

RESEARCH ARTICLE

Screening and verification of extranuclear genetic markers in green tide algae from the Yellow Sea

Chuner Cai^{1,2*}, Kai Gu¹, Hui Zhao¹, Sophie Steinhagen³, Peimin He¹, Thomas Wichard^{1,2*}

1 College of Marine Ecology and Environment, Shanghai Ocean University, Shanghai, China, **2** Institute for Inorganic and Analytical Chemistry, Jena School for Microbial Communication, Friedrich Schiller University Jena, Jena, Germany, **3** Department of Marine Sciences-Tjärnö Marine Laboratory, University of Gothenburg, Strömstad, Sweden

* cecai@shou.edu.cn (CC); thomas.wichard@uni-jena.de (TW)



OPEN ACCESS

Citation: Cai C, Gu K, Zhao H, Steinhagen S, He P, Wichard T (2021) Screening and verification of extranuclear genetic markers in green tide algae from the Yellow Sea. PLoS ONE 16(6): e0250968. <https://doi.org/10.1371/journal.pone.0250968>

Editor: Berthold Heinze, Austrian Federal Research Centre for Forests BFW, AUSTRIA

Received: August 4, 2020

Accepted: April 18, 2021

Published: June 1, 2021

Copyright: © 2021 Cai et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The mainly relevant data are contained within the paper and its [Supporting information](#) files. Others are held in NCBI database with accession number KX595275.

Funding: The project was funded by the Natural Science Foundation of Shanghai (18ZR1417400) (CC), the Shanghai Port and Offshore Ecological Environment Technology Service Platform Project (19DZ2292500) (CC), the Key Laboratory of Integrated Marine Monitoring and Applied Technologies for Harmful Algal Blooms, S.O.A. (MATHAB201817) (RJ) and by the Deutsche

Abstract

Over the past decade, *Ulva compressa*, a cosmopolitan green algal species, has been identified as a component of green tides in the Yellow Sea, China. In the present study, we sequenced and annotated the complete chloroplast genome of *U. compressa* (alpha-numeric code: RD9023) and focused on the assessment of genome length, homology, gene order and direction, intron size, selection strength, and substitution rate. We compared the chloroplast genome with the mitogenome. The generated phylogenetic tree was analyzed based on single and aligned genes in the chloroplast genome of *Ulva* compared to mitogenome genes to detect evolutionary trends. *U. compressa* and *U. mutabilis* chloroplast genomes had similar gene queues, with individual genes exhibiting high homology levels. Chloroplast genomes were clustered together in the entire phylogenetic tree and shared several forward/palindromic/tandem repetitions, similar to those in *U. prolifera* and *U. linza*. However, *U. fasciata* and *U. ohnoi* were more divergent, especially in sharing complementary/palindromic repetitions. In addition, phylogenetic analyses of the aligned genes from their chloroplast genomes and mitogenomes confirmed the evolutionary trends of the extranuclear genomes. From phylogenetic analysis, we identified the *petA* chloroplast genes as potential genetic markers that are similar to the *tufA* marker. Complementary/forward/palindromic interval repetitions were more abundant in chloroplast genomes than in mitogenomes. Interestingly, a few tandem repetitions were significant for some *Ulva* subspecies and relatively more evident in mitochondria than in chloroplasts. Finally, the tandem repetition [GAAATATATAATAATA × 3, abbreviated as TRg] was identified in the mitogenome of *U. compressa* and the conspecific strain *U. mutabilis* but not in other algal species of the Yellow Sea. Owing to the high morphological plasticity of *U. compressa*, the findings of this study have implications for the rapid non-sequencing detection of this species during the occurrence of green tides in the region.

Forschungsgemeinschaft (DFG, German Research Foundation) - SFB 1127/2 ChemBioSys - 239748522 of the Friedrich Schiller University Jena (TW). The funders had no involvement in the study design, data collection/analysis, or manuscript preparation or publication.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Green algae extranuclear genomes date back at least one billion years [1] and account for 0.1% of the nuclear genome. They are used to forecast the evolutionary trends of various species [2, 3]. A phylogenetic analysis of chloroplasts from green algae, for example, discovered a new lineage sister to the Sphaeropleales [4]. Trebouxiophyceae species were multiphyletic in the subphylum, and three clusters were included in the core trebouxiophyceans [5, 6]. Furthermore, the Chlorophyceae contain two robustly supported major lineages, including a clade that unites the Oedogoniales, Chaetophorales, and Chaetopeltidales, as well as a clade that unites the Sphaeropleales and Chlamydomonadales, where *Chlamydomonas reinhardtii* contains putative tyrosine kinases that function in phosphotyrosine signal, which most probably appeared before the animal and fungal lineages were diverged [7, 8].

Pedinophyceae and Trebouxiophyceae, on the other hand, are classified based on the sequences of their nuclear and plastid-encoded ribosomal RNA (rRNA) operons [9]. Ulvophyceae evolution was studied using phylogenetic analysis of seven nuclear genes, small subunit nuclear rDNA, and the plastid genes *rbcL* and *atpB* [10], whereas Ulvales evolution was studied using *rbcL* and small subunit nuclear rDNA [11]. Results from these studies revealed a few unusual structural features based on the chloroplast genomes of individual species. For example, although *Pseudendoclonium* cpDNA features a large inverted repeat, its quadripartite structure is unusual in that it displays an rRNA operon transcribed toward the large single-copy region and a small single-copy region containing 14 genes that are normally found in the large single-copy region [12]. Moreover, *Ulva* spp. share different numbers of introns [13]. Only a few studies have investigated chloroplast genome-based phylogeny within the genus *Ulva*, and no comparisons have been made with their corresponding mitogenomes [13, 14].

Several representatives of the genus *Ulva* exhibit strong morphological plasticity, and respective morphotypes can vary depending on changing environmental factors [15]. This often resulted in errors in the classical morphological criteria for identification [16–18], especially within subspecies. For instance, the species dominating the green tides in the Yellow Sea was first identified as *U. prolifera* in 2008 [19–22], and was subsequently amended as a new subspecies, namely, *U. prolifera* subsp. *qingdaoensis* [18]. Although internal transcribed spacers (ITS) and *rbcL* gene sequences are widely used in the classification of green tide algae, they cannot be used to distinguish the subspecies from the species complex “*Ulva linza-procera-prolifera*” [23–27]. However, the combination of ITS and 5S rDNA could be applied to characterize *U. prolifera* from *U. linza* [24, 26–35]. Moreover, in some cases, the *tufA* marker gene possessed, e.g. between *U. prolifera* and *U. linza*, a higher sensitivity than the DNA barcodes ITS and *rbcL* [18, 36, 37]. Nevertheless, there was a lack of systematic screening for suitable barcodes in characterized regions of whole mitochondrial and chloroplast genomes. Genetic markers were developed within the whole nuclear genomes to identify green-tide forming algae, including inter-simple sequence repeats [19], PCR-restriction fragment length polymorphisms [29], and *in situ* hybridization techniques [38]. Although these methods were accurate, they were time-consuming and entailed high costs.

In the present study, we aimed to (i) identify, sequence, and annotate the complete chloroplast genome of *U. compressa*, (ii) perform comparative phylogenetic analyses based on single and aligned genes in the chloroplast genome within the genus *Ulva* for comparison with those from mitogenomes, to determine evolutionary trends, and finally (iii) systematically screen for characteristic regions among the whole mitogenomes and chloroplast genomes of the genus *Ulva* for DNA barcoding purposes. To this end, we sequenced and annotated the complete chloroplast genome of *U. compressa* (alpha-numeric code: RD9023), a constitutive species of the green tide in the Yellow Sea, China. We hypothesized that specific nucleotide sequences

across the whole mitochondrial and chloroplast genomes can distinguish *U. compressa* from other species, considering a rapid non-sequencing approach, and subsequently be used for the identification of evolutionary trends. Phylogenetic analyses were performed and a comparative analysis was carried out for available chloroplast and mitochondrial genomes. Finally, systematic screening was performed for unique barcodes within characterized regions of mitogenomes and chloroplast genomes using algal strains collected from China (the Yellow Sea) and Europe (the North Sea and Baltic Sea).

Materials and methods

Sampling and species identification

The specimens of *U. compressa* used in this study to unravel the chloroplast genome were collected from the southern Yellow Sea near the Rudong Sea area in Jiangsu Province [39]. Healthy individual algae were selected and washed several times with sterile seawater autoclaved at 121 °C for 15 min. Following the removal of any surface attachments, they were cultivated in VSE medium [40] at 24 °C under a 12 h photoperiod and photon flux of 130–160 $\mu\text{mol}\cdot\text{m}^{-2}\text{s}^{-1}$ before being further characterised via ITS [41] and *tufA* molecular barcoding. Additionally, *Ulva* samples collected from either China (Yellow Sea) or Europe (North and Baltic Sea) were assessed for further comparison (e.g., using phylogenetic trees and barcodes) [16]. No endangered or protected species were targeted at the sampling sites mentioned, and no specific permissions were required for these locations and activities.

Sequencing and annotation

Several published and peer-reviewed chloroplast genomes (NC_035823 [13], NC_029040 [42], KP720616 [43], and NC_030312 [44]) from representatives of the genus *Ulva* were used as a reference for primer design to sequence the chloroplast genome of *U. compressa*. After the pre-test, selected primers were divided into 15 pairs (S1 Table in S1 File) for PCR amplification under the following conditions: pre-denaturation at 94 °C for 1 min, 35 cycles of amplification each at 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 8 min, followed by a final extension at 68 °C for 15 min. The PCR mixture contained 0.5 μL of DNA polymerase ($5\text{ U}\cdot\mu\text{L}^{-1}$), 5 μL of buffer (Mg²⁺ Plus; 10 \times), 8 μL of dNTP mixture (2.5 mM each), 0.2–1.0 μM (final concentration) of forward and reverse primers, and purified chloroplast DNA (<1 μg). RNase-free water was added to increase the final reaction volume to 50 μL . Takara *LA Taq* polymerase (Takara Biomedical Technology Co., Ltd, Beijing, China) was used for fragments longer than 4,000 bp, whereas Takara *Ex Taq* polymerase (Takara Biomedical Technology Co., Ltd, Beijing, China) used for shorter fragments. PCR products of long segments were fragmented using ultrasound, connected to the PMD19-T vector, and transformed into *E. coli* DH5 α (Sangon Biotech Co., Ltd., Shanghai, China). Positive clones were selected for shotgun sequencing using the ABI3730 sequencer (Applied Biosystems Co., US). Raw sequence reads were trimmed using `cross_match v1.09` and assembled using `phred v1.09` [45].

The *U. compressa* chloroplast genome was annotated with CpGAVAS in which the cutoff for the E-values of BLASTN and BLASTX was $1\cdot e^{-10}$ [46]. Meanwhile, transfer RNA (*tRNA*) genes were identified using *tRNAscan-SE* [42] and ARAGORN [47]. Inverted repeat (IR) sequences were predicted using Gepard [48], and the circular chloroplast genome map of *U. compressa* was drawn using OrganellarGenomeDRAW [49]. Furthermore, GC content was determined using the `Compseq` program provided by EMBOSS [50]. Final genome assembly and annotation results were deposited in GenBank (accession number: KX595275).

Simple sequence repeats (SSRs) were detected using the MISA Perl script available at <http://pgrc.ipk-gatersleben.de/misa/>, at the following thresholds: eight repeat units for

mononucleotide SSRs, four repeat units for di- and tri-nucleotide repeat SSRs, and three repeat units for tetra-, penta-, and hexa-nucleotide repeat SSRs. Tandem repeats were analyzed using the Tandem Repeats Finder program [51] by setting two parameters for matches and seven for indels and mismatches, in which 500 and 50 were set as maximum period size and minimum alignment score. After manual verification, irrelevant results were removed. The REPuter program [52] was employed to identify different IR sequences (forwarding, reverse, complementary, and palindromic) in *U. compressa*. We used 30 bp as the minimal repeat size, and the homology between repeat units was above 90%.

Comparative analysis of chloroplast genomes

Reference chloroplast genomes of *U. mutabilis* (txid498180), *U. flexuosa* (NC_035823), *U. linza* (NC_030312), *U. prolifera* (NC_036137), *U. ohnoi* (AP018696), *U. fasciata* (NC_029040), and *Ulva* sp. UNA00071828 (KP720616) were subjoined into a comparative genome analysis of *U. compressa* that assessed genome length and homology, gene order and direction, intron size, selection force, and substitution rate. The selection force and substitution rate assessment of 71 protein-coding genes (PCGs) extracted from *Ulva* species using the KaKs_Calculator Toolbox 2.0 [53] were performed using γ -NG methods and the standard genetic code [54]. *Ulva compressa* was considered as control during this assessment. In particular, *U. compressa* was compared with the conspecific *U. mutabilis* Føyn that was developed into a model organism [55, 56].

Phylogenetic tree analysis

Species with different gene combinations were subjected to phylogenetic analyses using either a group of or single genes. Firstly, the evolutionary status of seven species within genus *Ulva* (*U. compressa*, *U. mutabilis*, *U. flexuosa*, *U. linza*, *U. prolifera*, *U. ohnoi*, and *U. fasciata*) was deciphered from the aligned genes of the chloroplast genomes, including 71 PCGs. To compare the results of this analysis, the mitochondrial genome-related phylogenetic analysis was subsequently performed using similar types of mitogenomes from the NCBI database. For each species, 29 aligned PCGs were used. *Pseudendoclonium akinetum* was used as an out-group in each of these phylogenetic analyses.

Additionally, a phylogenetic tree showing PCG, rRNA (*rrn5*, *rrn16*, and *rrn23*), and the alignment of all tRNAs from the seven chloroplast genomes mentioned above was constructed separately.

Finally, the individual genes whose phylogenetic trees were in line with that of the whole chloroplast genome were selected for *in silico* PCR using designed primers. After comparison, two coding genes, namely *petA* and *tufA*, whose phylogenetic trees were in line with that of the whole chloroplast genome, were selected to classify the samples obtained from the Yellow Sea in China (S2 Table in S1 File).

Alignments were performed using ClustalW [57], manually examined, and adjusted. Each of the phylogenetic trees was analyzed based on maximum likelihood method using the best model tested in MEGA X [58]. The strength of support of branches in the phylogenetic tree was assessed by bootstrap testing with 1,000 replications.

Non-coding sequence analysis

Complement, forward, and palindromic matches from either chloroplast genomes (over 20 bp) or mitogenomes (over 30 bp) were compared separately to identify the consensus sequences among seven species in the genus *Ulva* (*U. compressa*, *U. mutabilis*, *U. flexuosa*, *U.*

linza, *U. prolifera*, *U. ohnoi*, and *U. fasciata*). Furthermore, unique tandem repeats in these genomes were revealed.

Twelve pairs of primers were designed across both sides of selected tandem repeats in the chloroplast (three pairs) and mitochondrial (nine pairs) DNA from *U. compressa*. These were used to amplify DNA from four constitutive green tide species in the Yellow Sea (S3 Table in [S1 File](#)). The PCR products were tested by agarose gel electrophoresis and sequenced to identify unique bands.

Two primer pairs (m-*cox3-trnR* and m-*trnG-trnY*) were further used to test 22 strains of green tide samples from the Yellow Sea (S4 Table in [S1 File](#)). The subsequent analysis focused on the exclusive pair of primers (m-*trnG-trnY*) that was tested on *U. mutabilis* (wild-type and “slender”—a naturally occurring developmental mutant), conspecific with *U. compressa* as well as on 27 *Ulva* samples from the North and Baltic Seas (Europe) (S5 Table in [S1 File](#)) and *U. ohnoi* (origin: Mediterranean Sea).

Results

Sequencing and annotation of the chloroplast genome

Sequencing analysis revealed the complete chloroplast genome of *U. compressa* (accession number: KX595275) with 96,808 bp and no IR region. The chloroplast genome clustered with peer-reviewed sequences (downloaded from GenBank) previously identified as *U. compressa* (S1 Fig in [S1 File](#)). One hundred genes were identified, including 71 PCGs, 26 *tRNA* genes, and three *rRNA* genes (S6 Table in [S1 File](#)). The general structure and location of these genes was typical for a chloroplast genome in *Ulva* ([Fig 1](#)). Six introns were predicted, one each in *petD*, *petB*, *atpB*, and *atpA*, and two in *rrn23*. Overall, 74.0% of the genome was composed of genes that encode proteins. The overall GC content of the genome was 26.2%, and 26.5% was allocated to protein-coding regions. Within the latter regions, the GC content for the first, second, and third positions of codons was 34.4%, 31.2%, and 13.8%, respectively.

A total of 235 SSRs were identified in the *U. compressa* chloroplast genome. Among these, the majority consisted of mono- and di-nucleotide repeats, which were detected 148 and 42 times, respectively. All of the mononucleotide repeats were A/T repeats. Similarly, the majority of dinucleotide repeats (92.9 percent) were AT/AT repeats. Tri-(19), tetra-(13), penta-(3), and hexa-(8) nucleotide repeats were detected at a lower frequency, with 24 localized in the intergenic regions, 17 in the coding regions, and others stretched across both the intergenic and coding regions (S7 Table in [S1 File](#)).

Eighteen tandem repeats longer than 20 bp were identified with a similarity > 90%. Thirteen tandem repeats were located in the intergenic regions, and the remaining within coding regions. Additionally, 10 forward repeats were identified with a size cutoff of 30 bp (S8 Table in [S1 File](#)). Among these, the most extended units were 63 bp and were located in the coding regions of *psaB* and *psaA*.

Comparative analysis of chloroplast genomes

The length of the chloroplast genome from the genus *Ulva* ranged from 86,726 bp (*U. linza*) to 119,855 bp (*U. mutabilis*) and the number of bases in non-coding regions varied largely compared to that of PCGs, *rRNAs*, and *tRNAs*, which accounted for 60%–80% of the whole genome ([Fig 2](#)).

The direction of gene transcription was slightly inconsistent in the case of 7 PCGs and 14 *tRNAs* from the selected 7 species. *Ulva linza* and *U. prolifera* had many in common, as did *U. ohnoi* and *U. fasciata*. However, this was not the case between *U. compressa* and *U. mutabilis* (S9 Table in [S1 File](#)).

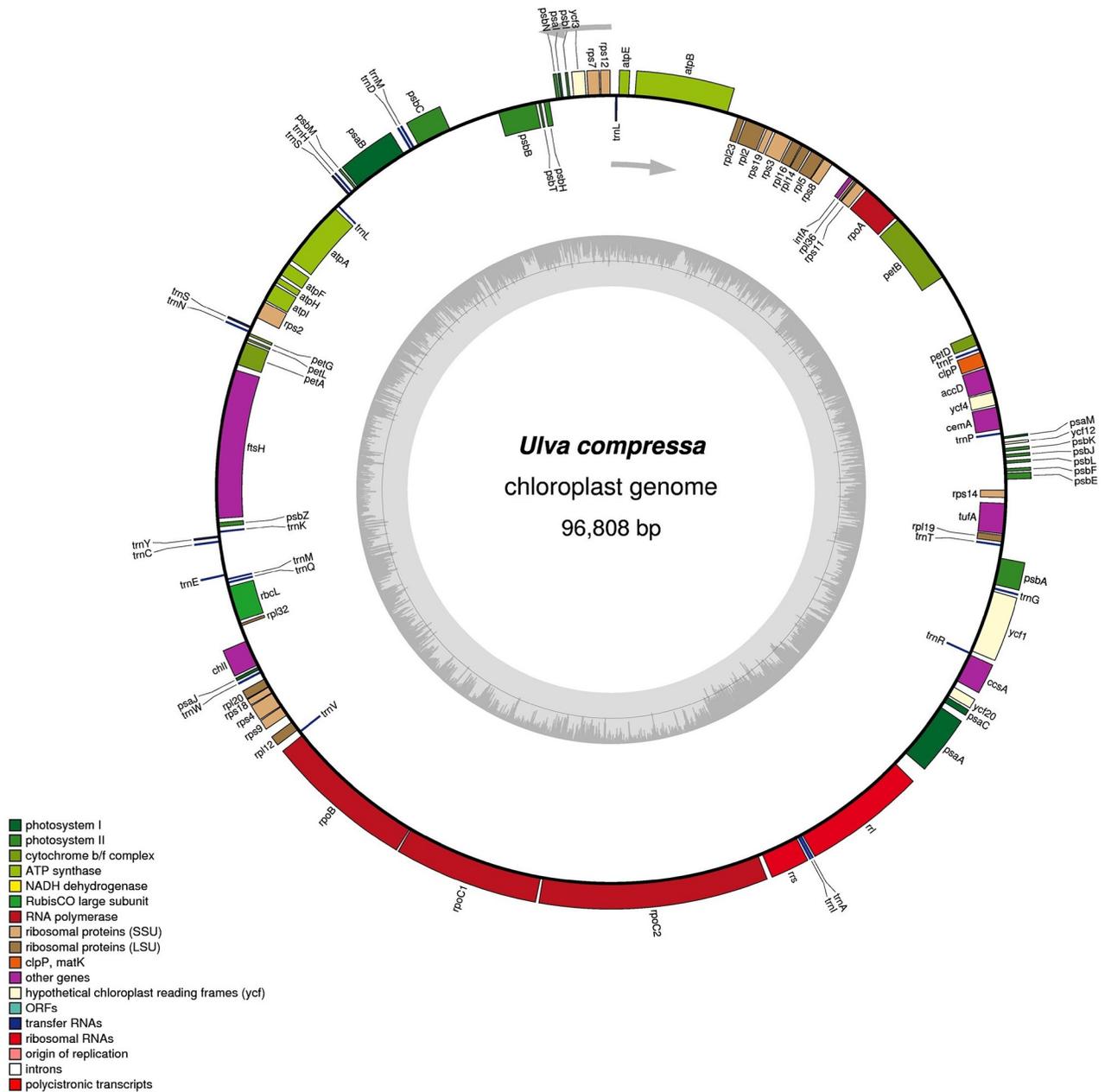


Fig 1. Schematic representation of the *U. compressa* chloroplast genome using OGDRAW. The predicted genes are shown, with the colors representing functional classifications at the bottom left. The genes outside the circle were transcribed counterclockwise. The inner-circle shows the GC content.

<https://doi.org/10.1371/journal.pone.0250968.g001>

All the assessed species, except for *U. linza*, contained introns that were spread irregularly within the genus, and a significantly higher number of intron regions were found in *U. mutabilis* compared to other *Ulva* species. Notably, 11 introns were included in *U. mutabilis*, whereas *U. compressa* contained only six. A loss of intronic genes may explain the difference in the number of introns among the species. However, these conclusions were derived from bioinformatic analyses that require further testing to be confirmed.

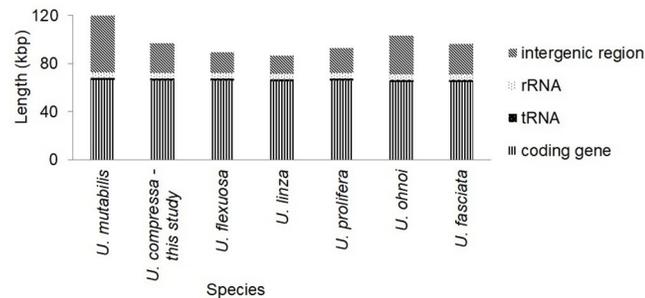


Fig 2. Length of the chloroplast genomes in genus *Ulva*. The total length of intergenic regions varied widely compared to that of the corresponding PCGs, rRNAs, and tRNAs.

<https://doi.org/10.1371/journal.pone.0250968.g002>

To understand the evolutionary trends of the investigated species, the Ka/Ks values of six species were compared against *U. compressa* (control) using 71 PCGs (Fig 3). The Ka/Ks ratio estimates the balance between neutral mutations, purifying selection, and beneficial mutations based on a set of homologous PCGs. Except for *ycf12*, *rps3*, *rps19*, *rpl23*, *atpF*, *petA*, *rpl32*, *chlI*, *rpl20*, and *rpl12* in *U. mutabilis*, the Ka value was always lower than the Ks value (i.e., $Ka/Ks \leq 1$). Moreover, both Ka and Ks values were distinctly lower (approximately 1/10) in *U. mutabilis* compared to the other analyzed species (Fig 3). The average selective constraint in *U. mutabilis* (0.38) and other *Ulva* species (0.17–0.20) is purifying selection, according to the Ka/Ks ratio. Furthermore, the jagged curve raised several hypotheses for tracing evolutionary changes in these genomes, with *U. linza* and *U. prolifera*, as well as *U. ohnoi* and *U. fasciata*, acting nearly identically according to the curve and gene queues.

Phylogenetic tree based on extranuclear genomes

We used aligned PCGs from chloroplast genomes to divide the seven *Ulva* species into four groups for the phylogenetic analysis (Fig 4A). The pairs of *Ulva linza* and *Ulva prolifera*, *Ulva ohnoi* and *Ulva fasciata* as well as *Ulva mutabilis* and *Ulva compressa* belonged to the same group, while *U. flexuosa* was more closely related to *Ulva linza* and *Ulva prolifera*.

The mitogenome-related phylogenetic analysis of 29 aligned PCGs elucidated the evolutionary status, which was consistent with that of the chloroplast genomes of the seven analyzed species (Fig 4B). The cluster was encountered within all the samples from the same species.

Identification of *tufA* and *petA* markers

A specific type of phylogenetic tree was discovered in a phylogenetic analysis of all PCGs from chloroplast genomes (S2 Fig in S1 File). In this type, each calculation that considered any of the PCGs, including *rbcL*, *rpoA*, *rpoB*, *rpoC1*, *ftsH*, *rpoC2*, *psaA*, *psbI*, *ycf12*, *petB*, *rps11*, *rps7*, *ycf4*, *atpB*, *cemA*, *chlI*, *clpP*, *petA*, *rpl5*, *tufA* and *ycf20*, resulted in a similar phylogenetic tree that was also consistent with the tree obtained from whole chloroplast genomes (Fig 4A). But phylogenetic tree from rRNAs (S3 Fig in S1 File) and tRNAs (S4 Fig in S1 File) were not the same as that from the whole chloroplast genomes.

We aimed to characterize genes from the above special coding gene group, which could potentially function as a suitable barcode for species identification, similar to *tufA* (1.2k bp). A suitably-sized barcoding marker (e.g., $1.2k \pm 0.3k$ bp) should contain the conserved regions and avoid most non-conserved ones. However, most of the elucidated genes were unsuitable. Some genes were too long (e.g., *rpoB*, *rpoC1*, *rpoC2*) or too short (e.g., *chlI*, *psaA*, *clpP*, *ftsH*, *psbI*, *rpl5*, *rps11*, *rps7*, *ycf12*, *ycf20*, *ycf4*) to work as a barcode. Others contained non-conserved

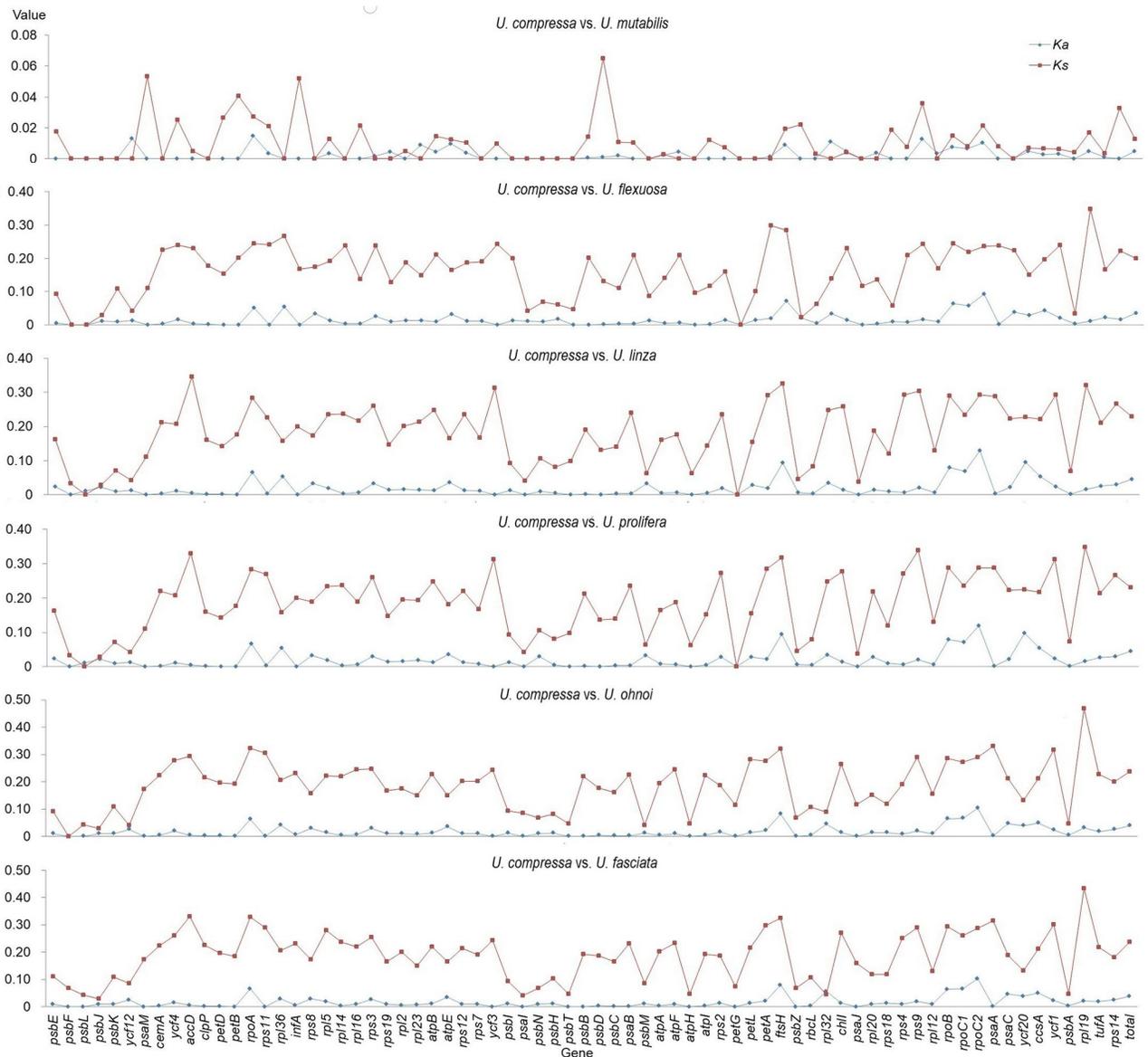


Fig 3. Ka/Ks value assessment of 71 PCGs in the chloroplast genomes of seven species. The value was measured using the $KaKs_Calculator$ Toolbox 2.0, the γ -NG method, and the standard genetic code. *Ulva compressa* (NCBI #KX595275) was used as a control.

<https://doi.org/10.1371/journal.pone.0250968.g003>

regions, e.g., introns, in their genome sequences (e.g., *petB*, and *atpB*). A few genes caused problems in the design of suitable primers, and the PCR simulation failed (e.g., *cemA* and *rpoA*). Finally, *petA* sequences allowed us to identify green tide samples from the Yellow Sea, as did the control analysis using *tufA* (Fig 5).

Duplication of non-coding sequences in extranuclear genomes

In addition to assessing the PCGs, they were screened for other regions focusing on non-coding sequences among the whole mitogenomes and chloroplast genomes for DNA barcoding purposes. IR sequences such as complementary, forward, or palindromic repeats and tandem repeat sequences were relatively conserved among the non-coding regions. The IR sequences

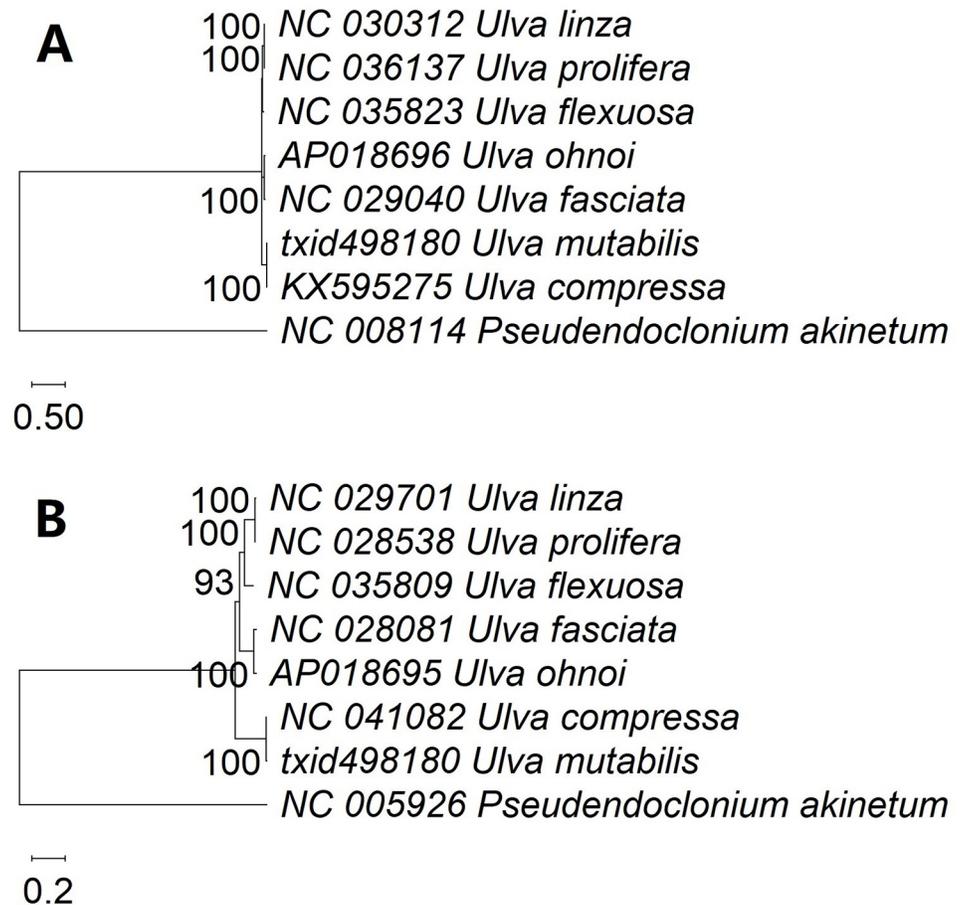


Fig 4. Molecular phylogenetic analysis of *Ulva* species using chloroplast or mitochondrial genomes and the maximum likelihood method based on the General Time Reversible+G+I model (A) and General Time Reversible+G model (B). (A) Molecular phylogenetic analysis with whole aligned coding genes of the *Ulva* chloroplast genomes. The coding genes include *accD*, *atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*, *ccsA*, *cemA*, *chlI*, *clpP*, *ftsH*, *infA*, *petA*, *petB*, *petD*, *petG*, *petL*, *psaA*, *psaB*, *psaC*, *psaI*, *psaJ*, *psaM*, *psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*, *rbcL*, *rpl12*, *rpl14*, *rpl16*, *rpl19*, *rpl2*, *rpl20*, *rpl23*, *rpl32*, *rpl36*, *rpl5*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps11*, *rps12*, *rps14*, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps9*, *tufA*, *ycf1*, *ycf12*, *ycf20*, *ycf3*, and *ycf4*. (B) Molecular phylogenetic analysis with whole aligned coding genes of the *Ulva* mitochondrial genomes. The coding genes include *atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *cob*, *cox1*, *cox2*, *cox3*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *rpl14*, *rpl16*, *rpl5*, *rps10*, *rps11*, *rps12*, *rps13*, *rps14*, *rps19*, *rps2*, *rps3*, and *rps4*.

<https://doi.org/10.1371/journal.pone.0250968.g004>

were not as species-specific as the tandem repeats, either in mitochondrial or chloroplast genomes. In some species of *Ulva*, shared IRs were discovered in the chloroplast genomes (S10 Table in S1 File, #1–13), including two complement (#1–2), three forward (#3–5), and eight palindromic matches (#6–13). Six IRs were shared by *U. linza* and *U. prolifera*, and three by *U. compressa* and *U. mutabilis*. However, *U. linza* and *U. prolifera* had only one IR (84 bp) in the mitogenome in common (S10 Table in S1 File, #14).

The tandem repeats appeared to differ among the species significantly but were identical within the subspecies. Forty repeats were found in mitogenomes, with a longer average length (16 bp) than that in chloroplast genomes (11 bp), which only contained six. In mitogenomes, *U. compressa* and *U. mutabilis* shared nine tandem repeats and four in chloroplast genomes, respectively, whereas *U. linza* and *U. prolifera* shared eight and two tandem repeats, respectively (S11 Table in S1 File). The longest tandem repeat (35 bp) was found in *Ulva fasciata*

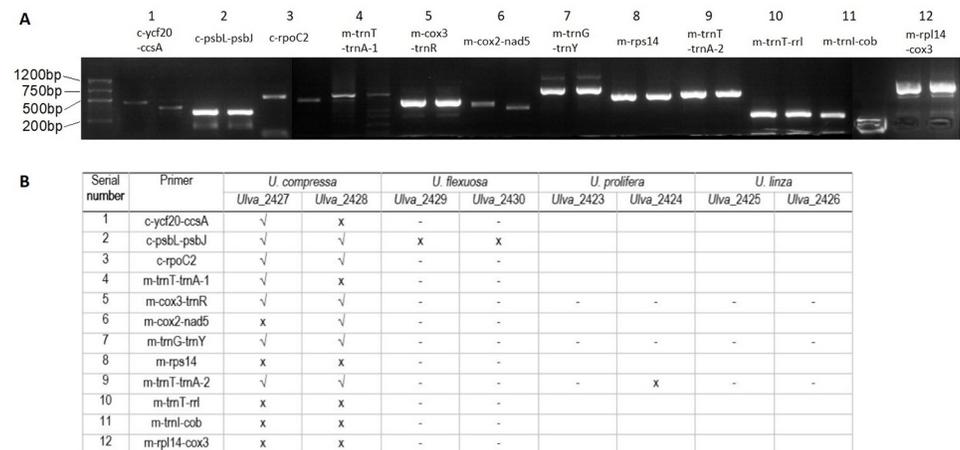


Fig 6. Testing 12 primer pairs on four constitutive green tide species in the Yellow Sea. (A) Electrophoresis results following PCR using *U. compressa* DNA and the 12 primer pairs. (B) Electrophoresis results following PCR using DNA from four constitutive green tide species in the Yellow Sea and the twelve primer pairs. Either “√” or “x” mean there are bands in the electrophoresis with related primers. Wherein “√” means the consistency of the band sequence with predicted tandem repeats, while “x” means not. “-” means no band; blank means not tested. Each primer pair was tested on two strains of the species in Panel A. Information about 12 pairs of primers was detailed in S3 Table in [S1 File](#).

<https://doi.org/10.1371/journal.pone.0250968.g006>

(*c-psbL-psbJ*) successfully amplified the *U. flexuosa* genome (Fig 6B), but the resulting sequence was not as predicted.

Following that, three primer pairs (*m-cox3-trnR*, *m-trnG-trnY*, and *m-trnT-trnA-2*) were tested on *U. linza* and *U. prolifera* to see if they generated PCR products from *U. compressa*, as evidenced by a distinct band on an agarose gel. The *m-trnT-trnA-2* primer only worked for *U. prolifera*, did not reveal the predicted sequence (Fig 6B). When tested against all the strains collected from the Yellow Sea’s green tide ($n = 22$), the *m-trnG-trnY* primer amplified the predicted ‘TRg’ sequence only in *U. compressa* and discriminated against *U. flexuosa*, *U. prolifera*, and *Blidingia minima* (S4 Table in S1 File and Fig 7A). As proof of principle, the TRg PCR product derived from *U. mutabilis* was conspecific to *U. compressa* and identical in size (800 bp), as predicted from the *U. mutabilis* genome [56] (Fig 7B).

Finally, the *m-cox3-trnR* primer produced non-specific bands in all samples when using green tide strains collected from the Yellow Sea (S4 Table in S1 File), which included *U. prolifera*, *U. compressa*, *U. flexuosa*, and *Blidingia minima*.

Samples from other regions reacted differently toward the specificity of the primers used. In addition to *U. compressa* and *U. mutabilis*, *U. ohnoi* (from the Mediterranean Sea) contained the same tandem repeat, as confirmed by sequencing analysis (S5 Fig in S1 File and Fig 7B). However, using the *m-trnG-trnY* primer, amplification of samples collected in the North Sea and Baltic Sea (Europe) resulted in various PCR products of different lengths independent of the *Ulva* species (Fig 8). As a result, the species specificity of this primer could only be attributed to Yellow Sea strains.

Discussion

Comparative analysis of chloroplast genomes

The conspecificity of *U. compressa* and *U. mutabilis*, which has been studied separately for years [55, 59–61], was recently characterized using morphological patterns, hybridization, and

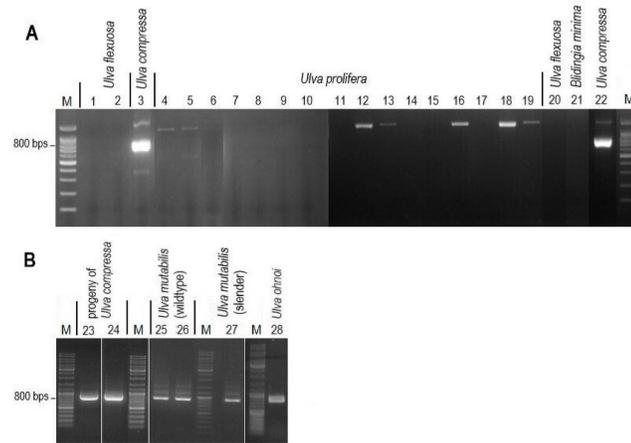


Fig 7. Application of the m-trnG-trnY primer on *U. compressa* and *U. mutabilis*. (A) Lanes 1–22 correspond to amplification using m-trnG-trnY with 22 green tide samples from the Yellow Sea, China, which were detailed in S4 Table in S1 File. Lanes 3 and 22 show the two brightest bands, characterized by the sequence GAAAATATAAATA (× 3), with a length of approximately 800 bp. In comparison with the DNA marker, the bands in lanes 4, 5, 12, 13, 16, 18 and 19 were higher than 800 bp. Subsequent sequencing using the m-trnG-trnY primer resulted in no result, suggesting the occurrence of non-specific amplification. (B) Lanes 23–28 correspond to PCR amplification using the m-trnG-trnY primer with *U. compressa* progeny, the candidate model *U. mutabilis*, and *U. ohnoi*. Two different types of standard genetic markers were used in (A) and (B) because the two experiments were performed independently.

<https://doi.org/10.1371/journal.pone.0250968.g007>

molecular phylogenetic methods [17]. Thus, it was not surprising that the chloroplast genomes of *U. mutabilis* and *U. compressa* were consistent in their structure in terms of gene sequence and homology. Both species clustered together in the phylogenetic tree (Fig 4) and shared several forward/palindromic/tandem repeats, including TRg, detected in both the wild-type and slender-morphotype of *U. mutabilis*. However, *U. compressa* and *U. mutabilis* (wild type) sequences differed in their total number of introns, length of intergenic regions (Fig 2), and transcription directions (S9 Table in S1 File), all of which were lesser conserved elements. The difference in the length of intergenic spacers may be caused by large-fragment insertions and deletions [62].

Ulva prolifera, *U. linza*, *U. compressa*, and *U. flexuosa* are the four main green-tide-forming *Ulva* species from the Yellow Sea that exhibit varying morphologies [35, 39]. *Ulva prolifera*

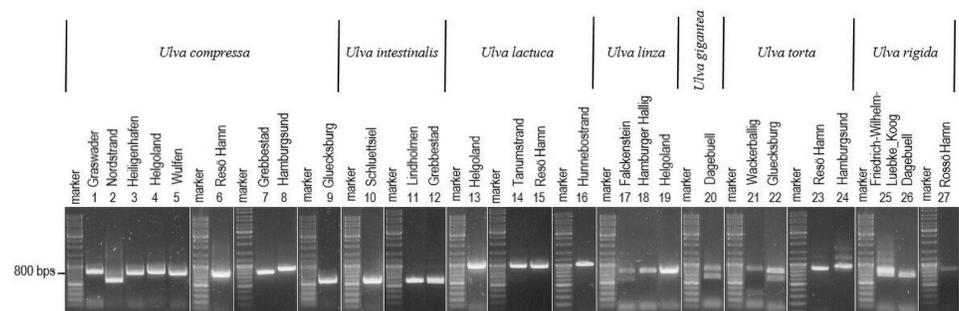


Fig 8. Application of primer m-trnG-trnY on 27 *Ulva* samples collected in the North and Baltic Seas (Europe). The standard genetic marker is shown corresponding to the individual gel electrophoresis analysis. The name and number of the sampling site for each *Ulva* isolate are indicated above each lane. Details of the 27 samples are given in S5 Table in S1 File.

<https://doi.org/10.1371/journal.pone.0250968.g008>

and *U. linza* clustered together following whole chloroplast- and mitogenome analyses (compare with Fig 4). Analysis of the chloroplast genomes revealed that the Ka and Ks graphs generated from the selection force and substitution rate evaluation charts were almost identical (Fig 3). Furthermore, the sequence of complementary/forward/palindromic/tandem repeats was almost consistent (S10 and S11 Tables in S1 File). This observation could explain the phenomenon that *U. prolifera* was derived from the hybridization of *U. linza* and *U. procera* in the Yellow Sea [18].

Ulva fasciata, a heterotypic synonym of *U. lactuca* [17], shares a few genotyping markers with *U. ohnoi*, and is also closely related to *U. ohnoi* based on chloroplast genome analysis [63]. However, *tufA* and *rbcl* analyses indicate that *U. fasciata* and *U. ohnoi* could be resolved into separate lineages. If more species were included in the analysis, they could be even separated by *U. spinulosa* and *U. reticulata* [64, 65]. *Ulva fasciata* and *U. ohnoi* shared only a few complementary/palindromic repeats and contained no forward or tandem repeats in their chloroplast genomes and mitogenomes. *U. mutabilis* and *U. compressa*, as well as *U. prolifera* and *U. linza* contained several such repeats (S10 and S11 Tables in S1 File). Therefore, we suggest that the repeat sequences within intergenic regions in the extranuclear genome contain unique signals that characterize the species clustered together in the phylogenetic tree (Fig 4, S10 and S11 Tables in S1 File). The phylogenetic trees derived from chloroplast genomes and mitogenomes were highly similar.

The Ka/Ks ratios of most of the single coding genes indicated a purifying selection [14], which shows that the corresponding PCS continue to be consistent. Furthermore, the jagged curve along the gene queue in these circular DNA sequences was in line with the original transcription point, which was randomly distributed within the entire chloroplast genome, indicating a mutation in uniform distribution [66].

Exploration of the *tufA* and *petA* markers

The mitogenomes and chloroplast genome contain potential barcodes [67–72]. *tufA* is a genetic marker [16] with higher sensitivity and applicability than ITS, *rbcl*, universal amplicon, and large ribosomal subunit [18, 36, 37]. Moreover, *tufA* can be used to characterize the evolutionary status of the extranuclear genome that has been confirmed in the present study by analyzing the aligned genes from chloroplast genomes and mitogenomes. *petA* is one of the large subunits of the cytochrome b6-f complex that covalently binds one heme group and mediates electron transfer between photosystems II and I [73, 74]. In the present study, *petA* possessed similar conserved molecular functions as *tufA* owing to its key roles in cells and could, therefore, be used to verify the results obtained from *tufA*-based phylogenetic analyses.

Duplication of the non-coding sequence in extranuclear genomes

According to the endosymbiosis hypothesis, mitochondria originate from the aerobic bacteria [75], while chloroplasts come from endosymbiotic and photoautotrophic cyanobacteria [76–78]. However, the non-endosymbiotic hypothesis states that the mitochondria evolved from the membrane system such as the endoplasmic reticulum [79]. Both mitochondria and chloroplasts are organelles that contain DNA which is different from that of the nucleus but similar to that of the bacteria (i.e., covalent, closed, circular in shape, and similarly sized). The internal inheritance, differentiation, and gene flow transfer between the nuclear and extranuclear genomes in plants, animals, and bacteria are still being intensely studied [80, 81].

Compared to certain land plants, the genus *Ulva* has no large typical IR within the chloroplast genome to inhibit itself from undergoing gene recombination [82–84]; this could partly

explain the introns presence in these prokaryote-derived chloroplasts [85]. Nevertheless, almost no recombination occurs between coding genes within the genus *Ulva*.

SSRs are valuable molecular markers with a high degree of intraspecific variation that are used in investigations focused on population genetics and polymorphism [86–89]. The present study results agreed with previous findings that reported that SSRs from chloroplast generally consist of short polyA or polyT repeats rather than tandem G or C repeats [90].

Based on the hypothesis of the endosymbiotic origin of chloroplasts [91], the non-species-specific IR (complementary/forward/palindromic) distributed among the chloroplast genomes of *Ulva* could be obtained in two ways: (1) We suggest that genes recombined in the ancestral cyanobacteria before entering relatively stable chloroplasts. (2) Alternatively, genes were acquired from the bacteria through horizontal gene transfer, as shown in the chloroplast genome of *Bryopsis plumosa* [92].

Specific spaces between coding genes are associated with the heterogeneity of evolutionary rates [93]. Examples include the *rnl-trnK*, *cox3-atp6*, and *trnP-rnl* spacers in the brown algae mitogenome [94], where tandem repeats enable the development of potential markers at various taxonomic levels [95]. The tandem repeats, which might be exclusive within the genus *Ulva*'s subspecies or neighboring species, are more abundant in the mitochondria than in the chloroplast. However, approximately half of all tested sequences in *U. compressa* did not contain the exact predicted tandem repeats owing to random or uneven variability, or fewer repeating numbers, which lowered its value as a potential species indicator (Fig 6). The fact that more intra-species shared tandem repeats instead of complementary/forward/palindromic repeats in fast-evolving mitogenomes might support the non-endogenous hypothesis that these tandem repeats are inherited from intracellular genetic material [79].

Typical tandem repeats in the extranuclear genome

Although the tandem repeat sequence TRg is not unique to *U. compressa*, considering the whole analyzed data set, it could be used to accurately identify the organism from the major green tide forming *Ulva* species of the Yellow Sea using a non-sequencing approach.

Although no such TRg sequence exists in the chloroplast or mitochondrial genomes of *U. ohnoi*, TRg and related sequences of the nuclear genome were amplified, with which the corresponding sequence in *U. compressa*'s mitochondrial genome shared 99.8% homology. Considering endosymbiotic gene transfer [96, 97] in *Ulva* genomes, it is tempting to speculate that TRg originated from an ancestral mitogenome and integrated into the nuclear genome potentially via replicative transposition (*U. mutabilis*) or conservative transposition (*U. ohnoi*) as previously suggested.

Certain *U. compressa* strains sampled from the Baltic and North Sea contained primer-derived sequences which varied in lengths. This demonstrated that mutations (via insertion, deletion, or recombination) were common in the mitochondrial genome, or that they entered the nuclear genome via transposition, a phenomenon similar to that observed in the chloroplast genome [97–101]. Sequences of various length were amplified in *U. torta* and *U. rigida* using the same primers (Fig 8). The primer are thus not so conserved, and work for several *Ulva* species.

We determined that TRg was absent in the extranuclear genomes of *U. linza* and *U. fasciata* (*U. lactuca*). Thus, the amplicons might result from the respective nuclear genome sequences (Fig 8). Conservative transposition might explain this occurrence similar to *U. ohnoi* (Fig 7B). However, this phenomenon was not observed in *U. linza* collected from the Yellow Sea, where the TRg-related region may have been lost entirely or mutated in the primer region. The high

reproduction rate that causes green tide blooms such in the Yellow Sea [102] might increase the probability of mutation events.

Conclusions

In the present study, the complete chloroplast genome of *U. compressa* (RD9023) was sequenced and annotated, including its protein coding regions and repeat sequences. The study focused on assessing genome length, homology, gene order and direction, intron size, selection strength, and substitution rate. The generated phylogenetic tree was analyzed on the basis of single and aligned genes in the chloroplast genome compared to mitogenome genes of *Ulva* to detect evolutionary trends. We systematically screened for specific regions among representative fragments of all mitogenomes and chloroplast genomes of the *Ulva* gene for DNA barcoding purposes. Several strains from China and Europe have complemented the screening process. The chloroplast genomes of *U. mutabilis* and *U. compressa* shared the same gene queue and high homology of each gene and were clustered together in the phylogenetic tree. In addition, there were several forward/palindromic/tandem repeats in common, similar to those in *U. prolifera* and *U. linza*, but different from those in *U. fasciata* and *U. ohnoi*.

PetA performed as an appropriate barcode marker gene following the results of the *tufA*-based analysis. Furthermore, the chloroplast genome and mitogenomes have been compared. In agreement with the extranuclear genomes verified by the phylogenetic analysis of the aligned genes of both chloroplast genomes and mitogenomes, *petA* showed the evolutionary trends of the reported species in the *Ulva* genome. Non-species-specific complementary/forward/palindromic interval repetitions have been more frequently distributed in chloroplast genomes than in mitogenomes. Conversely, tandem repetitions, which may be exclusive to neighboring species in the *Ulva* genus, were more abundant in mitochondria than in chloroplasts. The amplification of significant tandem repeats, such as TRG-related regions among mitogenomes, supported the identification of *U. mutabilis* and *U. compressa* within the green tide species sampled in this study of the Yellow Sea in China. However, this scenario was quite different for the European *Ulva* species, as this gene sequence varied across species boundaries and was not specific to the species collected in the North Sea and the Baltic Sea. It demonstrates the significance of *Ulva* sampling on a global scale for the development of a non-sequencing approach.

Our study is the foundation for future research on sequencing and characterization of other green tide algal species and screening for novel extranuclear markers.

Since there was a lack of strains in this study, more samples from different locations should be used to verify the results. It is also worth looking into the gene flow of high homology tandem repeats between nuclear and extranuclear genomes in future studies.

Supporting information

S1 File.
(PDF)

S2 File.
(PDF)

Acknowledgments

We gratefully appreciate Michiel Kwantes for fruitful discussions. We thank all three reviewers, who contributed to improving the manuscript

Author Contributions

Conceptualization: Chuner Cai, Peimin He, Thomas Wichard.

Data curation: Chuner Cai, Hui Zhao, Peimin He.

Formal analysis: Hui Zhao.

Funding acquisition: Chuner Cai, Thomas Wichard.

Investigation: Kai Gu, Sophie Steinhagen.

Methodology: Chuner Cai, Kai Gu, Hui Zhao.

Project administration: Chuner Cai, Thomas Wichard.

Resources: Kai Gu, Sophie Steinhagen.

Software: Chuner Cai.

Supervision: Peimin He, Thomas Wichard.

Visualization: Hui Zhao.

Writing – original draft: Chuner Cai, Kai Gu, Thomas Wichard.

Writing – review & editing: Chuner Cai, Kai Gu, Hui Zhao, Sophie Steinhagen, Peimin He, Thomas Wichard.

References

1. Tang Q, Pang K, Yuan X, Xiao S (2020) A one-billion-year-old multicellular chlorophyte. *Nature Ecology & Evolution*. 4: 543–549. <https://doi.org/10.1038/s41559-020-1122-9> PMID: 32094536
2. Leliaert F, Tronholm A, Lemieux C, Turmel M, DePriest MS, et al. (2016) Chloroplast phylogenomic analyses reveal the deepest-branching lineage of the Chlorophyta, Palmophyllophyceae class. nov. *Scientific Reports* 6: 25367. <https://doi.org/10.1038/srep25367> PMID: 27157793
3. Sun L, Fang L, Zhang Z, Chang X, Penny D, et al. (2015) Chloroplast phylogenomic inference of green algae relationships. *Scientific Reports* 6: 20528.
4. Lemieux C, Vincent AT, Labarre A, Otis C, Turmel M (2015) Chloroplast phylogenomic analysis of chlorophyte green algae identifies a novel lineage sister to the Sphaeropleales (Chlorophyceae). *BMC Evolutionary Biology* 15: 264. <https://doi.org/10.1186/s12862-015-0544-5> PMID: 26620802
5. Lemieux C, Otis C, Turmel M (2014) Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. *Bmc Evolutionary Biology* 14: 1–15. <https://doi.org/10.1186/1471-2148-14-1> PMID: 24382122
6. Turmel M, de Cambiaire J-C, Otis C, Lemieux C (2016) Distinctive Architecture of the Chloroplast Genome in the Chlorodendrophycean Green Algae *Scherffelia dubia* and *Tetraselmis* sp. CCMP 881. *PloS One* 11: e0148934. <https://doi.org/10.1371/journal.pone.0148934> PMID: 26849226
7. Turmel M, Brouard J-S, Gagnon C, Otis C, Lemieux C (2008) Deep division in the Chlorophyceae (Chlorophyta) revealed by chloroplast phylogenomic analyses. *Journal of Phycology* 44: 739–750. <https://doi.org/10.1111/j.1529-8817.2008.00510.x> PMID: 27041432
8. Wheeler GL, Miranda-Saavedra D, Barton GJ (2008) Genome analysis of the unicellular green alga *Chlamydomonas reinhardtii* indicates an ancient evolutionary origin for key pattern recognition and cell-signaling protein families. *Genetics* 179: 193–197. <https://doi.org/10.1534/genetics.107.085936> PMID: 18493051
9. Marin B (2012) Nested in the Chlorellales or independent class? Phylogeny AND CLASSIFICATION of the Pedinophyceae (Viridiplantae) revealed by molecular phylogenetic analyses of complete nuclear and plastid-encoded rRNA operons. *Protist* 163: 778–805. <https://doi.org/10.1016/j.protis.2011.11.004> PMID: 22192529
10. Cocquyt E, Verbruggen H, Leliaert F, De Clerck O (2010) Evolution and cytological diversification of the green seaweeds (Ulvophyceae). *Molecular Biology and Evolution* 27: 2052–2061. <https://doi.org/10.1093/molbev/msq091> PMID: 20368268
11. Hayden SH (2002) Phylogenetic systematics of the Ulvaceae (Ulvales, Ulvophyceae) using chloroplast and nuclear DNA sequences. *Journal of Phycology* 38: 1200–1212.

12. Pombert J-Fo, Otis C, Lemieux C, Turmel M (2005) The chloroplast genome sequence of the green alga *Peudendoctonium akinetum* (Uvophyceae) reveals unusual structural features and new insights into the branching order of chlorophyte lineages. *Molecular Biology and Evolution* 22: 1903–1918. <https://doi.org/10.1093/molbev/msi182> PMID: 15930151
13. Cai C, Wang L, Zhou L, He P, Jiao B (2017) Complete chloroplast genome of green tide algae *Ulva flexuosa* (Ulvophyceae, Chlorophyta) with comparative analysis. *PloS one*. pp. e0184196. <https://doi.org/10.1371/journal.pone.0184196> PMID: 28863197
14. Cai C, Liu F, Jiang T, Wang L, Jia R, et al. (2018) Comparative study on mitogenomes of green tide algae. *Genetica* 146: 529–540. <https://doi.org/10.1007/s10709-018-0046-7> PMID: 30377874
15. Rybak A, Czerwoniec A, Gabka M, Messyasz B (2014) *Ulva flexuosa* (Ulvaceae, Chlorophyta) inhabiting inland aquatic ecosystems: molecular, morphological and ecological discrimination of subspecies. *European Journal of Phycology* 49: 471–485.
16. Steinhagen S, Barco A, Wichard T, Weinberger F (2019) Conspecificity of the model organism *Ulva mutabilis* and *Ulva compressa* (Ulvophyceae, Chlorophyta). *Journal of Phycology* 55: 25–36. <https://doi.org/10.1111/jpy.12804> PMID: 30367499
17. Hughey JR, Maggs CA, Mineur F, Jarvis C, Miller KA, et al. (2019) Genetic analysis of the Linnaean *Ulva lactuca* (Ulvales, Chlorophyta) holotype and related type specimens reveals name misapplications, unexpected origins, and new synonymies. *J Phycol* 55: 503–508. <https://doi.org/10.1111/jpy.12860> PMID: 30907438
18. Cui J, Monotilla AP, Zhu W, Takano Y, Shimada S, et al. (2018) Taxonomic reassessment of *Ulva prolifera* (Ulvophyceae, Chlorophyta) based on specimens from the type locality and Yellow Sea green tides. *Phycologia* 57: 692–704.
19. Zhao J, Jiang P, Liu Z, Wei W, Lin H, et al. (2013) The yellow sea green tides were dominated by one species, *Ulva (Enteromorpha) prolifera*, from 2007 to 2011. *Chinese Science Bulletin* 58: 2298–2302.
20. Ye NH, Zhang XW, Mao YZ, Liang CW, Xu D, et al. (2011) 'Green tides' are overwhelming the coastline of our blue planet: taking the world's largest example. *Ecological Research* 26: 477–485.
21. Leliaert F, Zhang XW, Ye NH, Malta EJ, Engelen AH, et al. (2009) Identity of the Qingdao algal bloom. *Phycological Research* 57: 147–151.
22. Sun S, Wang F, Li C, Qin S, Zhou M, et al. (2008) Emerging challenges: Massive green algae blooms in the Yellow Sea. *Nature Precedings* 1: 2266.
23. Malta Erik-Jan, Draisma Stefano, Kamermans Pauline. (1999) Free-floating *Ulva* in the southwest Netherlands: species or morphotypes? a morphological, molecular and ecological comparison. *British Phycological Bulletin* 34: 443–454.
24. Wang J, Jiang P, Cui Y, Li N, Wang M, et al. (2010) Molecular analysis of green-tide-forming macroalgae in the Yellow Sea. *Aquatic Botany* 93: 25–31.
25. Huo Y, Han H, Shi H, Wu H, Zhang J, et al. (2015) Changes to the biomass and species composition of *Ulva* sp. on *Porphyra aquaculture* rafts, along the coastal radial sandbank of the Southern Yellow Sea. *Marine Pollution Bulletin* 93: 210–216. <https://doi.org/10.1016/j.marpolbul.2015.01.014> PMID: 25691340
26. Zhang J, Liu C, Yang L, Gao S, Ji X, et al. (2015) The source of the *Ulva* blooms in the East China Sea by the combination of morphological, molecular and numerical analysis. *Estuary Coast Shelf S* 164: 418–424.
27. Zhang J, Kim JK, Yarish C, He P (2016) The expansion of *Ulva prolifera* OF Muller macroalgal blooms in the Yellow Sea, PR China, through asexual reproduction. *Marine Pollution Bulletin* 104: 101–106. <https://doi.org/10.1016/j.marpolbul.2016.01.056> PMID: 26856643
28. Li Y, Song W, Xiao J, Wang Z, Fu M, et al. (2014) Tempo-spatial distribution and species diversity of green algae micro-propagules in the Yellow Sea during the large-scale green tide development. *Harmful Algae* 39: 40–47.
29. Xiao J, Li Y, Song W, Wang Z, Fu M, et al. (2013) Discrimination of the common macroalgae (*Ulva* and *Blidingia*) in coastal waters of Yellow Sea, northern China, based on restriction fragment-length polymorphism (RFLP) analysis. *Harmful Algae* 27: 130–137.
30. Liu F, Pang S, Chopin T, Gao S, Shan T, et al. (2013) Understanding the recurrent large-scale green tide in the Yellow Sea: Temporal and spatial correlations between multiple geographical, aquacultural and biological factors. *Marine Environmental Research* 83: 38–47. <https://doi.org/10.1016/j.marenvres.2012.10.007> PMID: 23176870
31. Liu D, Keesing JK, Dong Z, Zhen Y, Di B, et al. (2010) Recurrence of the world's largest green-tide in 2009 in Yellow Sea, China: *Porphyra yezoensis* aquaculture rafts confirmed as nursery for macroalgal blooms. *Marine Pollution Bulletin* 60: 1423–1432. <https://doi.org/10.1016/j.marpolbul.2010.05.015> PMID: 20541229

32. Cui JJ, Zhang JH, Huo YZ, Zhou LJ, Wu Q, et al. (2015) Adaptability of free-floating green tide algae in the Yellow Sea to variable temperature and light intensity. *Marine Pollution Bulletin* 101: 660–666. <https://doi.org/10.1016/j.marpolbul.2015.10.033> PMID: 26573134
33. Huo Y, Hua L, Wu H, Zhang J, Cui J, et al. (2014) Abundance and distribution of *Ulva* microscopic propagules associated with a green tide in the southern coast of the Yellow Sea. *Harmful Algae* 39: 357–364.
34. Huo Y, Zhang J, Chen L, Hu M, Yu K, et al. (2013) Green algae blooms caused by *Ulva prolifera* in the southern Yellow Sea: Identification of the original bloom location and evaluation of biological processes occurring during the early northward floating period. *Limnology Oceanography* 58: 2206–2218.
35. Han W, Chen LP, Zhang JH, Tian XL, Hua L, et al. (2013) Seasonal variation of dominant free-floating and attached *Ulva* species in Rudong coastal area, China. *Harmful Algae* 28: 46–54.
36. Saunders GW, Kucera H (2010) An evaluation of *rbcL*, *tufA*, UPA, LSU and ITS as DNA barcode markers for the marine green macroalgae. *Cryptogamie Algologie* 31: 487–528.
37. Song W, Han H, Wang Z, Li Y (2019) Molecular identification of the macroalgae that cause green tides in the Bohai Sea, China. *Aquatic Botany* 156: 38–46.
38. Zhang QC, Liu Q, Kang ZJ, Yu RC, Yan T, et al. (2015) Development of a fluorescence in situ hybridization (FISH) method for rapid detection of *Ulva prolifera*. *Estuarine, Coastal and Shelf Science* 163: 103–111.
39. He Peimin, Zhang Jianheng, Huo Yuanzi, Cai C (2019) Green-tide in China. Beijing: Science Press.
40. Ott FD (1965) Synthetic media and techniques for the xenic cultivation of marine algae and flagellate. *Virginia Journal of Science* 16: 205–218.
41. Coat G, Dion P, Noailles MC, Reviers BD, Fontaine JM, et al. (1998) *Ulva armoricana* (Ulvales, Chlorophyta) from the coasts of Brittany (France). II. Nuclear rDNA ITS sequence analysis. *European Journal of Phycology* 33: 81–86.
42. Melton JT III, Lopez-Bautista JM (2015) The chloroplast genome of the marine green macroalga *Ulva fasciata* Delile (Ulvophyceae, Chlorophyta). *Mitochondrial DNA* 28: 93–95. <https://doi.org/10.3109/19401736.2015.1110814> PMID: 26710262
43. Melton JT III, Leliaert F, Tronholm A, Lopez-Bautista JM (2015) The complete chloroplast and mitochondrial genomes of the green macroalga *Ulva* sp. una00071828 (Ulvophyceae, Chlorophyta). *Plos One* 10: e0121020 <https://doi.org/10.1371/journal.pone.0121020> PMID: 25849557
44. Wang L, Cai C, Zhou L, He P, Jiao B (2017) The complete chloroplast genome sequence of *Ulva linza*. *Conservation Genetics Resources* 9: 463–466.
45. Ewing B, Hillier LD, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Research* 8: 186–194. PMID: 9521922
46. Chang L, Shi L, Zhu Y, Chen H, Zhang J, et al. (2012) CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics* 13: 1–7. <https://doi.org/10.1186/1471-2164-13-1> PMID: 22214261
47. Laslett D, Canback B (2004) ARAGORN, a program to detect rRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* 32: 11–16. <https://doi.org/10.1093/nar/gkh152> PMID: 14704338
48. Krumsiek J, Arnold R, Rattei T (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* 23: 1026–1028. <https://doi.org/10.1093/bioinformatics/btm039> PMID: 17309896
49. Lohse M, Drechsel O, Bock R (2007) OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Current Genetics* 52: 267–274. <https://doi.org/10.1007/s00294-007-0161-y> PMID: 17957369
50. Rice P, Longden I, Bleasby A (2000) EMBOSS: The European molecular biology open software suite. *Trends in Genetics* 16: 276–277. [https://doi.org/10.1016/s0168-9525\(00\)02024-2](https://doi.org/10.1016/s0168-9525(00)02024-2) PMID: 10827456
51. Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27: 573–580. <https://doi.org/10.1093/nar/27.2.573> PMID: 9862982
52. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, et al. (2001) REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* 29: 4633–4642. <https://doi.org/10.1093/nar/29.22.4633> PMID: 11713313
53. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J (2010) KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genomics, Proteomics & Bioinformatics* 8: 77–80. [https://doi.org/10.1016/S1672-0229\(10\)60008-3](https://doi.org/10.1016/S1672-0229(10)60008-3) PMID: 20451164

54. Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3: 418. <https://doi.org/10.1093/oxfordjournals.molbev.a040410> PMID: 3444411
55. Wichard T, Charrier B, Mineur F, Bothwell JH, De Clerck O, et al. (2015) The green seaweed *Ulva*: a model system to study morphogenesis. *Frontiers in Plant Science* 6: 72. <https://doi.org/10.3389/fpls.2015.00072> PMID: 25745427
56. De Clerck O, Kao S-M, Bogaert KA, Blomme J, Foflonker F, et al. (2018) Insights into the evolution of multicellularity from the sea lettuce genome. *Current Biology* 28: 2921–2933. <https://doi.org/10.1016/j.cub.2018.08.015> PMID: 30220504
57. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7: 539. <https://doi.org/10.1038/msb.2011.75> PMID: 21988835
58. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096> PMID: 29722887
59. Ramanathan KR (1939) The morphology, cytology, and alternation of generations in *Enteromorpha compressa* (L.) Gr. var. *lingulata* (J. Ag.) Hauck. *Annals of Botany* 3: 375–398.
60. Eaton JW, Brown JG, Round FE (1966) Some observations on polarity and regeneration in *Enteromorpha*. *British Phycological Bulletin* 3: 53–62.
61. Müller-Stoll WR (1952) Über Regeneration und Polarität bei *Enteromorpha* Flora oder Allgemeine Botanische Zeitung 139: 148–180.
62. Wu M, Lan S, Cai B, Chen S, Chen H, et al. (2015) The complete chloroplast genome of *Guadua angustifolia* and comparative analyses of neotropical-paleotropical bamboos. *PloS One* 10: e0143792–e0143792. <https://doi.org/10.1371/journal.pone.0143792> PMID: 26630488
63. Mitsunashi C, Teramura H, Shimada H (2020) Construction of genomic marker sets based on the chloroplast genome of a green alga, *Ulva pertusa* (syn. *Ulva australis*), leads to simple detection of *Ulva* species. *Genes & Genetic Systems*.
64. Kazi MA, Kavale MG, Singh VV (2016) Morphological and molecular characterization of *Ulva chaugulii* sp. nov., *U. lactuca* and *U. ohnoi* (Ulvophyceae, Chlorophyta) from India. *Phycologia* 55: 45–54.
65. Lee HW, Kang JC, Kim MS (2019) Taxonomy of *Ulva* causing blooms from Jeju Island, Korea with new species, *U. pseudo-ohnoi* sp. nov. (Ulvales, Chlorophyta). *Algae* 34: 253–266.
66. Shi C, Wang S, Xia E-H, Jiang J-J, Zeng F-C, et al. (2016) Full transcription of the chloroplast genome in photosynthetic eukaryotes. *Scientific Reports* 6: 30135. <https://doi.org/10.1038/srep30135> PMID: 27456469
67. Duminil J (2014) Mitochondrial genome and plant taxonomy. *Methods in Molecular Biology* (Clifton, NJ) 1115: 121–140. https://doi.org/10.1007/978-1-62703-767-9_6 PMID: 24415472
68. Liu F, Li X, Che Z (2017) Mitochondrial genome sequences uncover evolutionary relationships of two *Sargassum* subgenera, *Bactrophyucus* and *Sargassum*. *Journal of Applied Phycology* 29: 3261–3270.
69. Curci PL, De Paola D, Danzi D, Vendramin GG, Sonnante G (2015) Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. *PloS One* 10: e0120589. <https://doi.org/10.1371/journal.pone.0120589> PMID: 25774672
70. Williams AV, Boykin LM, Howell KA, Nevill PG, Small I (2015) The complete sequence of the *Acacia ligulata* chloroplast genome reveals a highly divergent *clpp1* gene. *PLoS One* 10: e0125768. <https://doi.org/10.1371/journal.pone.0125768> PMID: 25955637
71. Yao X, Tan Y-H, Liu Y-Y, Song Y, Yang J-B, et al. (2016) Chloroplast genome structure in *Ilex* (Aquifoliaceae). *Scientific Reports* 6: 28559. <https://doi.org/10.1038/srep28559> PMID: 27378489
72. Jiang G-F, Hinsinger DD, Strijk JS (2016) Comparison of intraspecific, interspecific and intergeneric chloroplast diversity in Cycads. *Scientific Reports* 6: 31473. <https://doi.org/10.1038/srep31473> PMID: 27558458
73. Willey DL, Gray JC (1988) Synthesis and assembly of the cytochrome b-f complex in higher plants. *Photosynthesis Research* 17: 125–144. <https://doi.org/10.1007/BF00047685> PMID: 24429665
74. Hauska G (2004) The isolation of a functional cytochrome b6f complex: From lucky encounter to rewarding experiences. *Photosynthesis Research* 80: 277–291. <https://doi.org/10.1023/B:PRES.0000030425.69092.4f> PMID: 16328826
75. Gray MW, Burger G, Lang BF (2001) The origin and early evolution of mitochondria. *Genome Biology* 2: reviews1018.1011. <https://doi.org/10.1186/gb-2001-2-6-reviews1018> PMID: 11423013

76. Shih PM, Matzke NJ (2013) Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. *Proceedings of the National Academy of Sciences* 110: 12355. <https://doi.org/10.1073/pnas.1305813110> PMID: 23776247
77. Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, et al. (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Molecular Biology and Evolution* 25: 748–761. <https://doi.org/10.1093/molbev/msn022> PMID: 18222943
78. Zimorski V, Ku C, Martin WF, Gould SB (2014) Endosymbiotic theory for organelle origins. *Current Opinion in Microbiology* 22: 38–48. <https://doi.org/10.1016/j.mib.2014.09.008> PMID: 25306530
79. Lang BF, Gray MW, Burger G (1999) Mitochondrial genome evolution and the origin of eukaryotes. *Annual Review of Genetics* 33: 351–397. <https://doi.org/10.1146/annurev.genet.33.1.351> PMID: 10690412
80. Fučíková K, Lewis PO, González-Halphen D, Lewis LA (2014) Gene arrangement convergence, diverse intron content, and genetic code modifications in mitochondrial genomes of Sphaeropleales (Chlorophyta). *Genome Biology and Evolution* 6: 2170–2180. <https://doi.org/10.1093/gbe/evu172> PMID: 25106621
81. Rodríguez-Salinas E, Riveros-Rosas H, Li Z, Fučíková K, Brand JJ, et al. (2012) Lineage-specific fragmentation and nuclear relocation of the mitochondrial *cox2* gene in chlorophycean green algae (Chlorophyta). *Molecular Phylogenetics and Evolution* 64: 166–176. <https://doi.org/10.1016/j.ympev.2012.03.014> PMID: 22724135
82. Sugiura M (1992) The chloroplast genome. *Plant molecular biology* 19: 149–168. <https://doi.org/10.1007/BF00015612> PMID: 1600166
83. Lü F, Xü W, Tian C, Wang G, Niu J, et al. (2011) The *Bryopsis hypnoides* plastid genome: multimeric forms and complete nucleotide sequence. *PloS One* 6: e14663. <https://doi.org/10.1371/journal.pone.0014663> PMID: 21339817
84. Zheng W, Chen J, Hao Z, Shi J (2016) Comparative analysis of the chloroplast genomic information of *Cunninghamia lanceolata* (Lamb.) hook with sibling species from the genera *Cryptomeria* d. Don, *Taiwania* Hayata, and *Calocedrus* Kurz. *International Journal of Molecular Sciences* 17: 1084–1100.
85. McFadden GI (2001) Chloroplast origin and integration. *Plant Physiology* 125: 50–53. <https://doi.org/10.1104/pp.125.1.50> PMID: 11154294
86. Xue J, Wang S, Zhou SL (2012) Polymorphic chloroplast microsatellite loci in *Nelumbo* (Nelumbonaceae). *American Journal of Botany* 99: 240–244.
87. Pan L, Fu J, Zhang R, Qin Y, Lu F, et al. (2017) Genetic diversity among germplasms of *Pitaya* based on SSR markers. *Scientia Horticulturae* 225: 171–176.
88. Chiappetta A, Muto A, Muzzalupo R, Muzzalupo I (2017) New rapid procedure for genetic characterization of Italian wild olive (*Olea europaea*) and traceability of virgin olive oils by means of SSR markers. *Scientia Horticulturae* 226: 42–49.
89. El Zerey-Belaskri A, Ribeiro T, Alcaraz ML, El Zerey W, Castro S, et al. (2018) Molecular characterization of *Pistacia atlantica* Desf. subsp. *atlantica* (Anacardiaceae) in Algeria: Genome size determination, chromosome count and genetic diversity analysis using SSR markers. *Scientia Horticulturae* 227: 278–287.
90. Kuang DY, Wu H, Wang YL, Gao LM, Zhang SZ, et al. (2011) Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics. *Genome* 54: 663–673. <https://doi.org/10.1139/g11-026> PMID: 21793699
91. Xiao-Ming Z, Junrui W, Li F, Sha L, Hongbo P, et al. (2017) Inferring the evolutionary mechanism of the chloroplast genome size by comparing whole-chloroplast genome sequences in seed plants. *Scientific Reports* 7: 1555. <https://doi.org/10.1038/s41598-017-01518-5> PMID: 28484234
92. Leliaert F, Lopez-Bautista JM (2015) The chloroplast genomes of *Bryopsis plumosa* and *Tydemania expeditiones* (Bryopsidales, Chlorophyta): compact genomes and genes of bacterial origin. *BMC Genomics* 16: 20. <https://doi.org/10.1186/s12864-014-1200-y> PMID: 25612629
93. Cho KS, Yun BK, Yoon YH, Hong SY, Mekapogu M, et al. (2015) Complete chloroplast genome sequence of tartary buckwheat (*Fagopyrum tataricum*) and comparative analysis with common buckwheat (*F. esculentum*). *PloS One* 10: e0125332. <https://doi.org/10.1371/journal.pone.0125332> PMID: 25966355
94. Liu F, Pang S (2015) Mitochondrial genome of *Turbinaria ornata* (Sargassaceae, Phaeophyceae): comparative mitogenomics of brown algae. *Current Genetics* 61: 621–631. <https://doi.org/10.1007/s00294-015-0488-8> PMID: 25893565
95. Nazareno AG, Carlsen M, Lohmann LG (2015) Complete chloroplast genome of *Tanaecium tetragonolobum*: The first Bignoniaceae plastome. *PloS One* 10: e0129930. <https://doi.org/10.1371/journal.pone.0129930> PMID: 26103589

96. Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Reviews Genetics* 5: 123–135. <https://doi.org/10.1038/nrg1271> PMID: 14735123
97. Bock R, Timmis JN (2008) Reconstructing evolution: gene transfer from plastids to the nucleus. *BioEssays* 30: 556–566. <https://doi.org/10.1002/bies.20761> PMID: 18478535
98. Yao X, Tang P, Li Z, Li D, Liu Y, et al. (2015) The first complete chloroplast genome sequences in actinidiaceae: Genome structure and comparative analysis. *PloS one* 10: e0129347. <https://doi.org/10.1371/journal.pone.0129347> PMID: 26046631
99. Choi KS, Son OG, Park S (2015) The chloroplast genome of *Elaeagnus macrophylla* and trnh duplication event in Elaeagnaceae. *PloS One* 10: e0138727. <https://doi.org/10.1371/journal.pone.0138727> PMID: 26394223
100. Raman G, Park S (2015) Analysis of the complete chloroplast genome of a medicinal plant, *Dianthus superbus* var. *longicalyncinus*, from a comparative genomics perspective. *PloS One* 10: e0141329. <https://doi.org/10.1371/journal.pone.0141329> PMID: 26513163
101. Lei W, Ni D, Wang Y, Shao J, Wang X, et al. (2016) Intraspecific and heteroplasmic variations, gene losses and inversions in the chloroplast genome of *Astragalus membranaceus*. *Scientific Reports* 6: 21669. <https://doi.org/10.1038/srep21669> PMID: 26899134
102. Zhang Y, He P, Li H, Li G, Liu J, et al. (2019) *Ulva prolifera* green-tide outbreaks and their environmental impact in the Yellow Sea, China. *National Science Review* 6: 825–838.