phytoplankton through the genomic lens. *Sci Adv*. 6:eaay2587.

Yoon, K.S., K.P. Lee, T.A. Klochkova, and G.H. Kim. 2008. Molecular Characterization of the Lectin, Bryohealin, Involved in Protoplast Regeneration of the Marine Alga Bryopsis Plumosa (Chlorophyta)(1). *J Phycol*. 44:103-112.

Yoshida, M.W., M. Hakozaki, and G. Goshima. 2023. Armadillo repeat-containing kinesin represents the versatile plus-end-directed transporter in Physcomitrella. *Nat Plants*. 9:733-748.

Yoshida, M.W., M. Yamada, and G. Goshima. 2019. Moss Kinesin-14 KCBP Accelerates Chromatid Motility in Anaphase. *Cell Struct Funct*. 44:95-104.

Zan, J., Z. Li, M.D. Tianero, J. Davis, R.T. Hill, and M.S. Donia. 2019. A microbial factory for defensive kahalalides in a tripartite marine symbiosis. *Science*. 364.

Zheng, F., H. Liu, M. Jiang, Z. Xu, Z. Wang, C. Wang, F. Du, Z. Shen, and B. Wang. 2018. The complete mitochondrial genome of the Caulerpa lentillifera (Ulvophyceae, Chlorophyta): Sequence, genome content, organization structure and phylogenetic consideration. *Gene*. 673:225-238.

Zhou, L., L. Wang, J. Zhang, C. Cai, and P. He. 2016. Complete mitochondrial genome of Ulva prolifera, the dominant species of green macroalgal blooms in Yellow Sea, China. *Mitochondrial DNA B Resour*. 1:76-78.

Figure legends

Figure 1. Life cycle and regeneration of *Bryopsis* collected on Sugashima Island

(A) Life cycle of *Bryopsis*. Images are derived from *Bryopsis* sp. analysed in this study. Sequencing indicates a SNP in male and female lines (contig 3: nt 4124748–4124766). Note that both A and G were detected in the sporophyte (diploid). DAPI-stained (magenta) nuclei are shown in the middle. (B) Regeneration of *Bryopsis* sp. after extrusion of the cytoplasm into autoclaved seawater. See also Movie 1. Arrowheads indicate polarised tip growth of regenerated cells.

Figure 2. Nuclear and organelle genome assembly

(A) Circos plot of the 27 contigs and organelles assembled from *Bryopsis* sp. (From outmost to innermost lanes) (1) Contigs (cyan) and putative telomeric repeats (red bar, CCCTAAA) are shown. When the repeat was identified in both ends of the contig, the contig number was indicated in red. When just one end had the repeat, the contig was highlighted with a black bold letter. Blue bars indicate organelles of circular genome (mitochondrion: Mt, chloroplast: Cp). (2) Purple lines indicate G/C content per 10,000 bp. Two grey lines indicate 25% and 75%. (3) Black bars present non-telomeric repeat sequences. (4) Red and blue bars indicate genes from Watson and Crick strands, respectively. (5) Genes analysed in this study. (B) Phylogenetic tree of green algal species subjected to KEGG analysis in this study. Maximum Likelihood (ML) tree was constructed with LG+F+R4 selected as the best-fit model and the branch support was

estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

Figure 3. Massive duplication of BPL-1/Bryohealin in *Bryopsis* sp.
(A) Phylogenetic tree of BPL-1 proteins in green algae. Only partial sequences were available for *Bryopsis hypnoides* ACB47462.1 and *Pleodorina starrii* GCL49965.1, and therefore these were not included in the tree. ML tree was constructed with WAG+G4 selected as the best-fit model and the branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acids substitutions per site. (B) Alignment of amino acid sequences of BPL-1/Bryohealin of *Bryopsis* species. Asterisks indicate highly conserved residues.

Figure 4. Myosin and kinesin motors in *Bryopsis* sp.
(A) Repertoire of motors potentially involved in cargo transport and cytokinesis. Note that the number might be underestimated in some species, as the genome (RNA) coverage is not complete. (B) Schematic presentation of myosin motors. (C) Divergence of ARK-type motors in green plants. In case some species possess the motor but others in the same family do not, dotted lines were used.

Figure 5. Microtubule-dependent, but actin-independent, bidirectional motility of chloroplasts in *Bryopsis* sp.
(A) (Left) Device used for time-lapse imaging. (Right) Magnified view of the specimen (thalli) and a piece of net on the glass. (B) Time-lapse imaging of autofluorescent chloroplasts in the control DMSO-treated cell. Yellow and red arrowheads indicate unidirectional and bidirectional movement, respectively. Time is shown as min:sec. (C) Kymograph images of chloroplast motility in the presence or absence of microtubules or actin. Arrow indicates a point of directional switch. Horizontal bar, 10 μm; vertical bar, 120 s. (D) Rate of chloroplast motility. The mean rate was 339 ± 18 nm/s (control DMSO, ± SEM, n = 50), 45 ± 5 nm/s (+ oryzalin, ± SEM, n = 50), 369 ± 28 nm/s (+ latrunculin A [LatA], ± SEM, n = 50). P-values were calculated using a two-sided ART ANOVA; $P < 0.0001$ (control [DMSO] – oryzalin), $P < 0.0001$ (oryzalin – latrunculin A), $P = 0.7790$ (control [DMSO] – latrunculin A).

Figure S1. Microtubule and actin organisation in the cytoplasm
(A) Suppression of aggregation of the cytoplasmic extract by N-acetyl-D-glucosamine. Glucose was used as the control. (B) (Top) Schematic representation of the focal plane in microscopy. (Bottom) Three images acquired with 637 nm laser, each representing top, middle, or bottom section of the main axis. Autofluorescent chloroplasts are visualised. A large vacuole occupies the majority of the middle section. (C–E) Immmunostaining of

microtubules and actin filaments in the main axis of thalli in the presence or absence of oryzalin (10 μM) or latrunculin A (LatA, 10 μM). The control sample was treated with DMSO. Boxed regions are magnified on the right.

Figure S2. Phylogenetic tree based on mitochondrial genes

Bryopsis sp. formed a clade with other Bryopsidales species, whereas the registered '*Bryopsis pulmosa*' sequences (MN853874.1) were most similar to Ulvales sequences. ML gene tree was drawn using IQ-TREE v1.6.12 with LG+F+R4 selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

Figure S3. Overrepresenting gene pathway in Bryopsidales

(A) Number of the genes in 'MAPK signaling pathway - plant (KO04016)'. (B) Signal transduction pathway known in land plants. Figures are derived from 'MAPK signaling pathway - plant (KO04016)' in KEGG.

Figure S4. Phylogenetic tree of *BPL-2, 3, 4* genes

ML tree was drawn using IQ-TREE v1.6.12 with WAG+G4 (BPL-2) or LG+G4 (BPL-3/4) selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

Figure S5. Phylogenetic tree of myosin of green algae

ML tree was also drawn using IQ-TREE v1.6.12 with LG+I+G4 selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

Figure S6. Phylogenetic tree of the kinesin superfamily of green algae

Each page contains trees of a few kinesin subfamilies. Kinesin-GA is alga-specific subfamily. ML tree was also drawn using IQ-TREE v1.6.12 with LG+I+G4 and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

Figure S7. Structure of *cox1* gene encoded in the mitochondrial genome

Several ORFs were identified in the intron of *cox1* gene in *Bryopsis* sp. .

Figure S8. Length of intron in the mitochondrial genome

N = 72, 47, 29, 10, 18, 18, 26 (from left to right).

Figure S9. UGA codon likely encodes tryptophan in the mitochondrial genome

Based on the amino acid sequences of the Nad5 protein (this figure) and other conserved

proteins in green algae, the UGA of *Bryopsis* sp. likely represents a tryptophan codon, not a termination codon, in the mitochondrial genome.

Movie legends

Movie 1. Protoplast formation from extruded cytoplasm

Images were acquired using a stereomicroscope every 20 s immediately after the extrusion of the cytoplasm into seawater.

Movie 2. Chloroplast motility in the presence or absence of oryzalin or latrunculin A

Images were acquired every 10 s using a spinning-disc confocal microscope and a 40× 0.95 NA objective lens. Drugs or control DMSO were added at 2 min.

Supplementary tables

Table S1. Comparison of the genomes of green algae and land plant species.

Table S2. Genome and transcriptome data used in the comparative analysis.

Table S3. Number of unigenes based on KEGG pathway annotation.

Table S4. Number of genes in each species.

Table S5. BUSCO values after transcriptome assembly for Dasycladales and Cladophorales.

Table S6. Transcriptome results in the side branch, main axis and rhizoid.

Table S7. Comparison of the chloroplast genome of Chloroplastida including *Bryopsis*.

Table S8. Comparison of protein coding and ribosomal RNA genes encoded in the chloroplast genomes of Chloroplastida including Bryopsis.

Table S9. Comparison of the mitochondrial genome of Chloroplastida including *Bryopsis*.

Table S10. Genes encoded in the mitochondrial genome of Chloroplastida including *Bryopsis*.

Table S11. Protein-coding genes found on the intron of other genes in the mitochondrial genome.

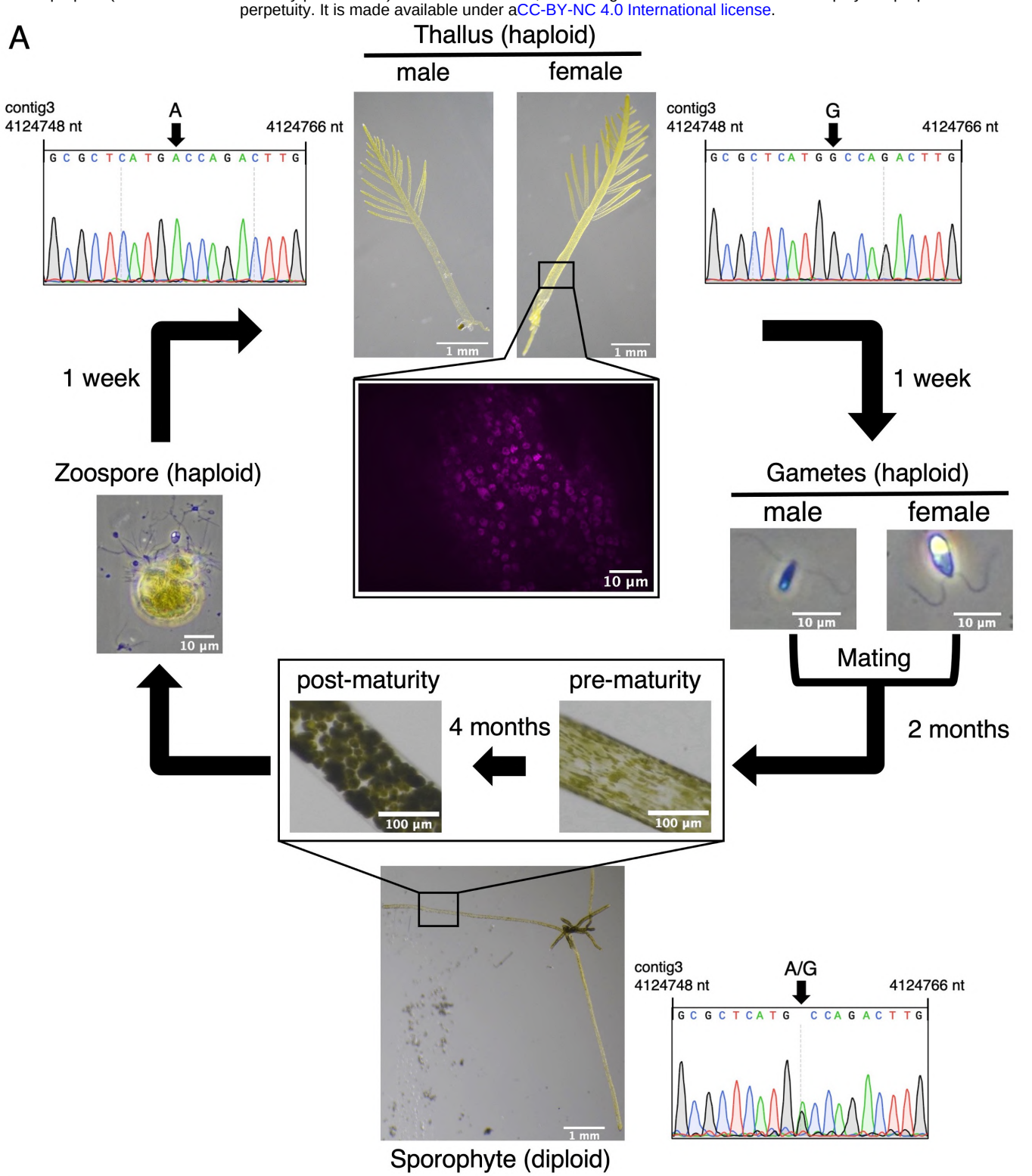


Figure 1. Life cycle and regeneration of *Bryopsis* collected on Sugashima Island

(A) Life cycle of *Bryopsis*. Images are derived from *Bryopsis* sp. analysed in this study. Sequencing indicates a SNP in male and female lines (contig 3: nt 4124748–4124766). Note that both A and G were detected in the sporophyte (diploid). DAPI-stained (magenta) nuclei are shown in the middle. (B) Regeneration of *Bryopsis* sp. after extrusion of the cytoplasm into autoclaved seawater. See also Movie 1. Arrowheads indicate polarised tip growth of regenerated cells.

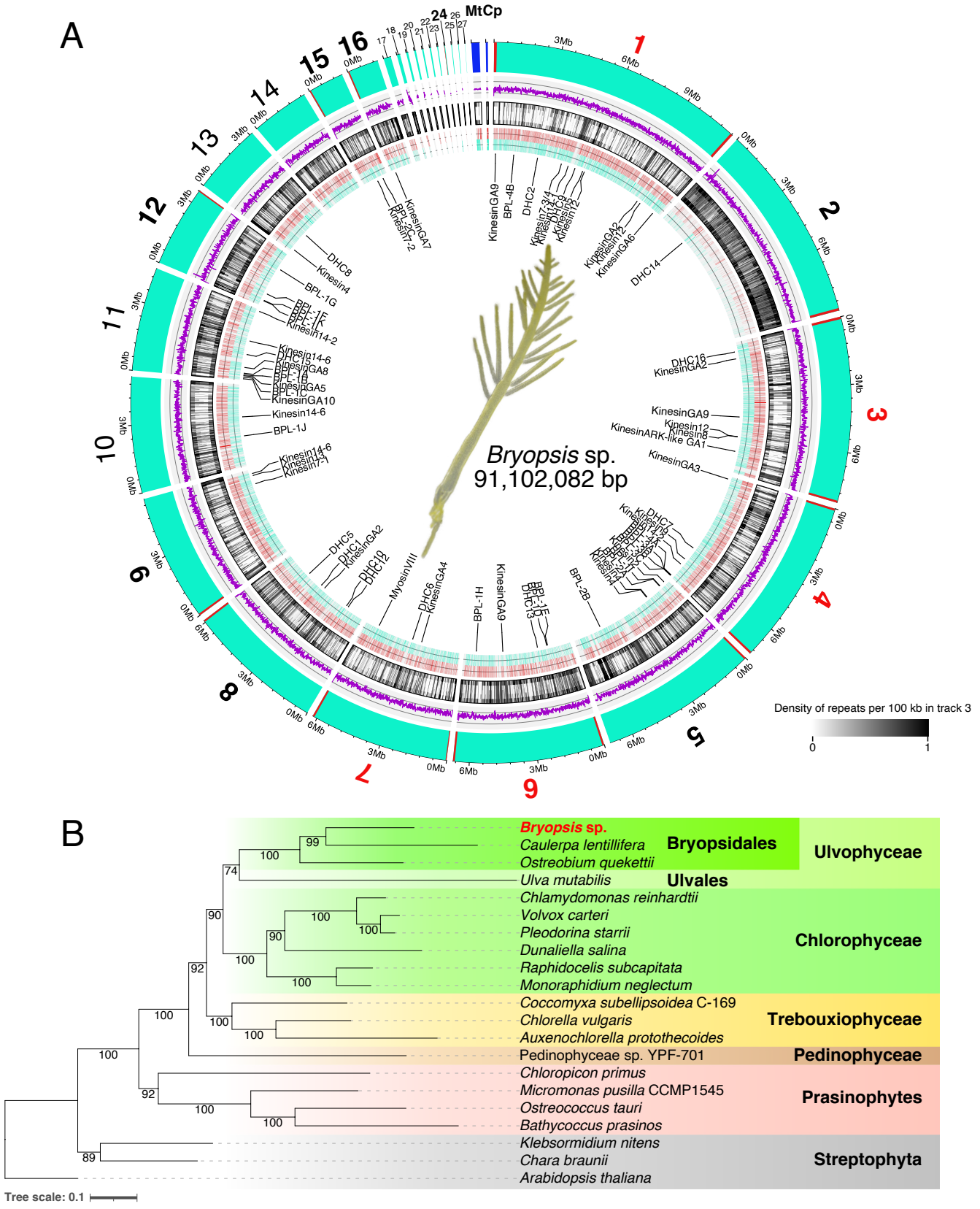
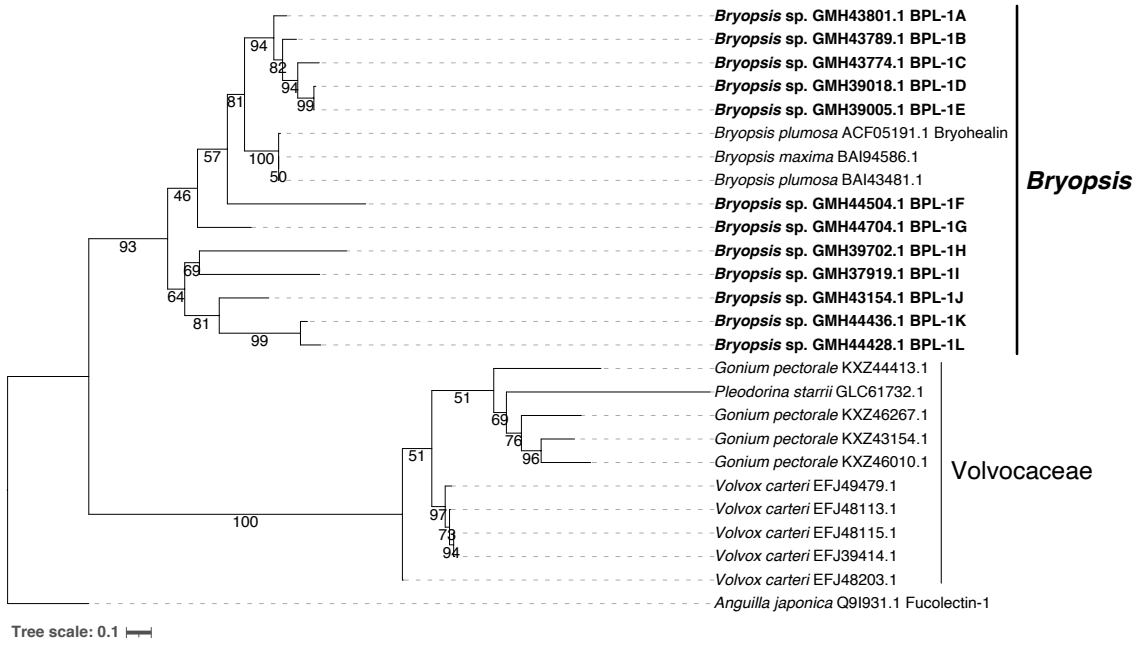


Figure 2. Nuclear and organelle genome assembly

(A) Circos plot of the 27 contigs and organelles assembled from *Bryopsis* sp. (From outmost to innermost lanes) (1) Contigs (cyan) and putative telomeric repeats (red bar, CCCTAAA) are shown. When the repeat was identified in both ends of the contig, the contig number was indicated in red. When just one end had the repeat, the contig was highlighted with a black bold letter. Blue bars indicate organelles of circular genome (mitochondrion: Mt, chloroplast: Cp). (2) Purple lines indicate G/C content per 10,000 bp. Two grey lines indicate 25% and 75%. (3) Black bars present non-telomeric repeat sequences. (4) Red and blue bars indicate genes from Watson and Crick strands, respectively. (5) Genes analysed in this study. (B) Phylogenetic tree of green algal species subjected to KEGG analysis in this study. Maximum Likelihood (ML) tree was constructed with LG+F+R4 selected as the best-fit model and the branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

A



B

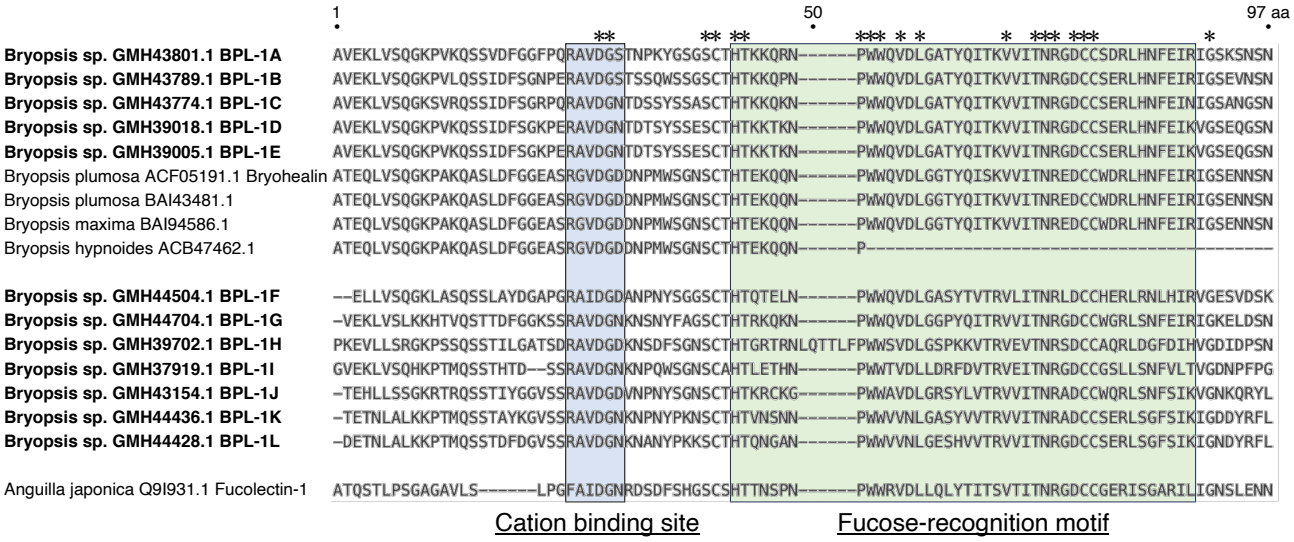


Figure 3. Massive duplication of BPL-1/Bryohealin in *Bryopsis* sp.
(A) Phylogenetic tree of BPL-1 proteins in green algae. Only partial sequences were available for *Bryopsis hypnoides* ACB47462.1 and *Pleodorina starrii* GCL49965.1, and therefore these were not included in the tree. ML tree was constructed with WAG+G4 selected as the best-fit model and the branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acids substitutions per site. (B) Alignment of amino acid sequences of BPL-1/Bryohealin of *Bryopsis* species. Asterisks indicate highly conserved residues.

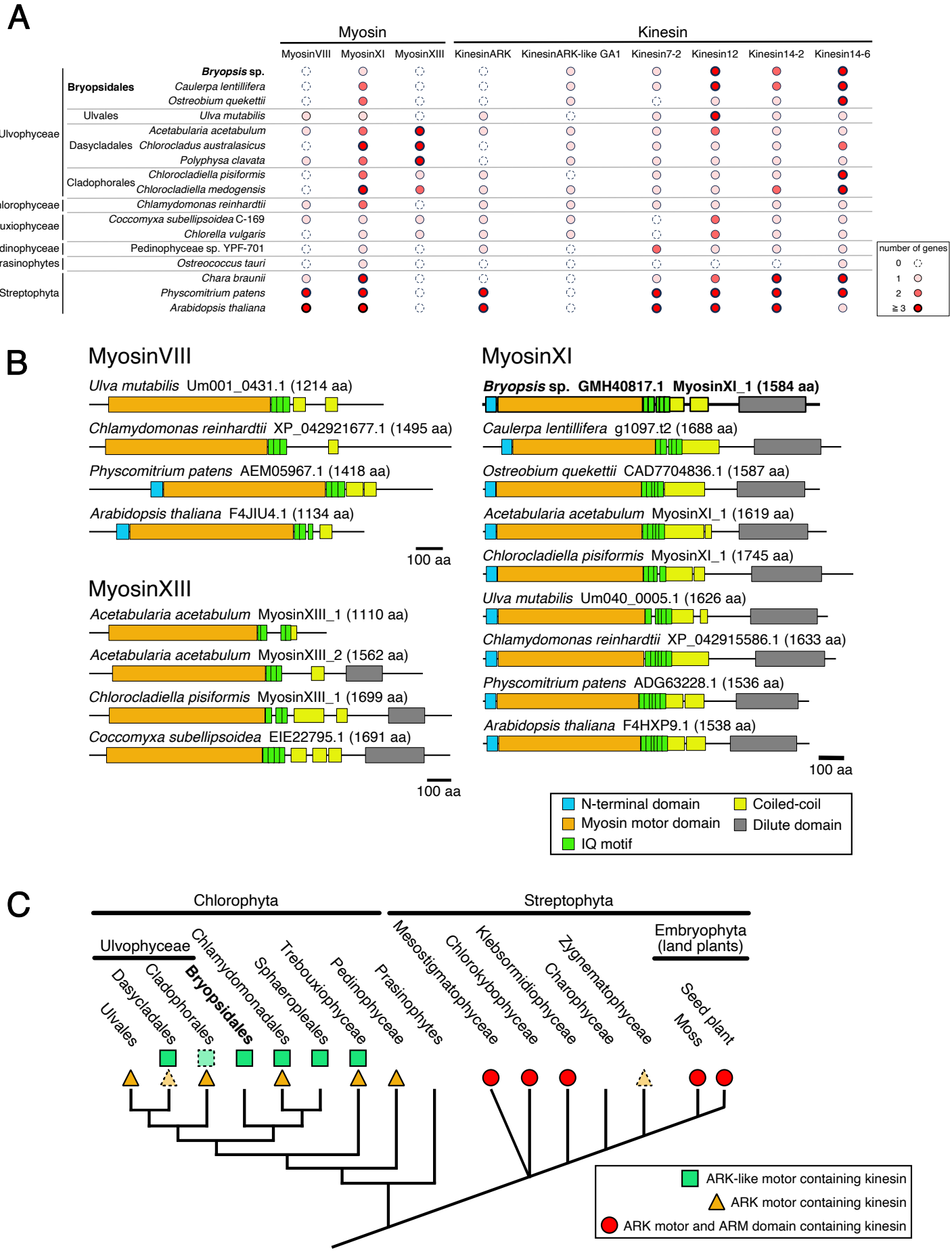


Figure 4. Myosin and kinesin motors in *Bryopsis* sp.
(A) Repertoire of motors potentially involved in cargo transport and cytokinesis. Note that the number might be underestimated in some species, as the genome (RNA) coverage is not complete. (B) Schematic presentation of myosin motors. (C) Divergence of ARK-type motors in green plants. In case some species possess the motor but others in the same family do not, dotted lines were used.

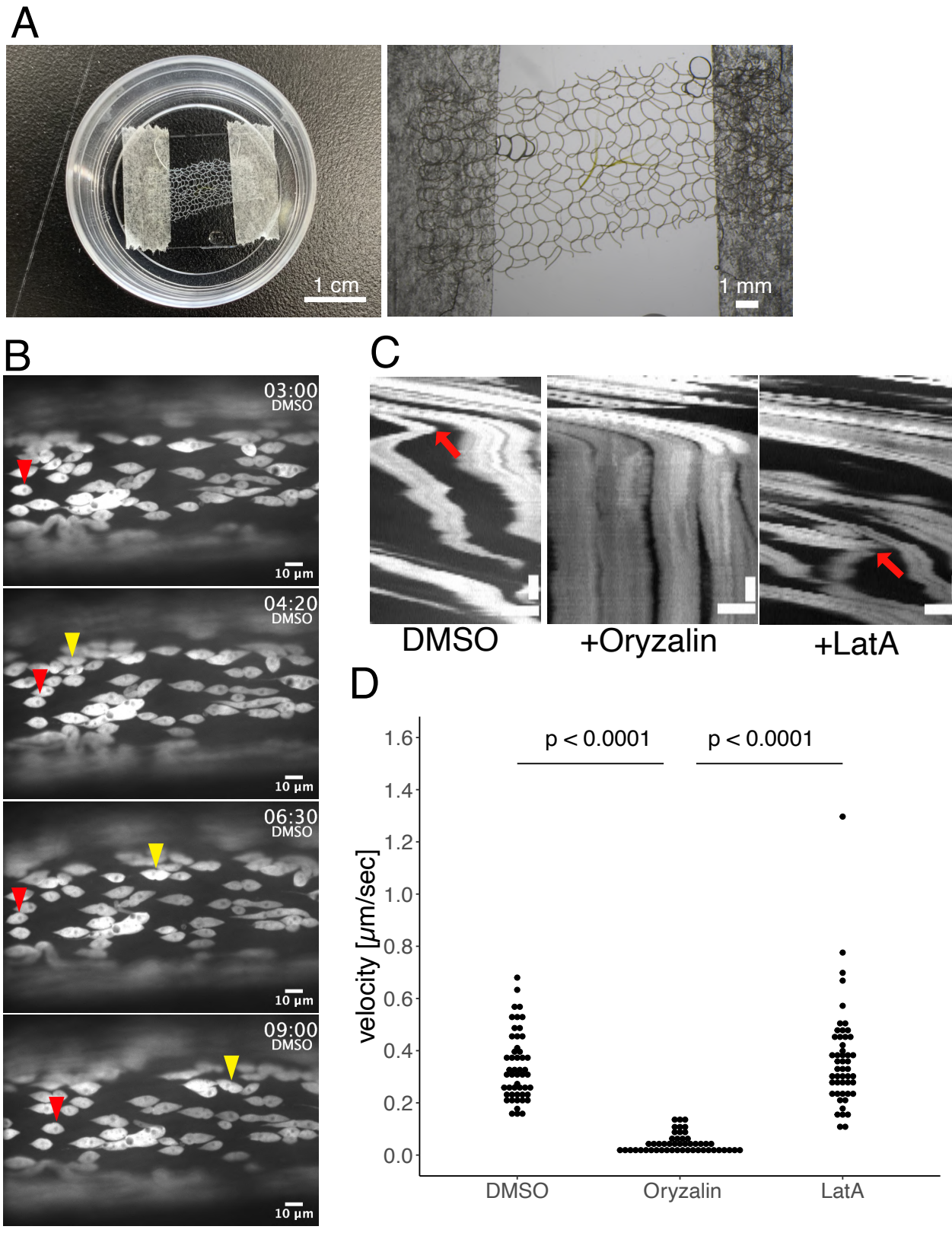


Figure 5. Microtubule-dependent, but actin-independent, bidirectional motility of chloroplasts in *Bryopsis* sp.

(A) (Left) Device used for time-lapse imaging. (Right) Magnified view of the specimen (thalli) and a piece of net on the glass. (B) Time-lapse imaging of autofluorescent chloroplasts in the control DMSO-treated cell. Yellow and red arrowheads indicate unidirectional and bidirectional movement, respectively. Time is shown as min:sec. (C) Kymograph images of chloroplast motility in the presence or absence of microtubules or actin. Arrow indicates a point of directional switch. Horizontal bar, 10 µm; vertical bar, 120 s. (D) Rate of chloroplast motility. The mean rate was 339 ± 18 nm/s (control DMSO, ± SEM, $n = 50$), 45 ± 5 nm/s (+ oryzalin, ± SEM, $n = 50$), 369 ± 28 nm/s (+ latrunculin A [LatA], ± SEM, $n = 50$). P-values were calculated using a two-sided ART ANOVA; $P < 0.0001$ (control [DMSO] – oryzalin), $P < 0.0001$ (oryzalin – latrunculin A), $P = 0.7790$ (control [DMSO] – latrunculin A).

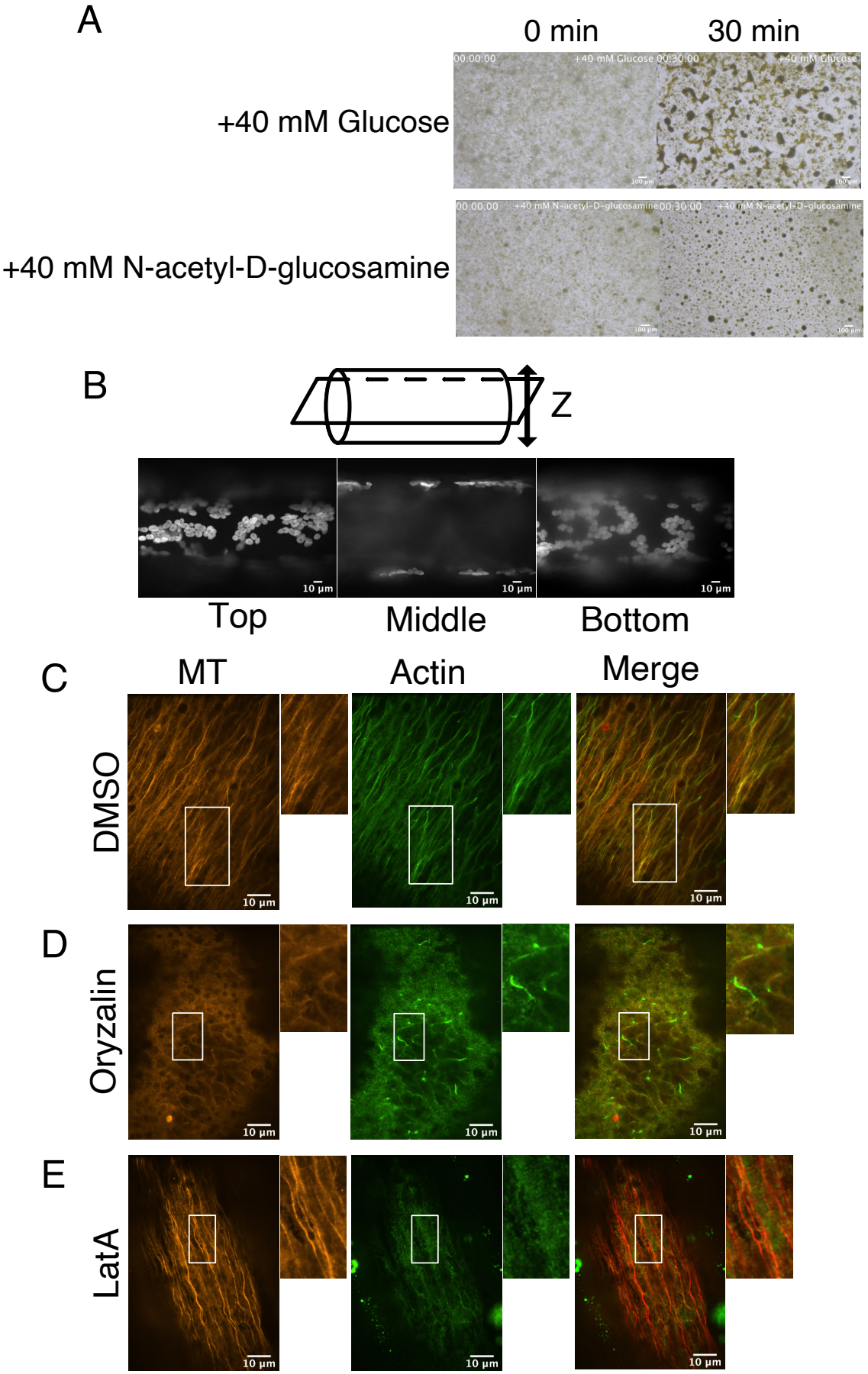


Figure S1. Microtubule and actin organisation in the cytoplasm

(A) Suppression of aggregation of the cytoplasmic extract by N-acetyl-D-glucosamine. Glucose was used as the control. (B) (Top) Schematic representation of the focal plane in microscopy. (Bottom) Three images acquired with 637 nm laser, each representing top, middle, or bottom section of the main axis. Autofluorescent chloroplasts are visualised. A large vacuole occupies the majority of the middle section. (C–E) Immmunostaining of microtubules and actin filaments in the main axis of thalli in the presence or absence of oryzalin (10 µM) or latrunculin A (LatA, 10 µM). The control sample was treated with DMSO. Boxed regions are magnified on the right.

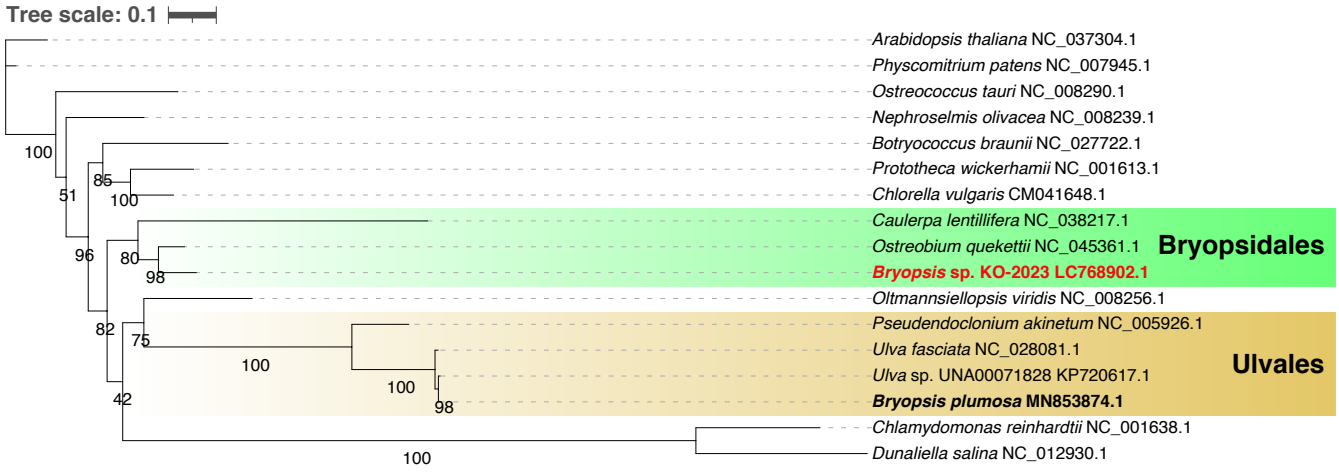
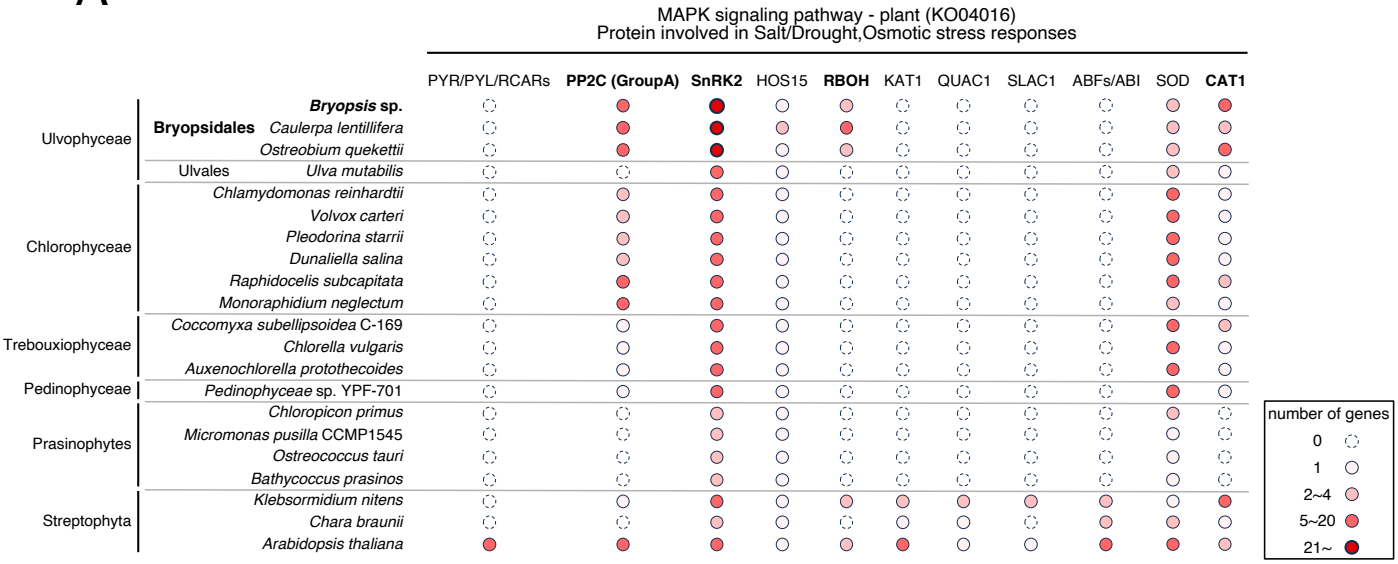


Figure S2. Phylogenetic tree based on mitochondrial genes

Bryopsis sp. formed a clade with other Bryopsidales species, whereas the registered '*Bryopsis pulmosa*' sequences (MN853874.1) were most similar to Ulvales sequences. ML gene tree was drawn using IQ-TREE v1.6.12 with LG+F+R4 selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

A



B

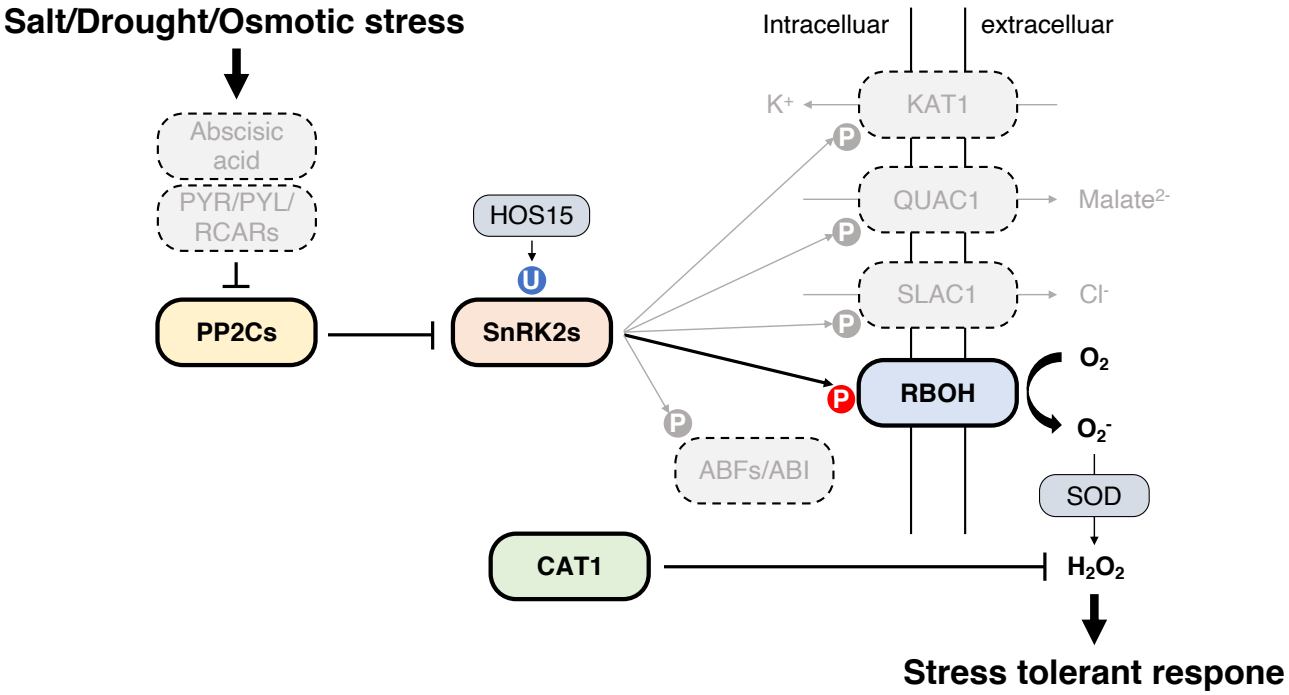


Figure S3. Overrepresenting gene pathway in Bryopsidales
(A) Number of the genes in 'MAPK signaling pathway - plant (KO04016)'. (B) Signal transduction pathway known in land plants. Figures are derived from 'MAPK signaling pathway - plant (KO04016)' in KEGG.

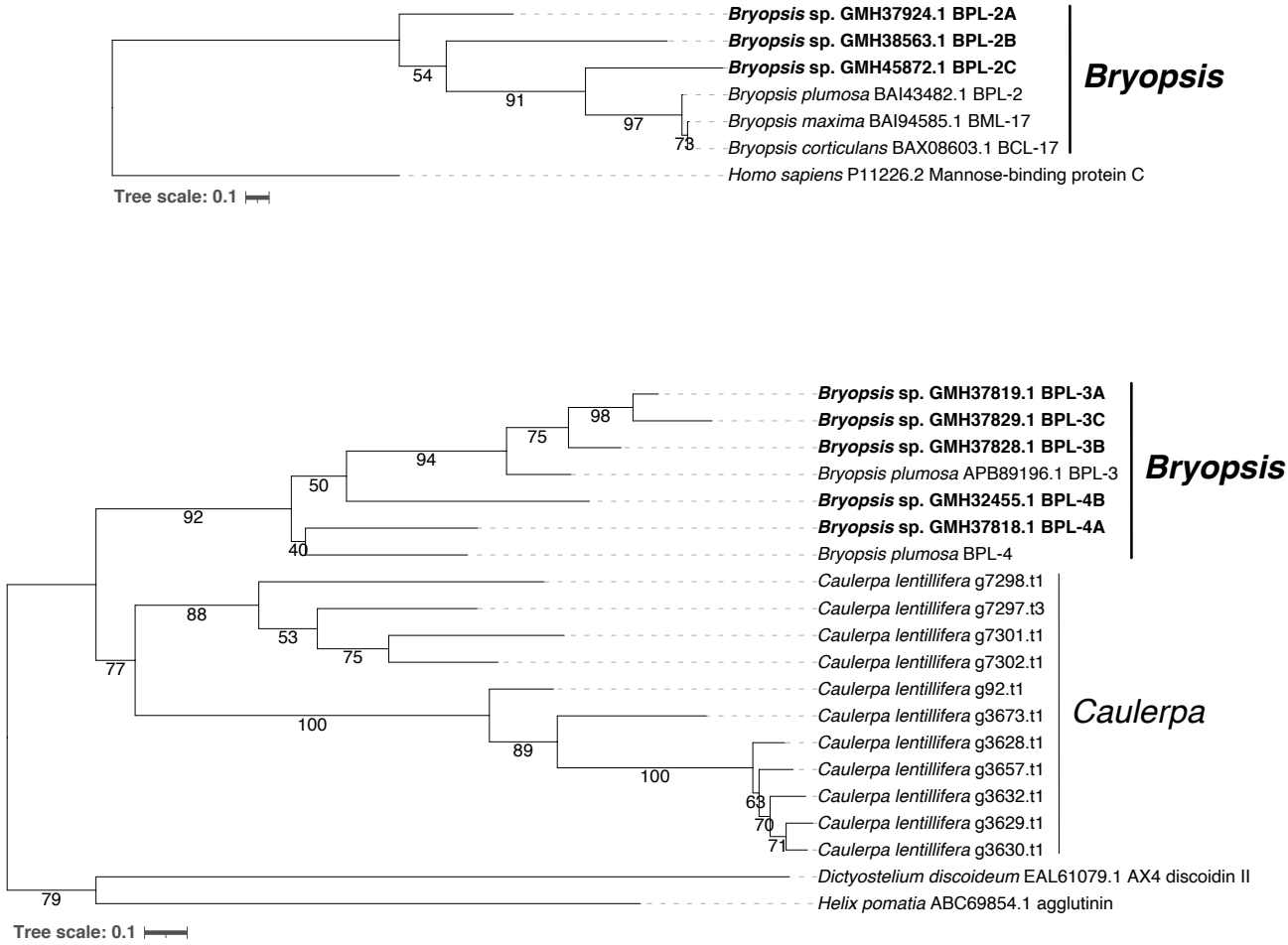


Figure S4. Phylogenetic tree of *BPL-2, 3, 4* genes

ML tree was drawn using IQ-TREE v1.6.12 with WAG+G4 (BPL-2) or LG+G4 (BPL-3/4) selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

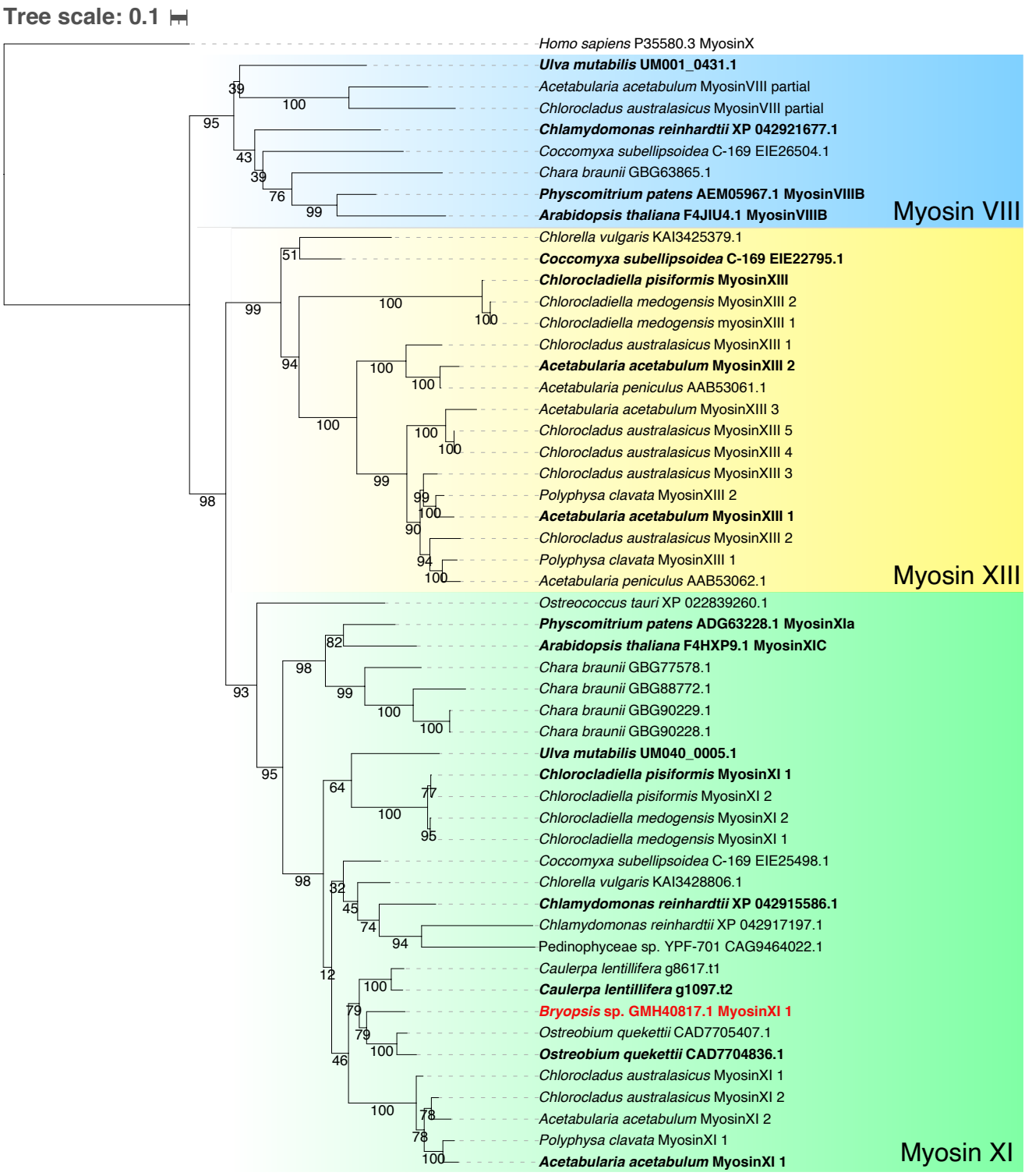


Figure S5. Phylogenetic tree of myosin of green algae
ML tree was also drawn using IQ-TREE v1.6.12 with LG+I+G4 selected as the best-fit model and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

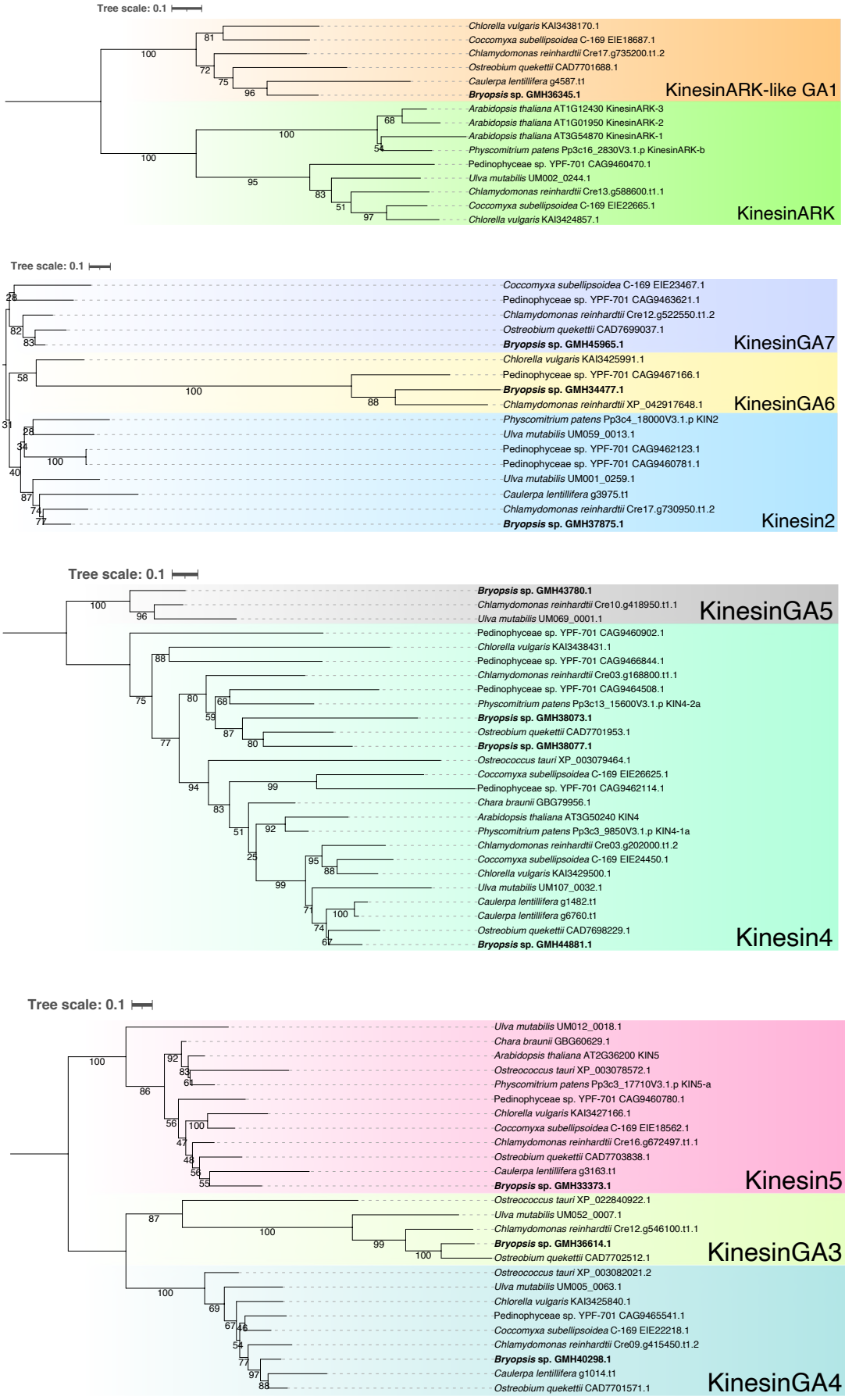


Figure S6.1 Phylogenetic tree of the kinesin superfamily of green algae
Each page contains trees of a few kinesin subfamilies. Kinesin-GA is alga-specific subfamily. ML tree was also drawn using IQ-TREE v1.6.12 with LG+I+G4 and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

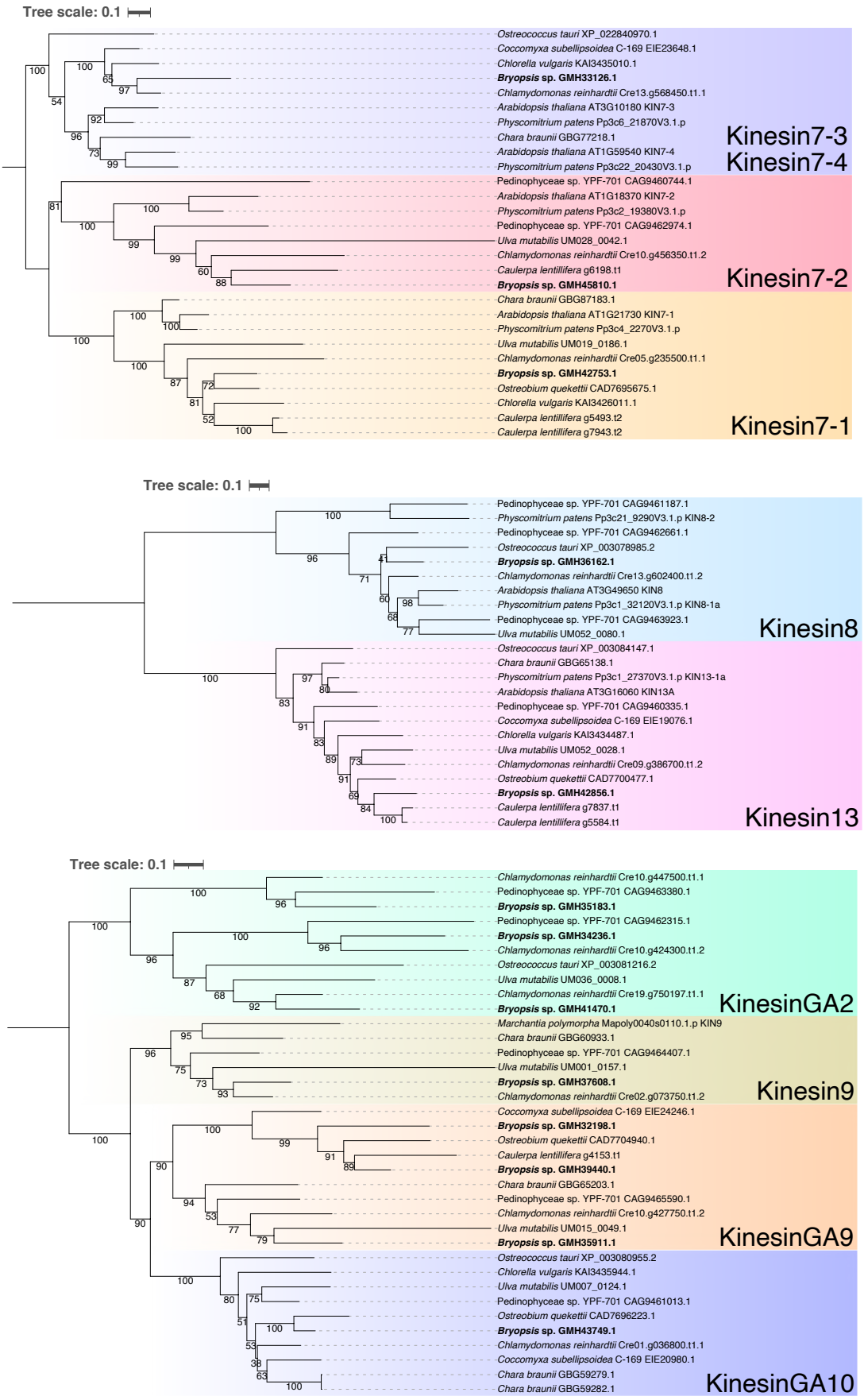


Figure S6.2 Phylogenetic tree of the kinesin superfamily of green algae
Each page contains trees of a few kinesin subfamilies. Kinesin-GA is alga-specific subfamily. ML tree was also drawn using IQ-TREE v1.6.12 with LG+I+G4 and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

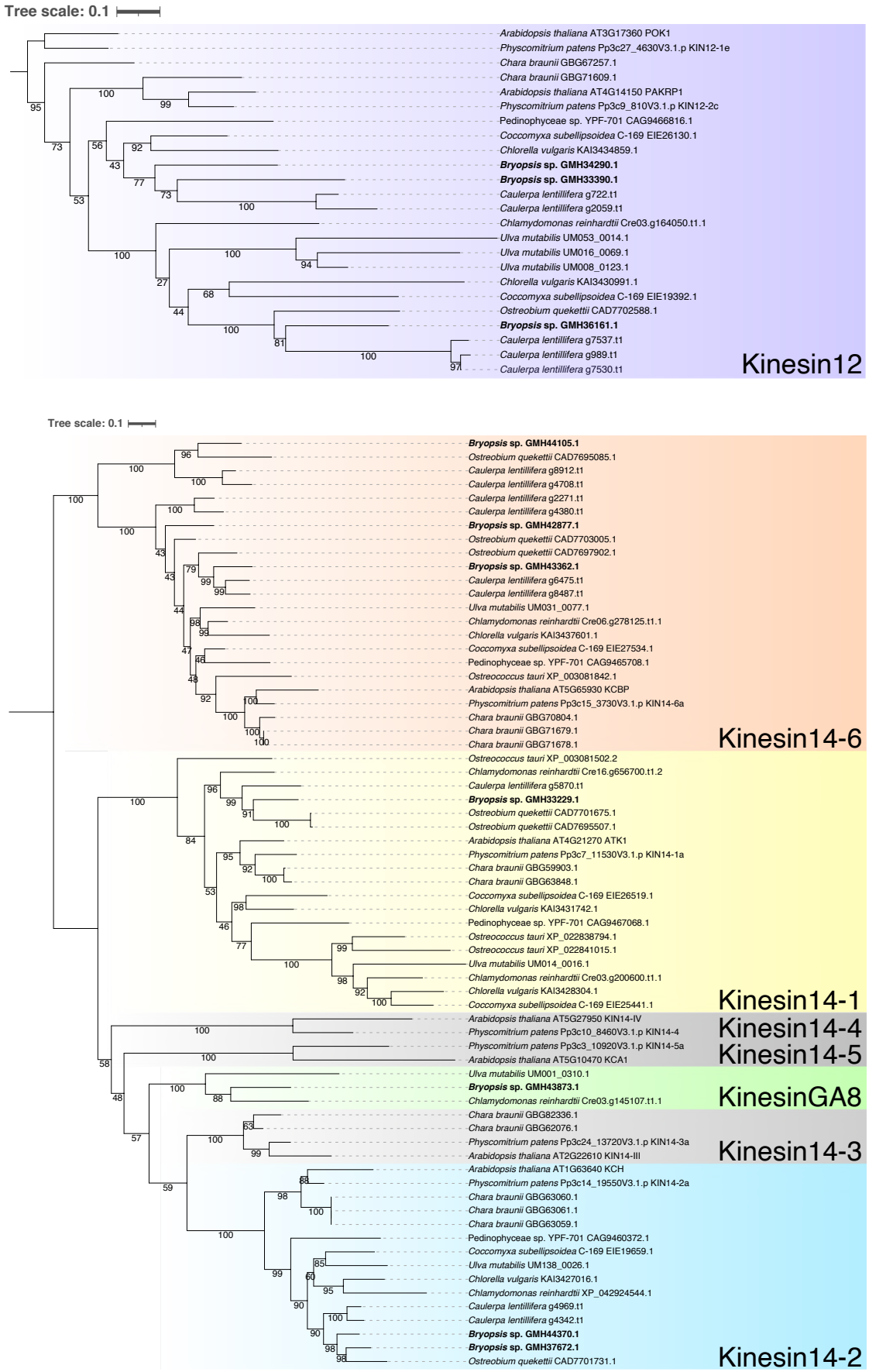


Figure S6.3 Phylogenetic tree of the kinesin superfamily of green algae
Each page contains trees of a few kinesin subfamilies. Kinesin-GA is alga-specific subfamily. ML tree was also drawn using IQ-TREE v1.6.12 with LG+I+G4 and branch support was estimated with 1,000 ultrafast bootstrap. The bar indicates 0.1 amino acid substitutions per site.

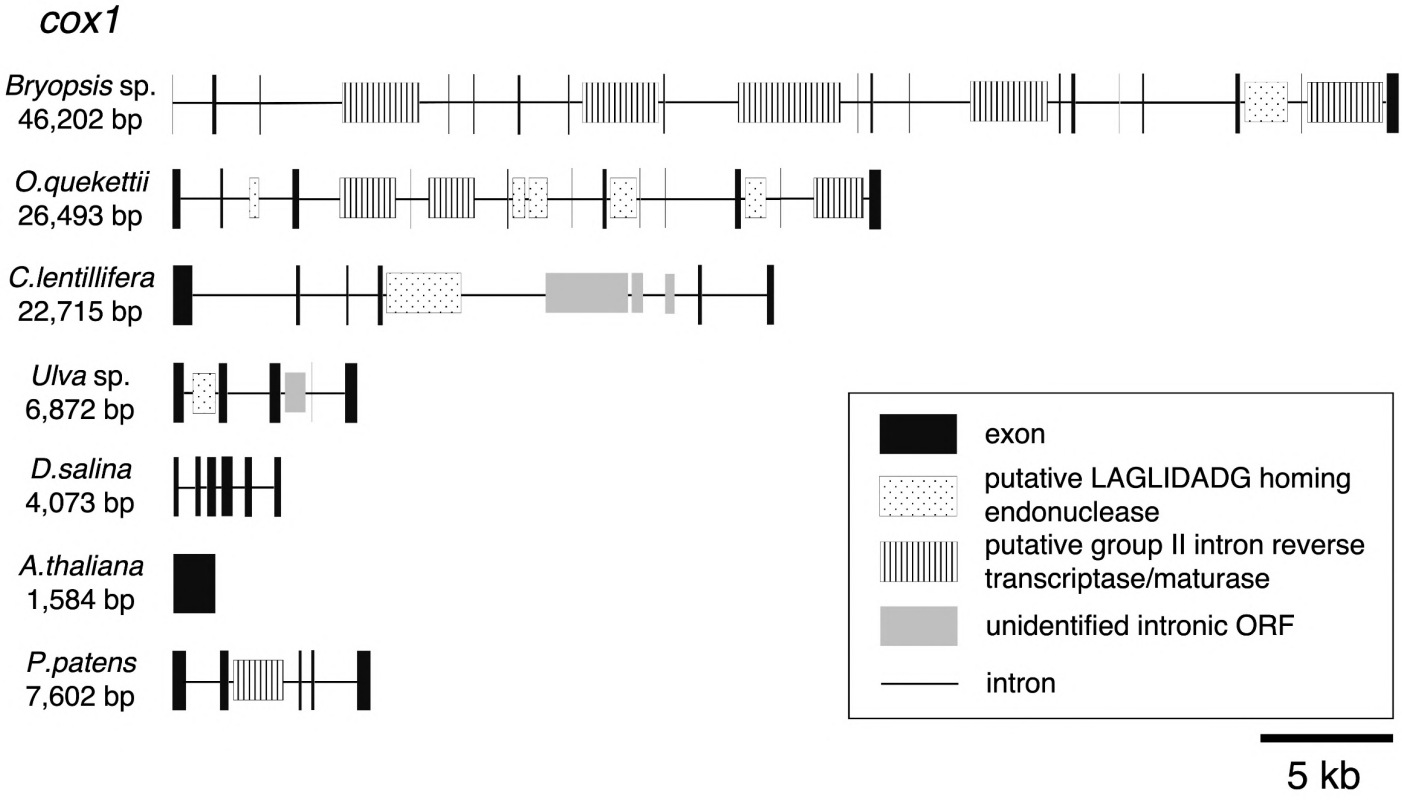


Figure S7. Structure of *cox1* gene encoded in the mitochondrial genome
Several ORFs were identified in the intron of *cox1* gene in *Bryopsis* sp..

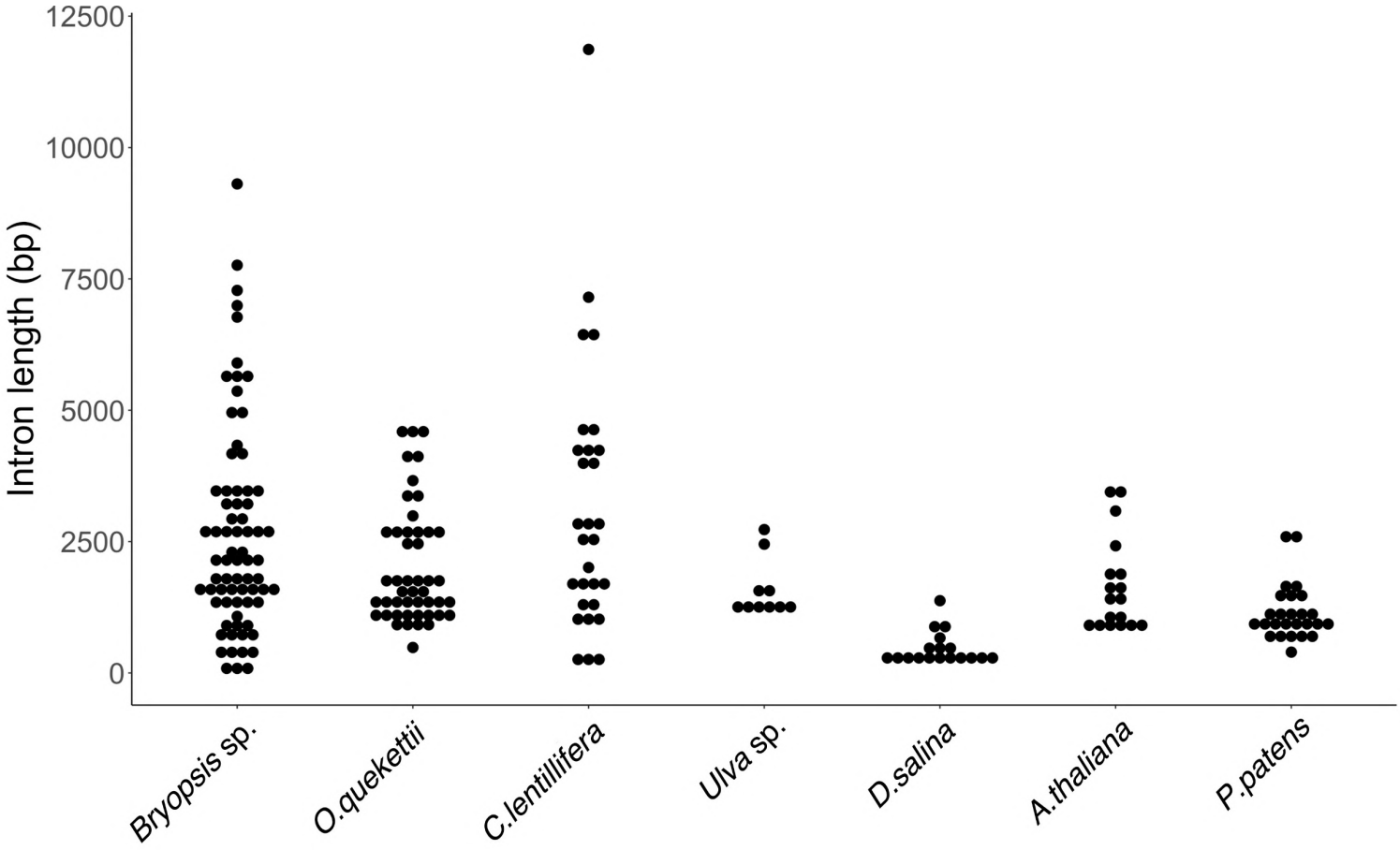


Figure S8. Length of intron in the mitochondrial genome

N = 72, 47, 29, 10, 18, 18, 26 (from left to right).

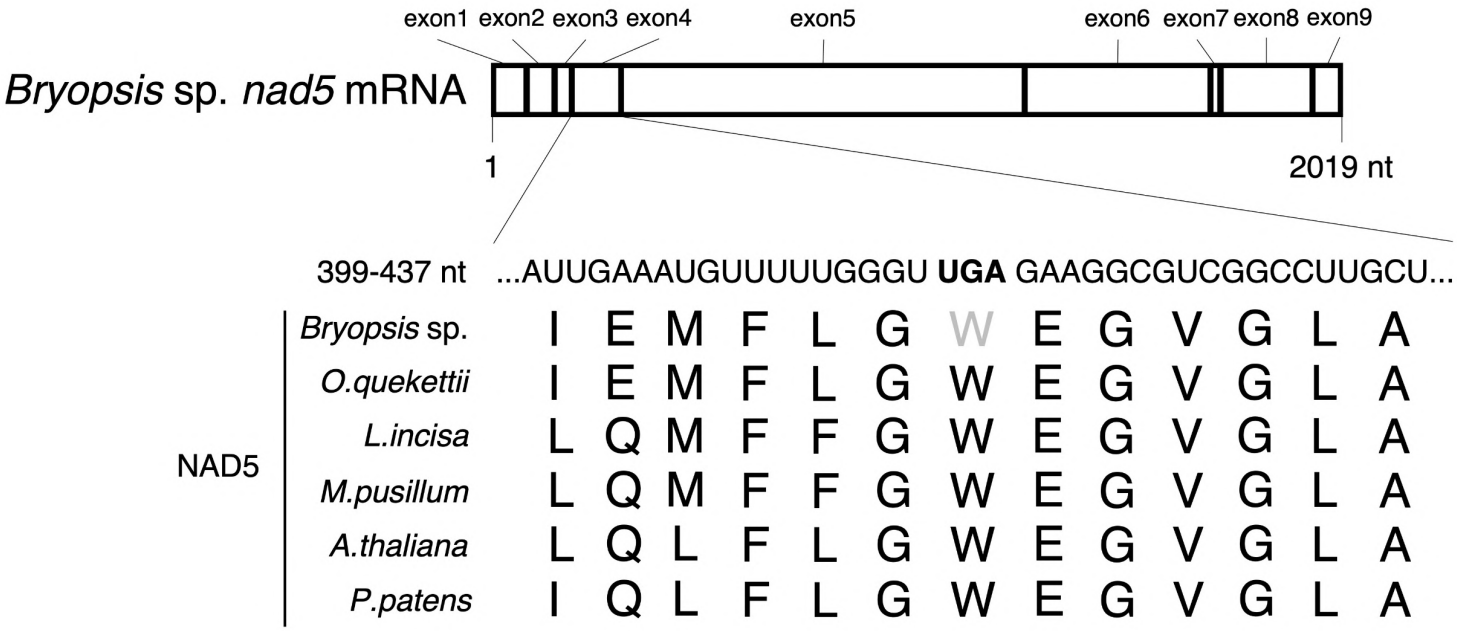


Figure S9. UGA codon likely encodes tryptophan in the mitochondrial genome

Based on the amino acid sequences of the Nad5 protein (this figure) and other conserved proteins in green algae, the UGA of *Bryopsis* sp. likely represents a tryptophan codon, not a termination codon, in the mitochondrial genome.