Plastid genome analysis of three Nemaliophycidae red algal species suggests environmental adaptation for iron limited habitats

Chung Hyun Cho¹, Ji Won Choi¹, Daryl W. Lam², Kyeong Mi Kim³, Hwan Su Yoon¹*

¹ Department of Biological Sciences, Sungkyunkwan University, Suwon, Korea, ² Department of Biological Sciences, University of Alabama, Tuscaloosa, Alabama, United States of America, ³ Marine Biodiversity Institute of Korea, Seocheon, Korea

* hsyoon2011@skku.edu

Abstract

The red algal subclass Nemaliophycidae includes both marine and freshwater taxa that contribute to more than half of the freshwater species in Rhodophyta. Given that these taxa inhabit diverse habitats, the Nemaliophycidae is a suitable model for studying environmental adaptation. For this purpose, we characterized plastid genomes of two freshwater species, Kumanoa americana (Batrachospermales) and Thorea hispida (Thoreales), and one marine species Palmaria palmata (Palmariales). Comparative genome analysis identified seven genes (ycf34, ycf35, ycf37, ycf46, ycf91, grx, and pbsA) that were different among marine and freshwater species. Among currently available red algal plastid genomes (127), four genes (pbsA, ycf34, ycf35, ycf37) were retained in most of the marine species. Among these, the pbsA gene, known for encoding heme oxygenase, had two additional copies (HMOX1 and HMOX2) that were newly discovered and were reported from previously red algal nuclear genomes. Each type of heme oxygenase had a different evolutionary history and special modifications (e.g., plastid targeting signal peptide). Based on this observation, we suggest that the plastid-encoded pbsA contributes to the iron controlling system in iron-deprived conditions. Thus, we highlight that this functional requirement may have prevented gene loss during the long evolutionary history of red algal plastid genomes.

Introduction

The red algal class Florideophyceae comprises 95% (6,748 spp. out of 7,100 spp.) of the Rhodophyta and encompass a biologically diversified group of taxa [1, 2]. Most of the red algal species (>95%) inhabit marine habitats, however about 5% are found in freshwater environments [3]. The Nemaliophycidae, one of five subclasses within Florideophyceae, contains more than half of these freshwater species. This subclass includes both marine and freshwater taxa with three exclusively freshwater orders (Balbianiales, Batrachospermales, Thoreales), six exclusively marine orders (Rhodachlyales, Balliales, Nemaliales, Entwisleiales, Colaconematales, Palmariales), and one order (Acrochaetiales) with both freshwater and marine species [1, 2, 4,
Among the freshwater orders, the Batrachospermales and Thoreales have more than half of the freshwater species in all of the Rhodophyta [6]. Thus, a comparison of Nemaliophycidae plastid genomes of taxa from freshwater and marine habitats may provide insights into environmental adaptation [1].

A previous study demonstrated that the freshwater angiosperm *Najas flexilis* (water nymph) adapted to the aquatic environment from its terrestrial ancestor by the complete loss of the *ndh* gene family in plastid genome [7]. The NDH gene complexes encode for the NAD (P)H dehydrogenase complex that increases photosynthetic efficiency at variable light intensities in terrestrial habitats. In its transition to the aquatic environment, *N. flexilis* did not require resistance to high light stress (due to the refractive properties of water) and therefore the *ndh* gene family had been lost. Likewise, similar gene loss or retention events may be present in the evolution of red algal plastid genomes during their transitions from marine habitats to freshwater systems.

To date, 99 florideophycean plastid genomes (cf. 127 red algal plastid genomes including three new genomes) are available in the NCBI organelle database, including 23 Nemaliophycidae that have been detailed in three recent papers [8–10]. To extend our understanding of red algal plastid evolution as it relates to the habitat adaptation, we completely sequenced and annotated three new plastid genomes for Nemaliophycidae, including one marine (*Palmaria palmata*) and two freshwater species (*Kumanoa americana, Thorea hispida*). From a comparative analysis of plastid genomes, we seek to identify plastid genes involved in the transition between marine and freshwater red algae and their physiological implications.

**Materials and methods**

**Whole genome sequencing and plastid genome construction**

Culture strains of two filamentous freshwater species of *Kumanoa americana* (hsy120, isolated by Franklyn D. Ott from a stream in Mississippi, USA) and *Thorea hispida* (hsy077, isolated by F. Ott from the Kaw river in Kansas, USA) were harvested with gentle centrifugations from the culture flask. Thalli of *Palmaria palmata* (commercially sold as dulse) were collected from Reid State Park in Maine, USA on 27 Aug. 2010 by HSY. Genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) and purified by LaboPass™ DNA Isolation Kit (Cosmo Genetech, Seoul, Korea). Genome sequence data were generated using the Ion Torrent PGM (Thermo Fisher Scientific, San Francisco, California, USA) Next-Generation Sequencing (NGS) platform. The sequencing libraries were prepared using the Ion Xpress Plus gDNA Fragment Library Preparation kit for 200 bp or 400 bp libraries. The library amplification and DNA sequencing were conducted by either Ion PGM Template OT2 200 or 400 Kits and Ion PGM Sequencing OT2 200 or 400 Kit for the Ion Torrent PGM platform.

From NGS genome sequencing data, short raw reads (< 50 bp) were removed completely from the analysis and the remainder of raw reads were de novo assembled into contigs using CLC Genomics Workbench 5.5.1 (CLC Bio., Aarhus, Denmark) and MIRA3 Assembler [11]. To obtain a plastid consensus sequence, contigs were sorted by tBLASTn (e-value: 1e-10) using the protein sequence of red algal plastid genes as a reference (i.e. *Chondrus crispus, Callicrathron tuberculatum*) [12, 13]. After the reassembly process, we obtained the circular plastid genomes and those circular genomes were confirmed by MUMmerplot [14] and by comparing with reference genomes to check for completeness. The sequences were verified with read-mapping tools implemented in CLC Genomics Workbench to correct for any sequencing errors or gaps.

Protein coding genes were manually annotated following the procedure described in Song et al., 2016 [15] using 'Bacteria and Archaea (11)’ genetic code and the NCBI database for non-
redundant (nr) protein sequences. To search the ribosomal RNA sequences, we applied RNAmmer v1.2 Server [16] using the option of Bacteria, and then re-confirmed by BLASTn. Introns, tRNAs, and other small RNA were searched by ARAGORN [17] and Rfam cmscan v1.1 [18]. Several important genes and ambiguous sequences were re-confirmed with PCR amplification followed by Sanger sequencing. The annotated plastid genomes were visualized using OrganellarGenomeDraw v1.2 [19]. Comparative analysis of the genome structure was accomplished using UniMoG v1.0 [20] and Mauve Genome Alignment v2.2.0 [21, 22] through the Geneious plug-in using the default setting [23].

A statistical test for the habitat-gene correlation

Habitat information of each species was referred from previous studies. Ott [5] summarized detail about the habitats and isolation history for most nemaliophycidaean species. AlgaeBase [2] provided information about authentic references with the type locality and its distribution. Based on the habitat and gene presence/absence information in the plastid genome, we performed a chi-square analysis that assessed the correlation between habitat and putative habitat-specific genes. We used the chi-square test incorporated into R package [24]. The p-value cut-off (<0.01) was adjusted to reject the null hypothesis.

Phylogenetic analysis

The red algal protein sequences of heme oxygenase were collected from NCBI public databases, which include transcriptome and genome data. The heme oxygenase homologs were obtained by BLASTp against the NCBI database with an adjusted threshold-cut to 500 top matches of red algal heme oxygenase. To discover the nuclear heme oxygenase, we also surveyed heme oxygenase genes from the published [25–28] or unpublished (H.S. Yoon et al.) whole genome data as described in S5 Table. Multiple sequence alignments were performed using MAFFT version 7 [29] with the default options. These alignments were refined manually based on conserved domains. Maximum likelihood-based phylogenetic analysis and bootstrap methods (MLB) were conducted using IQ-TREE [30] with 1,000 ultrafast bootstrap replications. The evolutionary model was ‘LG + I + G4’ [31], which was automatically selected by the model test option incorporated in IQ-TREE. Finally, highly divergent or contaminant sequences, which showed a long-branch or taxonomical mismatch to sister taxa, were removed from the following analysis.

Protein domain predictions

Protein domains were searched by the NCBI conserved domain searching tool [32] to predict their functions and conserved motifs. To predict gene localization, TargetP [33] and ChloroP [34] were used based on transit peptide sequences. Transmembrane domain regions were identified using TMHMM program [35]. The molecular function of heme oxygenase and its molecular interaction were surveyed based on KEGG pathway [36].

Gene network analysis

The dataset of heme oxygenase genes for network analysis was collected from NCBI based on BLASTp with an adjusted threshold-cut to the top 500 matches of red algal heme oxygenase (S2 Fig). The genes containing incomplete heme oxygenase domains were removed from the dataset, therefore, only complete domains were used for network analysis [32]. EGN (Evolutionary Gene and Genome Network) was performed to build a gene network based on protein similarity [37]. For comparative analysis, network connection thresholds were set at 1e-05 in
e-value, identities at 20%, hits with at least 20% of the shortest sequence, and query coverage of both sequences at 70%. The resulting network was visualized with the aid of Cytoscape program [38].

**Result and discussion**

**General features of three plastid genomes**

A total of 1.73 Gbp, 1.94 Gbp, and 1.10 Gbp of raw sequence data were produced for *Palmaria palmata*, *Kumanoa americana*, and *Thorea hispida*, respectively (see details in S1 Table). The average coverage for the plastid genomes was 1,003x in *T. hispida*, 215x in *K. americana*, and 343x in *P. palmata*.

Three complete plastid genomes (Fig 1A) were manually annotated based on published red algal plastid references [39]. Table 1 summarizes the characteristics of plastid genomes of the Florideophyceae [8–10, 12, 13, 39–53]. The total size of the plastid genome was 184,026 bp in *K. americana* (GenBank accession number NC_031178), containing 194 protein-coding genes (CDS), while *T. hispida* (GenBank accession number NC_031171) was 175,193 bp in size including 192 CDSs. The *P. palmata* plastid genome (GenBank accession number NC_031147) was 192,961 bp in size with 203 CDSs. The GC content of *K. americana* was 29.3%, which was similar to *T. hispida* (28.3%), but lower than that of *P. palmata* (33.9%). The high GC content in *P. palmata* was more similar to that of the Bangiophyceae (average of 11 spp.: 33.1%) than other Florideophyceae species (average of 102 spp.: 29.3%).

Comparing the genome architecture, two major differences were evident among the three Nemaliophycidae species (Fig 1B). First, two copies of ribosomal RNA (rRNA) operon were present in *P. palmata*, whereas *K. americana* and *T. hispida* have only a single rRNA operon (5S, 23S, 16S rRNA) like as most of florideophycean species (Table 1). It has been reported that the plastid genome structures are highly conserved among four florideophycean subclasses (i.e., Nemaliophycidae, Corallinophycidae, Ahnfeltiophycidae, Rhodymeniophycidae) [39]. Second, *K. americana* had a large inversion between chlL and rRNA operon region that differs from other two species.

**Specific gene loss in freshwater Nemaliophycidae species**

Gene contents were generally conserved among the three Nemaliophycidae plastid genomes, however, there were some differences (S1 Fig, S2 Table). For example, the ycf91 gene was present only in two freshwater species of *K. americana* and *T. hispida*, but absent in marine *P. palmata*. In addition, six genes (*ycf34, ycf35, ycf37, ycf46, grx*, and *pbsA*) were preserved only in the marine *P. palmata*, which were absent in the freshwater *K. americana* and *T. hispida*. Therefore, these seven genes were candidates of habitat-specific plastid gene (i.e., marine vs. freshwater specific).

In order to broaden the investigation of these putative habitat-specific genes, we extended the survey to include all currently available 127 red algal plastid genomes, which include 16 freshwater species, 109 marine species, and two brackish species (S3 Table). From the data in S3 Table, we calculated both the habitat-specific gene concordance rate and the p-value from the chi-square test for seven genes for whether these genes were correlated to the habitat type. Interestingly, we discovered that four of these genes (*pbsA, ycf34, ycf35, and ycf46*) have higher than 80% concordance rate with statistically significant support (p-values = < 0.01). These results suggest that the presence of these four genes is significantly different between marine and freshwater habitats. For example, the *pbsA* gene showed 84.1% of habitat-specific gene concordance rate (p-value = 3.17E-07). Except for four species (*Bangia atropurpurea, Sheathia arcuata, Paralemanea sp., Sirodotia delicatula*), 12 of 16 freshwater species lacked the *pbsA*
Three plastid genomes of Nemaliophyceae

A

Kumanoa americana
chloroplast genome
164,026 bp

Thorea hispida
chloroplast genome
176,193 bp

Palmaria palmata
chloroplast genome
192,961 bp

B

Palmaria palmata

\[ \text{dnaK} \sim \text{psaM} \rightarrow \text{rRNA Operon} \rightarrow \text{rps6} \sim \text{chL} \]

Thorea hispida

\[ \text{dnaK} \sim \text{psaM} \rightarrow \text{rRNA Operon} \rightarrow \text{rps6} \sim \text{chL} \]

Kumanoa americana

\[ \text{dnaK} \sim \text{psaM} \rightarrow \text{rRNA Operon} \rightarrow \text{rps6} \sim \text{chL} \]

Annotations:
- Photosystem I
- Photosystem II
- Cytochrome b/f complex
- ATP synthase
- Rubisco large subunit
- RNA polymerase
- Ribosomal proteins (SSU)
- Ribosomal proteins (LSU)
- clpP, matK
- Other genes
- Hypothetical chloroplast reading frames (ycf)
- ORFs
- Transfer RNAs
- Ribosomal RNAs
- Introns
gene in the plastid genomes. Likewise, the habitat-specific gene concordance rates of ycf34, ycf35, ycf46 and their p-value were 85.6%, 81.8%, 87.9% and 4.83E-10, 2.42E-03, 5.43E-12, respectively.

To inspect the association between the phylogenetic relationship and the habitat-gene pattern, we mapped the presence or absence of these four genes on the ML phylogeny with habitat information (Fig 2). According to the result, the losses (i.e., absence) of four habitat-specific genes were more likely due to a phylogenetic pattern as compared to random events. For instance, all four genes were absent in four Cyanidiales species, which are all freshwater species. Two exclusive freshwater orders of Thoreales (T. hispida) and Batrachospermales (eight species) were mostly absent of these genes that were clearly different from marine orders that contained these genes (i.e., one species of Palmariales and 16 spp. of Nemaliales). It is interesting that some mangrove species (i.e., Bulboplasting aphyrenoidosa, Caloglossa spp. Bostrychia spp.) [54–56] and two parasitic species (i.e., Polysiphonia infestans, Choreocolax polysiphoniae) [57, 58], which have significantly reduced plastid genomes, lost these genes. However, there were some exceptional cases: two marine species of Hildenbrandiales (Apophlaea sinclairii, Hildenbrandia rubra) together with one freshwater species (H. rivularis) were absent of all these genes, same as in two marine Corallinophycidae species (Sporolithon durum and Calliarthron tuberculatum). Based on this observation (Fig 2), we postulate that four habitat-specific genes (pbsA, ycf34, ycf35, and ycf46) were adapted to freshwater habitat during their evolutionary history.

To find a functional relevance of these genes in environmental adaptation, we focused on the in silico functional analysis. Because any functions were reported for ycf genes, a conserved hypothetical protein family, we selected only on the pbsA (heme oxygenase) gene for downstream analyses.

Heme oxygenase

The function of heme oxygenase is generally known for the degradation of a heme to a biliverdin and is involved in the production of phycobilins [59]. For instance, the heme oxygenase degrades heme that is an essential step in the phycobilin biosynthesis in Cyanidium caldarium (Cyanidiophyceae) [60]. During a heme degradation, iron ions are released and those ions play an essential role in the iron recycling pathway [61, 62].

While searching for pbsA genes in all available red algal genome data (S5 Table), two additional heme oxygenase genes were newly discovered from nuclear genome data. Both the newly discovered two heme oxygenases and pbsA had a conserved heme oxygenase domain (CDD name: HemeO superfamily), with conserved heme binding pockets (blue asterisk in Fig 3). According to recent studies about heme oxygenase in Chlamydomonas reinhardtii (Chlorophyta), two distinct types of nuclear-encoded heme oxygenase have been called as HMOX1 (plant type) and HMOX2 (animal type) [63]. However, there was no plastidal heme oxygenase (pbsA) in green algae and we could not find it from a currently available nuclear genome. Compared to the green algal heme oxygenase, we named three red algal heme oxygenase genes as HMOX1 and HMOX2 for nuclear copies and pbsA for the plastidial copy.

To identify the evolutionary history of three heme oxygenase isotypes, homologs of heme oxygenase were collected from the NCBI database (see details in Materials and Methods). These three distinct types of red algal heme oxygenase were grouped in three clades in the phylogenetic tree (Fig 4).
Table 1. Comparison of general features for 102 florideophycean plastid genomes.

| Subclass                      | Species                          | General Characteristics | RNAs | GenBank Accession | Reference |
|-------------------------------|----------------------------------|-------------------------|------|-------------------|-----------|
|                               |                                  | Total bp | GC % | Introns | CDS | tRNAs | rRNA |                             |                        |                        |                        |
| Hildenbrandio-physidae        | Apophlaea sinclairii             | 182,437  | 30.5% | 2       |     | 190    | 31   | 3   | NC_031172 [39]               |                        |                        |                        |
|                               | Hildenbrandia rivularis          | 189,725  | 32.4% | 2       |     | 186    | 31   | 3   | NC_031177 [39]               |                        |                        |                        |
|                               | Hildenbrandia rubra              | 180,141  | 31.4% | 2       |     | 191    | 31   | 3   | NC_031146 [39]               |                        |                        |                        |
|                               | Batrachospermum macrosporum      | 184,937  | 34.1% | 2       |     | 190    | 31   | 3   | NC_031165 [1]                |                        |                        |                        |
|                               | Batrachospermum virens           | 189,725  | 32.4% | 2       |     | 186    | 31   | 3   | NC_031166 [1]                |                        |                        |                        |
|                               | Derjuminaea rubra                | 180,141  | 31.4% | 2       |     | 191    | 31   | 3   | NC_031146 [39]               |                        |                        |                        |
|                               | Nematotaxis viridis              | 183,248  | 35.0% | 2       |     | 190    | 31   | 3   | NC_031160 [1]                |                        |                        |                        |
|                               | Karpovskia americana hsy120      | 184,025  | 29.3% | 2       |     | 201    | 32   | 3   | NC_031178 [This study]  |                        |                        |                        |
|                               | Karpovskia mahlacensis           | 181,361  | 29.8% | 1       |     | 187    | 31   | 3   | MG25486 [9]                  |                        |                        |                        |
|                               | Liagora brachyphylla             | 180,393  | 34.7% | 2       |     | 190    | 31   | 3   | NC_031165 [9]                |                        |                        |                        |
|                               | Sheathia arcuata                 | 187,354  | 29.9% | 2       |     | 190    | 31   | 3   | NC_031166 [1]                |                        |                        |                        |
|                               | Sheathia calosiphila             | 185,555  | 29.1% | 2       |     | 187    | 31   | 3   | NC_031166 [1]                |                        |                        |                        |
|                               | Thorea hispida hsy077            | 175,191  | 28.3% | 2       |     | 194    | 31   | 3   | NC_031171 [This study]  |                        |                        |                        |
|                               | Titanophyllum setchelli           | 183,356  | 32.7% | 2       |     | 206    | 31   | 3   | NC_031165 [9]                |                        |                        |                        |
|                               | Trichogloiosphalus pedicelletta  | 183,497  | 31.9% | 2       |     | 206    | 31   | 3   | NC_031166 [9]                |                        |                        |                        |
|                               | Yamadakia caespitifera           | 182,480  | 35.9% | 2       |     | 206    | 31   | 3   | NC_031165 [9]                |                        |                        |                        |
|                               | Corallinophytae                  | 182,933  | 33.9% | 2       |     | 205    | 34   | 6   | NC_031147 [9]                |                        |                        |                        |
|                               | Calliarthron tuberculatum         | 190,451  | 32.5% | 1       |     | 207    | 31   | 3   | NC_031145 [9]                |                        |                        |                        |
|                               | Caloglossa caespitifera          | 180,393  | 30.5% | 1       |     | 187    | 31   | 3   | NC_031166 [9]                |                        |                        |                        |
|                               | Caloglossa intermedia            | 176,064  | 25.5% | 0       |     | 206    | 33   | 3   | NC_035260 [51]               |                        |                        |                        |
|                               | Ceramium japonicum               | 171,923  | 27.6% | 0       |     | 202    | 34   | 3   | NC_035268 [51]               |                        |                        |                        |
|                               | Chondrus crispus                 | 175,347  | 29.5% | 0       |     | 216    | 34   | 3   | NC_035276 [51]               |                        |                        |                        |
|                               | Chlorella calcitrapida           | 182,437  | 30.5% | 2       |     | 190    | 31   | 3   | NC_031172 [39]               |                        |                        |                        |
|                               | Chlorella minutissima            | 170,750  | 28.4% | 0       |     | 214    | 33   | 3   | NC_035268 [51]               |                        |                        |                        |
|                               | Chlorella simpliciculosa         | 167,514  | 26.5% | 0       |     | 202    | 34   | 3   | NC_035268 [51]               |                        |                        |                        |
|                               | Chlorella tenella                | 170,809  | 28.6% | 0       |     | 206    | 34   | 3   | NC_035264 [51]               |                        |                        |                        |
|                               | Bryophyllum asparaginatum         | 184,279  | 32.1% | 2       |     | 190    | 31   | 3   | NC_035265 [51]               |                        |                        |                        |
|                               | Caloglossa beccarri              | 165,038  | 26.9% | 0       |     | 201    | 33   | 3   | NC_035269 [51]               |                        |                        |                        |
|                               | Caloglossa mossea                | 166,379  | 31.0% | 1       |     | 207    | 31   | 3   | NC_035268 [51]               |                        |                        |                        |
|                               | Caloglossa nanimontana           | 157,121  | 22.3% | 0       |     | 209    | 32   | 3   | NC_035265 [51]               |                        |                        |                        |
|                               | Ceramium cinnabarinum            | 171,923  | 27.6% | 1       |     | 190    | 28   | 3   | NC_031121 [18]               |                        |                        |                        |
|                               | Ceramium japonicum              | 171,634  | 27.8% | 1       |     | 202    | 31   | 3   | NC_031174 [39]               |                        |                        |                        |
|                               | Chondrus crispus                 | 170,809  | 28.6% | 0       |     | 206    | 34   | 3   | NC_035264 [51]               |                        |                        |                        |
|                               | Choreocalyx polyisophorae        | 90,243   | 20.5% | 0       |     | 72     | 0    | 3   | KP080906 [42]                |                        |                        |                        |
|                               | Chlorella pygmaea                | 174,482  | 28.0% | 0       |     | 214    | 34   | 3   | NC_035294 [51]               |                        |                        |                        |

(Continued)
Table 1. (Continued)

| Subclass                | Species                      | General Characteristics | GenBank Accession | Reference |
|-------------------------|------------------------------|-------------------------|------------------|----------|
|                         |                              | Total bp | GC % | Introns | CDS | tRNAs | rRNA | Reference |
| Coeloseira compressa   | 176,291                      | 29.0%    | 0    | 0       | 202 | 30    | 3     | NC_030338 | [45]      |
| Dasya binghamiae       | 177,213                      | 25.6%    | 0    | 0       | 199 | 29    | 3     | NC_031161 | [40]      |
| Dasya naccarioides     | 170,970                      | 27.4%    | 0    | 0       | 204 | 1     | 3     | NC_035287 | [51]      |
| Dasyclonium flaccidum  | 168,768                      | 28.3%    | 0    | 0       | 201 | 34    | 3     | NC_035297 | [51]      |
| Dictyomenia sonderi    | 174,848                      | 29.8%    | 0    | 0       | 207 | 34    | 3     | NC_035298 | [51]      |
| Digena simplex          | 173,119                      | 26.2%    | 0    | 0       | 201 | 34    | 3     | NC_035257 | [51]      |
| Dipetrocladia arubensis| 169,341                      | 28.8%    | 0    | 0       | 205 | 33    | 3     | NC_035288 | [51]      |
| Dipetrocladia australis| 174,748                      | 30.2%    | 1    | 202    | 30  | 3     | 3     | NC_030338 | [45]      |
| Gelidium vagum         | 170,813                      | 29.9%    | 0    | 0       | 204 | 31    | 3     | NC_035287 | [51]      |
| Gracilaria chilensis    | 185,637                      | 29.3%    | 1    | 204    | 30  | 3     | 3     | NC_029860 | [44]      |
| Gracilaria firma        | 187,001                      | 28.1%    | 1    | 219    | 32  | 3     | 3     | NC_033877 | [43]      |
| Gracilaria salicornia   | 179,757                      | 28.8%    | 1    | 206    | 31  | 3     | 3     | NC_023785 | [53]      |
| Gracilaria tenuistipitata| 183,883                     | 29.2%    | 0    | 205    | 30  | 3     | 3     | NC_006137 | [49]      |
| Gracilaria vancouverensis| 203                          | 27.4%    | 1    | 206    | 32  | 3     | 3     | NC_030149 | [53]      |
| Gracilaria vancouverensis| 183,013                     | 27.4%    | 0    | 204    | 30  | 3     | 3     | NC_035290 | [51]      |
| Herposiphonia versicolor| 178,949                      | 29.6%    | 0    | 206    | 34  | 3     | 3     | NC_035279 | [51]      |
| Kuetzingia canaliculata | 178,949                      | 28.1%    | 1    | 218    | 34  | 3     | 3     | NC_035293 | [51]      |
| Membranoptera platyphylla| 176,031                     | 26.3%    | 0    | 193    | 29  | 3     | 3     | NC_032041 | [46]      |
| Membranoptera tenuis    | 176,031                      | 26.2%    | 0    | 192    | 29  | 3     | 3     | NC_032399 | [46]      |
| Membranoptera simplex   | 176,070                      | 26.2%    | 1    | 199    | 29  | 3     | 3     | NC_031179 | [39]      |
| Ophidoclada kuetzingii  | 175,085                      | 26.9%    | 0    | 221    | 34  | 3     | 3     | NC_035292 | [51]      |
| Mastocarpus papillatus  | 184,382                      | 29.1%    | 0    | 206    | 30  | 3     | 3     | NC_031167 | [43]      |
| Melanothamnus harveyi  | 164,979                      | 29.8%    | 0    | 207    | 33  | 3     | 3     | NC_035281 | [51]      |
| Membranoptera platyphylla| 176,159                     | 26.4%    | 0    | 193    | 29  | 3     | 3     | NC_032041 | [46]      |
| Ophiocladus simplicissculus | 168,531                   | 28.1%    | 0    | 203    | 34  | 3     | 3     | NC_035258 | [51]      |
| Osmundaria timbriata   | 183,995                      | 28.3%    | 0    | 224    | 34  | 3     | 3     | NC_035262 | [51]      |
| Periphyton beckeri     | 168,290                      | 28.8%    | 0    | 206    | 33  | 3     | 3     | NC_035261 | [51]      |
| Polysiphonia schneideri | 163,271                      | 28.1%    | 1    | 203    | 33  | 3     | 3     | NC_035296 | [51]      |
| Polysiphonia scapulorum| 168,001                      | 29.2%    | 0    | 206    | 34  | 3     | 3     | NC_035282 | [51]      |
| Polysiphonia sertularioides | 166,000                    | 29.6%    | 0    | 203    | 33  | 3     | 3     | NC_035270 | [51]      |
| Polysiphonia stricta    | 169,061                      | 29.0%    | 0    | 201    | 34  | 3     | 3     | NC_035270 | [51]      |
| Polysiphonia sertularioides | 166,000                    | 29.6%    | 0    | 203    | 33  | 3     | 3     | NC_035270 | [51]      |
| Polysiphonia sertularioides | 165,361                    | 29.0%    | 0    | 201    | 34  | 3     | 3     | NC_035270 | [51]      |
| Rhodomelaceae confervoides | 175,951                     | 29.0%    | 0    | 210    | 34  | 3     | 3     | NC_035271 | [51]      |
| Rhodymenia pseudopalmata | 194,153                     | 32.0%    | 1    | 202    | 32  | 6     | 3     | NC_031144 | [39]      |
| Riquetophycus sp.       | 180,384                      | 28.8%    | 1    | 205    | 30  | 4     | 3     | NC_029859 | [44]      |
| Schizymenia dubyi       | 183,959                      | 30.0%    | 0    | 206    | 30  | 3     | 3     | NC_031169 | [39]      |
| Sebdenia flabellata     | 169,619                      | 26.0%    | 0    | 201    | 34  | 3     | 3     | NC_035289 | [51]      |
| Spondelina linearis     | 169,341                      | 28.8%    | 0    | 205    | 33  | 3     | 3     | NC_035288 | [51]      |
| Spondelina linearis     | 169,341                      | 28.8%    | 0    | 205    | 33  | 3     | 3     | NC_035288 | [51]      |

(Continued)
A gene. The plastid encoded \textit{pbs} A gene was present in most marine red algal species. Red algal \textit{pbs} A genes were grouped in a highly supported clade with diverse cyanobacteria (94% MLB) (Fig 4), suggesting a cyanobacterial origin. Additionally, \textit{pbs} A was present in the red algal derived plastids of cryptophytes. The \textit{pbs} A homolog was also found in the nuclear genome of the \textit{Cyanophora paradoxa} (Glaucophyta) with an additional extension of the N-terminal transit peptide (e-value = 1e-61; compare to \textit{pbs} A gene of \textit{Porphyridium purpureum}). The alignment for \textit{pbs} A and its homologous proteins were identified to have a functional domain of heme oxygenase (see Fig 3A). Nonetheless, each type of heme oxygenase proteins contained a few differences, such as putative transmembrane domain (red box) in C-terminal extension or targeting domain (green box) in N-terminal extension (Fig 3A). For \textit{pbs} A protein sequences, a transmembrane region was predicted as shown in Fig 3B, where most of the red algal \textit{pbs} A proteins had putative transmembrane domains in the C-terminal. This C-terminal transmembrane domain was also present in \textit{HMOX2} genes, but \textit{HMOX2} genes contained additional putative transmembrane domains inside the functional heme oxygenase domain. None of \textit{HMOX1} genes in red algae were predicted to have a transmembrane domain region in their protein sequences.

One noteworthy discovery was that the \textit{pbs} A gene is generally absent in freshwater species (75%; 12 of 16 spp.), but present in most of the marine red algal species (86%; 100 of 116) (see S3 Table). Heme oxygenase is well known for acting as an iron-controlling factor [61, 64]. It has been demonstrated that transcription of the \textit{pbs} A gene in a unicellular red alga, \textit{Rhodella violacea} (Rhodophyceae), was up-regulated in iron deprivation conditions [65]. Given these observations, it is highly likely that the \textit{pbs} A gene in red algae assists to uptake of iron in iron deprived marine environment. On the other hand, most freshwater red algae have lost the \textit{pbs} A gene, likely due to this gene being unnecessary or redundant in freshwaters that are typically not as iron limited as marine environments (iron composition in freshwater is ~1,400 times higher than seawater [66]). Although gene content in red algal plastid genomes was highly conserved [39], \textit{pbs} A gene may be a good example for the genomic response to environmental adaption.

\textit{HMOX1} gene. Although the nuclear heme oxygenase gene, \textit{HMOX1}, contained a conserved heme oxygenase domain and heme binding motif, it was clearly different from the other heme oxygenase genes in the phylogenetic and gene network analyses. The ML tree (Fig 4) showed that orthologs of \textit{HMOX1} gene form a strongly supported monophyletic group (93% MLB) with relatively long-branches. In addition, the gene network analysis (S2 Fig) also indicated \textit{HMOX1} genes to be clearly separated from the groups of \textit{HMOX2} or \textit{pbs} A despite

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
Subclass & Species & General Characteristics & RNAs & GenBank Accession & Reference \\
\hline
 & & Total bp & GC % & Introns & CDS & tRNAs & rRNA \\
\hline
Symphyocladia dendroidea & 171,837 & 28.4% & 0 & 210 & 29 & 0 & NC_035267 [51] \\
Taenioma perpusillum & 163,418 & 27.6% & 0 & 200 & 32 & 3 & NC_035295 [51] \\
Thaumatella adunca & 169,659 & 26.5% & 0 & 203 & 34 & 3 & NC_035291 [51] \\
Thuretia quercifolia & 274,510 & 25.9% & 0 & 212 & 34 & 3 & NC_035291 [51] \\
Tolyposcladus gloriosulata & 165,623 & 29.2% & 0 & 206 & 33 & 3 & NC_035299 [51] \\
Vertebrata australis & 167,318 & 28.3% & 0 & 199 & 33 & 3 & NC_035283 [51] \\
Vertebrata isogona & 167,445 & 28.3% & 0 & 205 & 33 & 3 & NC_035278 [51] \\
Vertebrata lanosa & 167,158 & 30.0% & 0 & 193 & 28 & 3 & KP308097 [42] \\
Vertebrata thyesides & 168,951 & 28.6% & 0 & 208 & 33 & 3 & NC_035273 [51] \\
\hline
\end{tabular}
\caption{Continued}

\end{table}

https://doi.org/10.1371/journal.pone.0196995.t001

\textit{pbs} A gene. The plastid encoded \textit{pbs} A gene was present in most marine red algal species. Red algal \textit{pbs} A genes were grouped in a highly supported clade with diverse cyanobacteria (94% MLB) (Fig 4), suggesting a cyanobacterial origin. Additionally, \textit{pbs} A was present in the red algal derived plastids of cryptophytes. The \textit{pbs} A homolog was also found in the nuclear genome of the \textit{Cyanophora paradoxa} (Glaucophyta) with an additional extension of the N-terminal transit peptide (e-value = 1e-61; compare to \textit{pbs} A gene of \textit{Porphyridium purpureum}). The alignment for \textit{pbs} A and its homologous proteins were identified to have a functional domain of heme oxygenase (see Fig 3A). Nonetheless, each type of heme oxygenase proteins contained a few differences, such as putative transmembrane domain (red box) in C-terminal extension or targeting domain (green box) in N-terminal extension (Fig 3A). For \textit{pbs} A protein sequences, a transmembrane region was predicted as shown in Fig 3B, where most of the red algal \textit{pbs} A proteins had putative transmembrane domains in the C-terminal. This C-terminal transmembrane domain was also present in \textit{HMOX2} genes, but \textit{HMOX2} genes contained additional putative transmembrane domains inside the functional heme oxygenase domain. None of \textit{HMOX1} genes in red algae were predicted to have a transmembrane domain region in their protein sequences.

One noteworthy discovery was that the \textit{pbs} A gene is generally absent in freshwater species (75%; 12 of 16 spp.), but present in most of the marine red algal species (86%; 100 of 116) (see S3 Table). Heme oxygenase is well known for acting as an iron-controlling factor [61, 64]. It has been demonstrated that transcription of the \textit{pbs} A gene in a unicellular red alga, \textit{Rhodella violacea} (Rhodophyceae), was up-regulated in iron deprivation conditions [65]. Given these observations, it is highly likely that the \textit{pbs} A gene in red algae assists to uptake of iron in iron deprived marine environment. On the other hand, most freshwater red algae have lost the \textit{pbs} A gene, likely due to this gene being unnecessary or redundant in freshwaters that are typically not as iron limited as marine environments (iron composition in freshwater is ~1,400 times higher than seawater [66]). Although gene content in red algal plastid genomes was highly conserved [39], \textit{pbs} A gene may be a good example for the genomic response to environmental adaption.

\textit{HMOX1} gene. Although the nuclear heme oxygenase gene, \textit{HMOX1}, contained a conserved heme oxygenase domain and heme binding motif, it was clearly different from the other heme oxygenase genes in the phylogenetic and gene network analyses. The ML tree (Fig 4) showed that orthologs of \textit{HMOX1} gene form a strongly supported monophyletic group (93% MLB) with relatively long-branches. In addition, the gene network analysis (S2 Fig) also indicated \textit{HMOX1} genes to be clearly separated from the groups of \textit{HMOX2} or \textit{pbs} A despite
Fig 2. The distribution of four putative habitat-specific genes in the phylogenetic tree. The absence/presence of four genes (pbsA, ycf34, ycf35, ycf46) in 127 species is visualized with habitat information. The maximum likelihood phylogenetic tree was reconstructed based on the concatenated 190 orthologous plastid gene alignment. The dataset used in this analysis is shown in S3 Table.

https://doi.org/10.1371/journal.pone.0196995.g002
Fig 3. The sequence alignment of heme oxygenase proteins. (A) The alignment of pbsA and its homologous proteins. The alignment shows the conserved heme oxygenases amino acid sequences in different lineages. Conserved heme binding pockets are marked as a blue asterisk. N-terminal transit peptides (green) are unique for HMOX1 proteins, with an exceptional transit peptide of pbsA gene in Cyanophora paradoxa, which was likely transferred to the nuclear genome independently. Heme oxygenase domain (grey) and putative transmembrane domain (red) are shown. (B) HMOX2 and pbsA contain putative transmembrane domain(s) (TM domain; red box). Multiple TM domains were found in HMOX2.

https://doi.org/10.1371/journal.pone.0196995.g003
Fig 4. Maximum likelihood tree and schematic diagrams of heme oxygenase proteins. Maximum likelihood (ML) tree based on an alignment of 510 heme oxygenase amino acid sequences from 1,678 taxa. Red algal heme oxygenases had three isotypes of heme oxygenase: \( \text{HMOX1} \) \( \text{HMOX2} \) \( \text{pbsA} \). \( \text{HMOX1} \) and \( \text{HMOX2} \) were located in the nuclear genome whereas the \( \text{pbsA} \) was encoded in the plastid genome in red algae.

https://doi.org/10.1371/journal.pone.0196995.g004
the conservation of heme oxygenase domain. Unlike HMOX2 and pbsA, protein sequences of the HMOX1 gene displayed a low similarity (20–30%) to the cyanobacterial heme oxygenase. Interestingly, the HMOX1 gene had unique N-terminal peptides, which were predicted (via ChloroP software) to target the plastid (Fig 3). N-terminal peptides were present in the photosynthetic eukaryote lineages including the primary endosymbiotic lineages (red algae, Viridiplantae, and a glaucophyte alga Cyanophora paradoxa) and red algal-derived secondary plastid groups (i.e., haptophytes, cryptophytes, stramenopiles). While the transit peptides were absent in HMOX2 and pbsA, pbsA in C. paradoxa (Glaucophyta) was located in the nucleus and possesses the additional N-terminal transit peptide like those of HMOX1 genes. In Chlorophyta species, they retained both HMOX1 and HMOX2 in the nucleus, but only the HMOX1 was found in the streptophytes (charophytes and land plants) [67, 68]. Streptophyta species had several copies of the heme oxygenase gene (HMOX1) in their nuclear genome and those genes formed a monophyletic group with other HMOX1 in the phylogenetic tree (Fig 4). For instance, Arabidopsis thaliana contained the biochemically well-characterized heme oxygenase HY1 with three additional putative heme oxygenase copies (HO2, HO3 and HO4) [69]. Analysis of the major biochemical parameters (i.e., enzyme activity depends on pH, temperature, conversion time of heme to biliverdin) demonstrated that activities of these three heme oxygenases (HO2, HO3 and HO4) do not different from that of HY1 [70]. Therefore, unlike other lineages, the four isotypes of HMOX1 (HY1, HO2, HO3 and HO4) in land plants likely plays an important role in synthesizing phycobilin, and these four isotypes likely originated from gene duplication events from an ancestral land plant HMOX1 gene [71, 72].

Because of their low similarities (protein identities: 22.41~27.68%; 1e-05~1e-10) to other homologous proteins, including the cyanobacterial heme oxygenase, the origin of HMOX1 were still ambiguous. However, HMOX1 was only present in plastid-bearing eukaryotes and this was the only gene that possesses the plastid-targeting transit peptide among the heme oxygenase families. Therefore, we concluded that HMOX1 was involved in plastidial function. The phylogenetic position of HMOX1 in red algal-derived secondary endosymbionts suggested that HMOX1 was likely derived from secondary endosymbiosis events followed by the gene transfer to the nuclear genome of the host (Fig 4). Indeed, a trafficking experiment in Chlamydomonas reinhardtii showed that Chlamydomonas nuclear-encoded HMOX1 gene targets to the plastid [63]. Although none of these studies showed the trafficking of red algal HMOX1, it is highly likely that red algal HMOX1 targets the plastid because of the monophyly of red algae with the Viridiplantae. In Viridiplantae, Chlamydomonas reinhardtii, Arabidopsis thaliana, and Ceratodon purpureus provided experimental evidence for plastid trafficking with the N-terminal extension HMOX1 [70, 73, 74]. Nevertheless, function of HMOX1 genes in red algae needs further experimental investigation to elucidate its metabolic pathway.

HMOX2 gene. The red algal HMOX2, another nuclear encoded heme oxygenase gene, had one or more putative transmembrane domains in C-terminus (Fig 3). Transmembrane domain structures of red algae were homologous to those of metazoan (e.g., mammalian, amphibian) heme oxygenase [63, 75]. Within the HMOX2 clade in the phylogenetic trees (Fig 4), red algae were grouped with diverse eukaryotes including chlorophytes, cryptophytes, and stramenopiles as well as non-photosynthetic fungi and Metazoa. Therefore, we would suggest that HMOX2 was derived from an ancient eukaryotic common ancestor. Gene network analysis and the sequence conservation between the HMOX2 and pbsA in protein alignment supported that the HMOX2 is related to the iron uptake function. Indeed, it has been reported that C. reinhardtii captured extracellular heme as an iron source with an association between the HMOX2 protein and the cytosolic membrane [63].
Conclusion

Three Nemaliophycidae plastid genomes were completely sequenced and annotated. These included two exclusively freshwater species *Kumanoa americana* and *Thorea hispida* and the marine species *Palmaria palmata*. Until recently, plastid genome data were used mainly for phylogenomic analysis (e.g., [47]), divergence time estimation (e.g., [10]), comparative structural analysis (e.g., [39]), and the development of red algal molecular markers for DNA bar-coding studies (e.g., [12]). In this study, we focused on finding genomic clues for an environmental adaptation between marine and freshwater red algal species.

Based on the environment-specific genes of the heme oxygenase family in red algae, we postulate that red algae have adapted efficiently in differing iron concentration conditions through the retention or loss of the heme oxygenase genes. Although this study included only a few freshwater red algal species and was not able to present the direct evidence of correlation between habitat and gene retention, we demonstrated the general trend of red algal plastid gene loss (ycf34, ycf35, ycf46, and pbsA) in iron-limited marine habitats.

We also demonstrated different evolutionary strategies of three types of heme oxygenase genes in different lineages (e.g., presence of *HMOX1* with gene duplications [HY1, HO2, HO3 and HO4], but the absence of *HMOX2* in the streptophytes, which include charophytes and land plants) and habitat conditions (e.g., *pbsA* genes in the marine and freshwater red algal species). It is generally known that plastid genes are in the process of reduction (either complete loss or gene transfer to the host nucleus) after its endosymbiotic origin [76]. However, gene loss from the plastid genome appears functionally constrained as demonstrated in *pbsA* and its gene homologs. Through selective gene retention, red algae successfully adapted to different aquatic environments over billions of years of evolutionary history.

Supporting information

S1 Fig. Venn diagram visualization of comparing gene contents within the three Nemaliophycidae genomes. Six unique genes (*pbsA, grx, ycf35, ycf36, ycf37, ycf46*) were only found in marine *Palmaria palmata*. For freshwater species, there is only one gene (*ycf91*) that found in *Thorea hispida* and *Kumanoa americana*, but not in *Palmaria palmata*. (EPS)

S2 Fig. Gene network for heme oxygenase. The gene network was constructed by EGN with heme oxygenase database from the public database. We performed analysis with 1e-05 of e-value, 20% of protein identities and 70% of query coverage. The result shows that *HMOX1* are clearly separated from other red algal heme oxygenase. (EPS)

S1 Table. The sequencing information for the plastid genome of three Nemaliophycidae species. (XLSX)

S2 Table. A list of genes in the plastid genomes of *Palmaria palmata, Kumanoa americana, and Thorea hispida*. (XLSX)

S3 Table. The habitat-specific gene survey in 127 plastid genomes of red algae. 1) Marine unique gene (only in *Palmaria palmata: pbsA, grx, ycf35, ycf36, ycf37, ycf46*); 2) Freshwater unique gene (not in *Palmaria palmata* but in *Thorea hispida* and *Kumanoa americana*: *ycf91*. (XLSX)
Acknowledgments

We would like to thank JunMo Lee, Louis Graf, and Eun Chan Yang for technical assistance in genome sequencing and data analysis as well as Morgan Vis for supplying algal culture and helpful comments on the manuscript. Thanks for the constructive comments from reviewers to revise the manuscript.

Author Contributions

Conceptualization: Chung Hyun Cho, Daryl W. Lam, Hwan Su Yoon.
Data curation: Chung Hyun Cho, Ji Won Choi, Kyeong Mi Kim.
Formal analysis: Chung Hyun Cho, Ji Won Choi, Kyeong Mi Kim.
Funding acquisition: Hwan Su Yoon.
Investigation: Hwan Su Yoon.
Supervision: Hwan Su Yoon.
Writing – original draft: Chung Hyun Cho, Hwan Su Yoon.
Writing – review & editing: Chung Hyun Cho, Ji Won Choi, Daryl W. Lam, Kyeong Mi Kim, Hwan Su Yoon.

References

1. Lam DW, Verbruggen H, Saunders GW, Vis ML. Multigene phylogeny of the red algal subclass Nematophycidae. Molecular Phylogenetics and Evolution. 2016; 94, Part B:730–6. https://doi.org/10.1016/j.ympev.2015.10.015 PMID: 26518739
2. Guiry MD, Guiry GM. AlgaeBase. World-wide electronic publication 2017 [cited 2017 October. 18]. Available from: http://www.algaebase.org.
3. Dixon PS. Biology of the Rhodophyta. Reprint ed ed. Koenigstein: Koeltz: Oliver & Boyd Edinburgh; 1973.
4. Le Gall L, Saunders GW. A nuclear phylogeny of the Florideophyceae (Rhodophyta) inferred from combined EF2, small subunit and large subunit ribosomal DNA: establishing the new red algal subclass Corallinophycidae. Molecular Phylogenetics and Evolution. 2007; 43(3):1118–30. https://doi.org/10.1016/j.ympev.2006.11.012 PMID: 17197199
5. Ott FD. Handbook of the taxonomic names associated with the non-marine Rhodophycophyta: J. Cramer in der Gebrüder Borntraeger Verlagsbuchhandlung; 2009.
6. Kapraun DF, Braly KS, Freshwater DW. Nuclear DNA content variation in the freshwater red algal orders Batrachospermales and Thoreales (Florideophyceae, Nematophycidae). Phycologia. 2007; 46 (1):54–62.
7. Peredo EL, King UM, Les DH. The plastid genome of Najas flexilis: Adaptation to submersed environments is accompanied by the complete loss of the NDH complex in an aquatic angiosperm. PLoS One. 2013; 8(7):e68591. https://doi.org/10.1371/journal.pone.0068591 PMID: 23861923
8. Costa JF, Lin S-M, Macaya EC, Fernández-Garcia C, Verbruggen H. Chloroplast genomes as a tool to resolve red algal phylogenies: a case study in the Nemaliales. BMC Evolutionary Biology. 2016; 16 (1):205. https://doi.org/10.1186/s12862-016-0772-3 PMID: 27724867
9. Paiano MO, Del Cortona A, Costa JF, Liu SL, Verbruggen H, De Cierck O, et al. Organization of plastid genomes in the freshwater red algal order Batrachospermales (Rhodophyta). Journal of Phycology. 2018; 54(1):25–33. https://doi.org/10.1111/jpy.12802 PMID: 29077982
10. Nan F, Feng J, Lv J, Liu Q, Fang K, Gong C, et al. Origin and evolutionary history of freshwater Rhodophyta: further insights based on phylogenomic evidence. Scientific Reports. 2017; 7:2934. https://doi.org/10.1038/s41598-017-03235-5 PMID: 28592899

11. Chevreux B, Wetter T, Suhai S, editors. Genome sequence assembly using trace signals and additional sequence information. German Conference on Bioinformatics; 1999.

12. Janouskovec J, Liu SL, Martone PT, Carre W, Leblanc C, Collen J, et al. Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. PLoS One. 2013; 8(3):e59001. https://doi.org/10.1371/journal.pone.0059001 PMID: 23536846

13. Collen J, Porcel B, Carre W, Ball SG, Chaparro C, Tonon T, et al. Genome structure and metabolic features in the red seaweed Chondrus crispus shed light on evolution of the Archaeplastida. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110(13):5247–52. https://doi.org/10.1073/pnas.1212591110 PMID: 23503846

14. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biology. 2004; 5(2):R12. https://doi.org/10.1186/gb-2004-5-2-r12 PMID: 14759262

15. Song HJ, Lee J, Graf L, Rho M, Qiu H, Bhattacharya D, et al. A novice’s guide to analyzing NGS-derived organelle and metagenome data. Algae. 2016; 31(2):137–54.

16. Lagesen K, Hallin P, Rødland EA, Stærlund H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Research. 2007; 35(9):3100–8. https://doi.org/10.1093/nar/gkm160 PMID: 17452365

17. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Research. 2004; 32(1):11–6. https://doi.org/10.1093/nar/gkh152 PMID: 14704338

18. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, et al. Rfam 12.0: updates to the RNA families database. Nucleic Acids Research. 2015; 43(D1):D130–D7.

19. Lohse M, Drechsel O, Bock R. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Current Genetics. 2007; 52(5–6):267–74. https://doi.org/10.1007/s00294-007-0161-y PMID: 17957369

20. Hilker R, Sickinger C, Pedersen CN, Stoye J. UniMoG—a unifying framework for genomic distance calculation and sorting based on DCJ. Bioinformatics. 2012; 28(19):2509–11. https://doi.org/10.1093/bioinformatics/bts199 PMID: 22815356

21. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research. 2004; 14(7):1394–403. https://doi.org/10.1101/gr.2289704 PMID: 15231754

22. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010; 5(6):e11147. https://doi.org/10.1371/journal.pone.0011147 PMID: 20593022

23. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS Biology. 2012; 10(6):e1001889. https://doi.org/10.1371/journal.pbio.1001889 PMID: 24959919

24. Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber AP, et al. Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. Science. 2012; 335(6070):843–7. https://doi.org/10.1126/science.1213561 PMID: 22344442

25. Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, et al. Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Current Biology. 2011; 21(4):328–33. https://doi.org/10.1016/j.cub.2011.01.037 PMID: 21315598

26. Bhattacharya D, Price DC, Chan CX, Qiu H, Rose N, Ball S, et al. Genome of the red alga Porphyridium purpureum. Nature Communications. 2013; 4:1941. https://doi.org/10.1038/ncomms2931 PMID: 23770768

27. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 2013; 30(4):772–80. https://doi.org/10.1093/molbev/msb010 PMID: 23329690
30. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution. 2015; 32(1):268–74. https://doi.org/10.1093/molbev/msu300 PMID: 25371430

31. Le SQ, Gascuel O. An improved general amino acid replacement matrix. Molecular Biology and Evolution. 2008; 25(7):1307–20. https://doi.org/10.1093/molbev/msn067 PMID: 18367465

32. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Research. 2011; 39 (suppl 1):D225–D9.

33. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. Journal of Molecular Biology. 2000; 300(4):1005–16. https://doi.org/10.1006/jmbi.2000.4315 PMID: 11152613

34. Emanuelsson O, Nielsen H, Von Hejine G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Science. 1999; 8(05):978–84.

35. Krogh A, Larsson B, Von Hejine G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. Journal of Molecular Biology. 2001; 305 (3):567–80. https://doi.org/10.1006/jmbi.2000.4315 PMID: 11152613

36. Kanelhisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research. 2000; 28(1):27–30. PMID: 10592173

37. Halary S, McInerney J, Lopez P, Bapteste E. EGN: a wizard for construction of gene and genome similarity networks. BMC Evolutionary Biology. 2013; 13(1):146.

38. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research. 2003; 13 (11):2498–504. https://doi.org/10.1101/gr.1239303 PMID: 14597658

39. Lee J, Cho CH, Park SI, Choi JW, Song HS, West JA, et al. Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. BMC Biology. 2016; 14(1):1–16.

40. Tamayo DA, Hughey JR. Organellar genome analysis of the marine red alga Dasya binghamiae (Dasyraceae, Rhodophyta) reveals an uncharacteristic florideophyte mitogenome structure. Mitochondrial DNA Part B. 2016; 1(1):510–1.

41. Sissini MN, Navarrete-Fernández TM, Murray FM, Freese JM, Gentilhomme AS, Huber SR, et al. Mitochondrial and plastid genome analysis of the heteromorph red alga Mastocarpus papillatus (C. Agardh) Kützing (Phyllophoraceae, Rhodophyta) reveals two characteristic florideophyte organellar genomes. Mitochondrial DNA Part B. 2016; 1(1):676–7.

42. Salomaki ED, Nickles KR, Lane CE. The ghost plastid of Choreocolax polysiphoniae. Journal of Phycology. 2015; 51(2):217–21. https://doi.org/10.1111/jpy.12283 PMID: 26986516

43. Nq P-K, Lin S-M, Lim P-E, Liu L-C, Chen C-M, Pai T-W. Complete chloroplast genome of Gracilaria firma (Graciliaceae, Rhodophyta), with discussion on the use of chloroplast phylogenomics in the sub-class Rhodymenioiophyceae. BMC Genomics. 2017; 18(1):40. https://doi.org/10.1186/s12864-016-3453-0 PMID: 28061748

44. Lee J, Kim KM, Yang EC, Miller KA, Boo SM, Bhattacharya D, et al. Reconstructing the complex evolutionary history of mobile plasmids in red algal genomes. Scientific Reports. 2016; 6:23744. https://doi.org/10.1038/srep23744 PMID: 27030297

45. Kilpatrick ZM, Hughey JR. Mitochondrial and plastid genome analysis of the marine red alga Coeloseira compressa (Champiaceae, Rhodophyta). Mitochondrial DNA Part B. 2016; 1(1):456–8.

46. Hughey JR, Hommersand MH, Gabrielson PW, Miller KA, Fuller T. Analysis of the complete plastomes of three species of Membranoptera (Ceramiaceae, Rhodophyta) from Pacific North America. Journal of Phycology. 2017; 53(1):32–43. https://doi.org/10.1111/jphy.12472 PMID: 27690326

47. Hughey JR, Gabrielson PW, Rohmer L, Tortolani J, Silva M, Miller KA, et al. Minimally destructive sampling of type specimens of Pyropia (Bangiales, Rhodophyta) recovers complete plastid and mitochondrial genomes. Scientific Reports. 2014; 4:5113. https://doi.org/10.1038/srep05113 PMID: 24894641

48. Hughey JR, Boo GH. Genomic and phylogenetic analysis of Ceramium cimbrium (Ceramiaceae, Rhodophyta) from the Atlantic and Pacific Oceans supports the naming of a new invasive Pacific entity Ceramium sungninthoii sp. nov. Botanica Marina. 2016; 59(4):211–22.

49. Hagopian JC, Reis M, Kitajima JP, Bhattacharya D, de Oliveira MC. Comparative analysis of the complete plastid genome sequence of the red alga Gracilaria tenuistipitata var. liui provides insights into the evolution of rhodoplasts and their relationship to other plastids. Journal of Molecular Evolution. 2004; 59 (4):464–77. https://doi.org/10.1007/s00239-004-2638-3 PMID: 15638458

50. Du Q, Bi G, Mao Y, Sui Z. The complete chloroplast genome of Gracilariosis lemaneiformis (Rhodophyta) gives new insight into the evolution of family Graciliaceae. Journal of Phycology. 2016; 52 (3):441–50. https://doi.org/10.1111/jphy.12406 PMID: 27273536
51. Diaz Tapia P, Maggs CA, West JA, Verbruggen H. Analysis of chloroplast genomes and a supermatrix inform reclassification of the Rhodomelaceae (Rhodophyta). Journal of Phycology. 2017; 53:920–37. https://doi.org/10.1111/jpy.12553 PMID: 28561261

52. DePriest MS, Bhattacharya D, Lopez-Bautista JM. The plastid genome of the red macroalga Grateloupia taiwanensis (Halyomniaceae). PLoS One. 2013; 8(7):e68246. https://doi.org/10.1371/journal.pone.0068246 PMID: 23894297

53. Campbell MA, Presting G, Bennett MS, Sherwood AR. Highly conserved organelar genomes in the Gracilariales as inferred using new data from the Hawaiian invasive alga Gracilaria salicornia (Rhodophyta). Phycologia. 2014; 53(2):109–16.

54. Kushibiki A, Yokoyama A, Iwataki M, Yokoyama J, West JA, Hara Y. New unicellular red alga, Bulboplasma apyrenoidosa gen. et sp. nov. (Rhodophyta) from the mangroves of Japan: phylogenetic and ultrastructural observations. Photobiology. 2012; 60(2):114–22.

55. Kamiya M, West JA, Karsten U, Ganesan E. Molecular and morphological delineation of Caloglossa beccarii and related species (Delesseriaceae, Rhodophyta). Phycologia. 2016; 55(6):640–9.

56. Zuccarello GC, West JA. Insights into evolution and speciation in the red alga Bostrychia: 15 years of research. Algae. 2011; 26(1):21–32.

57. Womersley H. Southern Australian species of Polysiphonia Greville (Rhodophyta). Australian Journal of Botany. 1979; 27(4):459–528.

58. Kugrens P, West JA. The ultrastructure of an alloparasitic red alga Choreococcol polysiphoniae. Phycologia. 1973; 12(3):175–86.

59. Kikuchi G, Yoshida T, Noguchi M. Heme oxygenase and heme degradation. Biochemical and Biophysical Research Communications. 2005; 338(1):558–67. https://doi.org/10.1016/j.bbrc.2005.08.020 PMID: 16115609

60. Rhie G-e, Beale SI. Biosynthesis of phycobilins. Ferredoxin-supported nadph-independent heme oxygenase and phycocyanin-forming activities from Cyanidium caldarium. Journal of Biological Chemistry. 1992; 267(23):16088–93. PMID: 1644795

61. Frankenberg-Dinkel N. Bacterial heme oxygenases. Antioxidants & Redox Signaling. 2004; 6(5):825–34.

62. Poss KD, Tonegawa S. Heme oxygenase 1 is required for mammalian iron reutilization. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94(20):10919–24. PMID: 9380735

63. Duanmu D, Casero D, Dent RM, Gallaher S, Yang W, Rockwell NC, et al. Retrograde bilin signaling enables Chlamydomonas greening and phototrophic survival. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110(9):3621–6. https://doi.org/10.1073/pnas.1222375110 PMID: 23345435

64. Hentze MW, Muckenthaler MU, Andrews NC. Balancing acts: molecular control of mammalian iron metabolism. Cell. 2004; 117(3):285–97. PMID: 15109490

65. Richaud C, Zabulon G. The heme oxygenase gene (pbsA) in the red alga Rhodella violacea is discontinuous and transcriptionally activated during iron limitation. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94(21):11736–41. PMID: 9326680

66. Heyck R, Kilham P. Nutrient limitation of phytoplankton in freshwater and marine environments: a review of recent evidence on the effects of enrichment. Limnology and Oceanography. 1988; 33(4part2):796–822.

67. Emborg TJ, Walker JM, Noh B, Vierstra RD. Multiple heme oxygenase family members contribute to the biosynthesis of the phytochrome chromophore in Arabidopsis. Plant Physiology. 2006; 140(3):856–68. https://doi.org/10.1104/pp.105.074211 PMID: 16428602

68. Flagel LE, Wendel JF. Gene duplication and evolutionary novelty in plants. New Phytologist. 2009; 183(3):557–64. https://doi.org/10.1111/j.1469-8137.2009.02923.x PMID: 19555435

69. Garcia-Mata C, Lamattina L. Gasotransmitters are emerging as new guard cell signaling molecules and regulators of leaf gas exchange. Plant Science. 2013; 201:66–73. https://doi.org/10.1016/j.plantsci.2012.11.007 PMID: 23352403

70. Gisk B, Yasui Y, Kohchi T, Frankenberg-Dinkel N. Characterization of the haem oxygenase protein family in Arabidopsis thaliana reveals a diversity of functions. Biochemical Journal. 2010; 425(2):425–34. https://doi.org/10.1042/BJ20090775 PMID: 19860740

71. Panchy N, Lehti-Shiu M, Shiu S-H. Evolution of gene duplication in plants. Plant Physiology. 2016; 171(4):2294–316. https://doi.org/10.1104/pp.16.00523 PMID: 27288366

72. Shekhawat GS, Verma K. Haem oxygenase (HO): an overlooked enzyme of plant metabolism and defence. Journal of Experimental Botany. 2010; 61(9):2255–70. https://doi.org/10.1093/jxb/erq074 PMID: 20378668
73. Muramoto T, Kohchi T, Yokota A, Hwang I, Goodman HM. The Arabidopsis photomorphogenic mutant hy1 is deficient in phytochrome chromophore biosynthesis as a result of a mutation in a plastid heme oxygenase. The Plant Cell. 1999; 11(3):335–47. PMID: 10072395

74. Brücker G, Mittmann F, Hartmann E, Lamparter T. Targeted site-directed mutagenesis of a heme oxygenase locus by gene replacement in the moss Ceratodon purpureus. Planta. 2005; 220(6):864–74. https://doi.org/10.1007/s00425-004-1411-6 PMID: 15578218

75. Kadish KM, Smith KM, Guilard R. The Porphyrin Handbook: The iron and cobalt pigments: biosynthesis, structure, and degradation: Elsevier; 2003.

76. Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature Reviews Genetics. 2004; 5(2):123–35. https://doi.org/10.1038/nrg1271 PMID: 14735123