Marine Life Science & Technology (2025) 7:717–729
https://doi.org/10.1007/s42995-025-00327-4

RESEARCH PAPER

Species-specific marker development for accurate identification of three red algae (*Grateloupia asiatica*, *Pachymeniopsis lanceolata* and *Polyopes affinis*) based on complete organelle genomes

Yong Jin Lee[1,2] · Yeon Mi Kim[2] · Cheol Min Kim[2] · Sung-Je Choi[3] · Cheol Seong Jang[1,2]

Received: 14 May 2024 / Accepted: 13 October 2025 / Published online: 28 October 2025
© The Author(s) 2025

Abstract
Red alga species in the phylum Rhodophyta are ecologically and economically important, which are widely used as food and medicinal products because they are rich in bioactive compounds. Unfortunately, researchers often meet challenges in identifying red algal species and understanding the evolutionary relationship among them. Shooting away these limitations necessitates detailed genomic studies, including those of cellular organelles. In this study, we sequenced circular organelle genomes of three red macroalgal species, *Grateloupia asiatica*, *Pachymeniopsis lanceolata* and *Polyopes affinis*, yielding two complete chloroplast genomes of *P. lanceolata* and *P. affinis* and one mitochondrial genome of *P. affinis*. The average chloroplast and mitochondrial genome sizes were 192,724 bp and 29,699 bp, respectively, which encoded 202 and 25 proteins, respectively, on average. The short- and long-repeat sequences, gene rearrangements, the nucleotide diversity, and phylogenetic relationship among these three species were analyzed.Memo: revise conjunctions and prepositions to form a clearer list (A, B, C, and D). The 194 chloroplast and 23 mitochondrial protein-coding genes shared by the species belonging to Florideophyceae and Bangiophyceae were used to reconstruct a maximum-likelihood phylogenetic tree. In addition, we developed 15 species-specific PCR markers, five for each species, using single-nucleotide polymorphism information. Our results should aid in identifying these species and deciphering the evolutionary relationship among species in the phylum Rhodophyta.

Keywords Halymeniaceae · Chloroplast genome · Mitochondrial genome

Introduction

Florideophyceae, a major class of red algae in phylum Rhodophyta, contributes about 95% of the total diversity of the phylum Rhodophyta, which encompasses > 7500 documented species in AlgaeBase (https://www.algaebase.org/) (Guiry 2024). These red algal species play pivotal roles in diverse marine ecosystems and have been extensively used in food, pharmaceutical and biotechnological industries; they are rich in high-value bioactive compounds (Cui et al. 2019; Kandale et al. 2011; Lee et al. 2023; Villanueva et al. 2014). The global demand for red alga-derived products is increasing constantly, and accordingly the sustainable exploitation of these red algal species requires accurate species identification.

Researchers have met the challenges in identifying these red algal species. The traditional methods often fail to distinguish them due to intra- and inter-specific variations, simple and convergent morphologies and phenotypic plasticity with environments (Kim et al. 2023; Saunders 2005; Verlaque et al. 2005). Occasional modifications in the scientific names of red algal species like *Porphyra dentata* (now *Neoporphyra dentata*) and *Grateloupia lanceolata* (now *Pachymeniopsis lanceolata*) have also complicated the identification of red algal species, which

Edited by Jiamei Li.

✉ Cheol Seong Jang
csjang@kangwon.ac.kr

[1] Agriculture and Life Sciences Research Institute, Kangwon National University, Chuncheon 24341, Korea

[2] Plant Genomics Laboratory, Interdisciplinary Program in Smart Agriculture, Kangwon National University, Chuncheon 24341, Korea

[3] Department of Aquaculture and Fisheries Convergence, Korea National College of Agriculture and Fisheries, Jeonju 54874, Korea

is even worsened by the frequent application of outdated names (Lee et al. 2023; Li et al. 2019; Yang et al. 2022). In addition, morphological characteristics alone are often insufficient for the accurate identification of Rhodophyta species; they may lead to unintended mixing or misuse of different species during manufacturing processes (Yang and Kim 2015). To address these challenges, the species boundaries and genetic relationships among Rhodophyta species should be untangled using molecular tools and genomic resources.

The genomes of chloroplasts and mitochondria promise to be the resources for winning these challenges. These genomes are relatively small and high in copy numbers, making them extremely applicable for developing species-specific molecular markers (Greiner et al. 2015; Liu et al. 2024; Rawal et al. 2020; Wang et al. 2021). However, the organelle genomic data for Florideophyceae remain limited, hindering our endeavors to develop reliable markers for species identification.

Recent advancements in high-throughput sequencing technologies have made the analysis of organellar genomes convenient. These technologies allow researchers to characterize the genomes in their gene content, structural variation, repeat element identification among others, supporting the development of species-specific PCR primers.

In this study, we enriched the genomic resources of economically important red algae by newly sequencing the organelle genomes of three species, *Grateloupia asiatica*, *Pachymeniopsis lanceolata* and *Polyopes affinis*. These species are used in South Korea as edible seaweeds. They are known also for their health benefits such as anti-inflammation and atopic dermatitis treatment (Ha et al 2022; Hwang and Park 2020; Khan et al 2008). These species were selected as the studying materials due to their economic significance and health application and less described genomic characteristics. The primary goal of this study was to develop reliable species-specific markers for accurate species identification. We sequenced and assembled the organellar genomes of these three species, and then conducted comparative analyses and deciphered phylogenetic relationships among species in Florideophyceae. We analyzed their mitochondrial (mt) genomes to reinforce the accuracy of their classification. We subsequently developed species-specific markers based on the SNPs found among the chloroplast (cp) genomes. We generated the first complete cp genomes for *P. lanceolata* and *P. affinis* and the first mt genome for *P. affinis*. Our results bridged the gaps between the genomic data and the species identification, and provided the foundation to the species identification. Furthermore, our findings enabled further explorations of the organelle genomes of other red algal species.

Materials and methods

Plant materials and DNA isolation

Algal samples of *G. asiatica*, *P. lanceolata* and *P. affinis* were collected from the coastal area of Jeju Island, Republic of Korea. The collection of algal samples was permitted by the Ministry of Ocean and Fisheries, Republic of Korea. This study was conducted in compliance with local, national, Kangwon National University (Chuncheon, Republic of Korea), and Ministry of Oceans and Fisheries (Sejong, Republic of Korea) regulations. The collected samples were identified by professional experts on Korean algae. The total DNA was extracted using the Exgene Plant SV Kit (GeneAll, Seoul, Republic of Korea) following the manufacturer's directions.

Complete genome construction and annotation

Genomic library construction and DNA sequencing were performed using an Illumina NovaSeq 6000 Platform yielding 151 bp paired-end reads by Macrogen Inc. (Seoul, Republic of Korea). Circular organelle genomes were constructed by comparing and integrating two assemblers, NOVOPlasty and GetOrganelle (Dierckxsens et al. 2017; Jin et al. 2020c) and with one mapping tool, Geneious Prime software. The adapter sequences were trimmed using fastp (Chen et al. 2018) for NOVOPlasty following the developer's recommendation, and quality filtering was conducted for GetOrganelle and Geneious Prime. The reference sequences were provided for Geneious Prime (mandatory) and NOVOPlasty (optional). For the construction of the organelle genomes of newly completed *G. asiatica* and *P. lanceolata* in this study, we used the *G. asiatica* cp genome (AP018129.2) and mt genome (OR221180.1) as the references. For *P. affinis*, the *Halymenia maculata* chloroplast genome (NC_046751.1) and the *P. affinis* mitochondrial genome (NC_063816.1) were used as references. The *rbcL* and *cox3* genes were used as the seeds in assembling cp and mt genomes, respectively.

Genome annotation was performed by comparing the annotations from two tools, GeSeq (Tillich et al. 2017), a web-based organelle genome annotation tool, and Geneious Prime. The tRNAscan-SE (Chan and Lowe 2019) was used to predict tRNA genes, and OGDRAW (Greiner et al. 2019) was used to draw organelle genome maps. GeSeq, including tRNAscan-SE and OGDRAW, was performed on the website https://chlorobox.mpimp-golm.mpg.de/index.html. The names and symbols of the genes were manually standardized by comparing the organelle genomes across the species used in this study.

Repeat sequence analysis

Simple sequence repeats (SSRs) were identified using MISA (Beier et al. 2017) with the parameters set 10 > for mono-, > 5 for di- and tri-, and > 3 for tetra-, penta-, and hexa-nucleotide. Dispersed long repeats of four types (forward, palindromic, reverse and complementary) were calculated using REPuter (Kurtz et al. 2001) with the parameters of hamming distance 3, maximum computed repeats 50, and minimal repeat size 30 bp.

Comparative analysis

To detect the rearrangements, inversions and synteny of genes among Florideophyceae species, the progressive-Mauve (Darling et al. 2004) implemented in Geneious was used with the newly assembled *G. asiatica* organelle genome (OR635816.1 for cp genome and OR635819.1 for mt genome) as the references. The similarities of the newly assembled organelle genomes of the three species were plotted using the mVISTA web-based program with Shuffle-LAGAN mode using the *G. asiatica* organelle genomes (AP018129.2 for cp genome and OR221180.1 for mt genome) downloaded from NCBI as the references. To estimate the nucleotide diversity value (Pi), the organelle genomes of the three species were aligned using MAFFT v7.520 (Katoh and Standley 2013), and then a sliding window analysis was performed using DnaSP v6.12.03 (Rozas et al. 2017). The window length was set to 300 bp for cp genomes and 200 bp for mt genomes, and the step size was set to 200 bp for cp genomes and 20 bp for mt genomes. Relative synonymous codon usage (RSCU) was analyzed for all protein-coding genes (PCGs) using CodonW program (http://codonw.sourceforge.net/) and visualized with Phylosuite (Zhang et al. 2020).

To investigate the selective pressure on the PCGs of organelle genomes, the amino acid sequences of the PCGs were aligned using MUSCLE v5.1 (Edgar 2022), and the aligned sequences were converted to nucleotide sequences using PAL2NAL (Suyama et al. 2006) with the nogap option. The alignment files were converted to pairwise aligned axt format by the script AXTConvertor implemented in KaKs_Calculator 2.0 (Wang et al. 2010), and the files were used to calculate non-synonymous (Ka), synonymous (Ks) substitutions and Ka/Ks value with Yang–Nielsen (YN) algorithm calculation method.

Phylogenetic analysis

The organelle genome data of 15 species in class Florideophyceae, 12 species in Bangiophyceae, a sister taxon of Florideophyceae, and one in Compsopogonophyceae, *Compsopogon caeruleus* as the outgroup were retrieved from NCBI GenBank to reconstruct the phylogenetic trees (Supplementary Table S1). The 194 cp and 23 mt PCGs (Supplementary Table S2) shared by the species were aligned using MAFFT, trimmed using trimAl (Capella-Gutiérrez et al. 2009), and concatenated using SequenceMatrix (Vaidya et al. 2011). Maximum-likelihood (ML) phylogenetic trees were constructed using IQ-TREE 2 (Minh et al. 2020) with 1000 bootstrap replicates, 1000 SH-aLRT replicates, and an approximate Bayes test. The best-fit models were chosen by ModelFinder implemented in IQ-TREE 2 using the option "-m TESTNEW".

Development of species-specific markers

The single-nucleotide polymorphisms (SNPs) of the aligned PCG sequences were detected, and the Beacon Designer (PRIMER Biosoft, Palo Alto, CA, USA) was used to design species-specific primer pairs based on polymorphic sites. qRT-PCR was performed on a Quantstudio 3 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with AccuPower® 2X GreenStar qPCR Master Mix (Bioneer, Daejeon, Korea).

Results

Characteristics of the organelle genomes of *Grateloupia asiatica*, Pachy*meniopsis lanceolata* and *Polyopes affinis*

For the three newly sequenced species, Illumina PE sequencing generated 100,694,720 (*G. asiatica*), 123,689,536 (*P. lanceolata*), and 124,205,314 (*P. affinis*) raw reads. Complete circular organelle genomes of the three species were constructed by comparing and combining the results of two assemblers, GetOrganelle and NOVOPlasty, and Geneious mapping (Fig. 1). These two assemblers sometimes include ambiguous nucleotides in their results and fail to generate a complete genome sequence. Additionally, in the case of mapping, a deletion in a genome relative to the reference sequence can result in addition of sequence to the map of the corresponding region. In our previous studies, we observed a deletion in the rRNA region (one *rrs* and one *rrl*) of *N. seriata* compared to that of other two species, *N. yezoensis* and *N. dentata*. However, unlike the results of the two assemblers, the mapping results placed virtual sequences in the deleted region (Lee et al. 2023). We thus compared the results of these three methods to derive the complete sequence. The cp and mt genomes of *P. lanceolata* and the cp genome of *P. affinis* (OR635818.1, OR635821.1, and OR635817.1, respectively) were documented for the first time here.

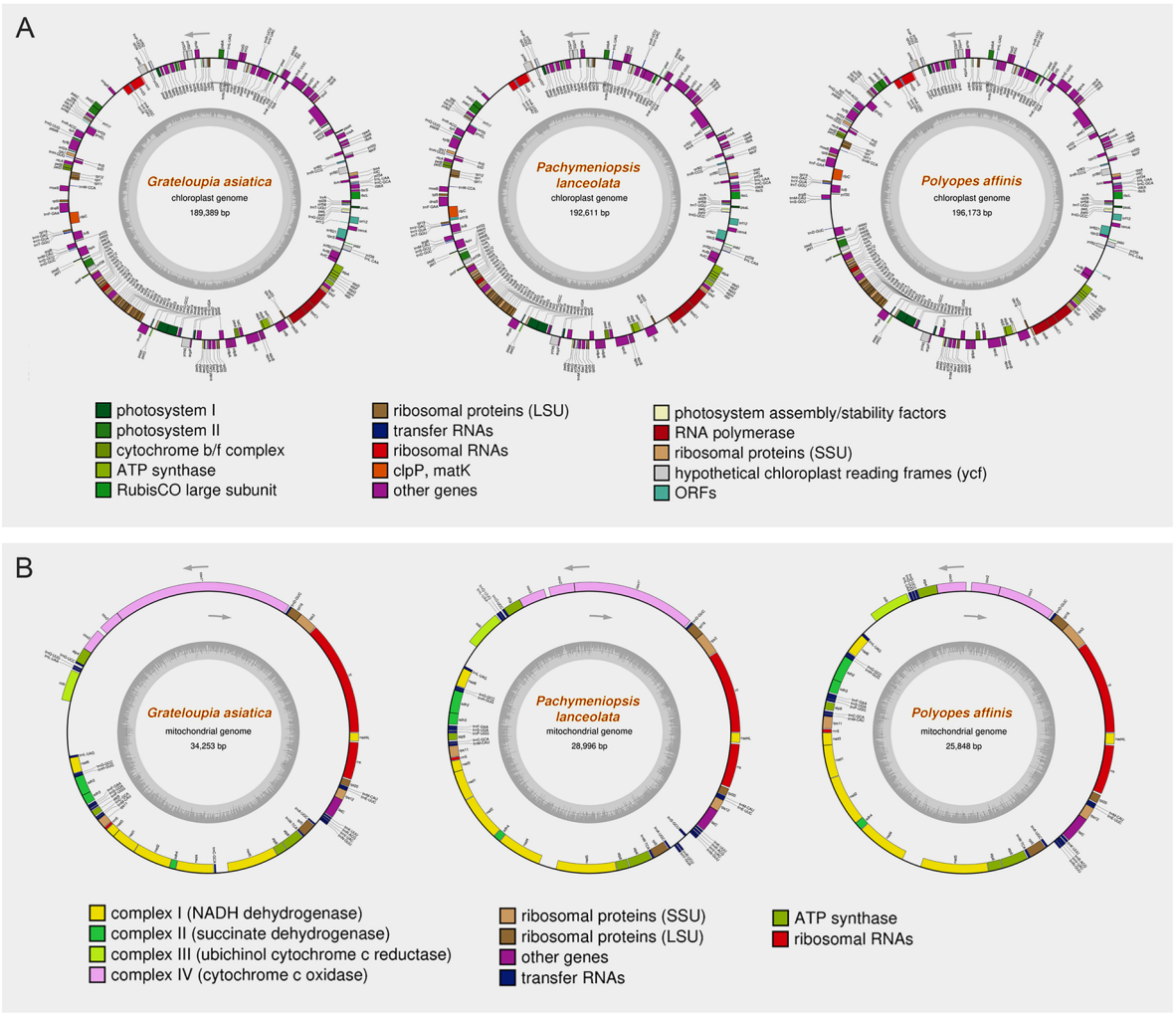


Fig. 1 The complete circular genome maps of chloroplast **A** and mitochondria **B** of three species (*Grateloupia asiatica*, *Pachymeniopsis lanceolata*, and *Polyopes affinis*, from left to right). Genes inside and outside of the maps are transcribed in a clockwise and counterclockwise direction

According to the yield of NOVOPlasty assembler, the coverage varied among the organelle genomes of three species (*G. asiatica* cp, 1842X; *P. lanceolata* cp, 3536X; *P. affinis* cp, 8416X; *G. asiatica* mt, 2148X; *P. lanceolata* mt, 2447X; *P. affinis* mt, 10,276X). The cp genome size was 189,389 bp (*G. asiatica*), 192,611 bp (*P. lanceolata*) and 196,173 bp (*P. affinis*), respectively. The average GC content was 29.93%. Across three species, there were 206 PCGs, 29 tRNA genes and three rRNA genes (Table 1). Although tRNA genes and rRNA genes were located identically across three species, some missing PCGs were found between species. Among a total of 206 PCGs, two genes in *G. asiatica* (*orf16* and *orf18*), four genes in *P. lanceolata* (*orf13*, *orf16*, *orf24*, and *orf48*) and seven genes in *P. affinis* (*grx*, *orf13*, *orf18*, *orf25*, *orf48*, *ycf21*, and *ycf56*) were missing.

The average length of mt genomes was 29,699 bp (34,253 bp for *G. asiatica*; 28,996 bp for *P. lanceolata*; 25,848 bp for *P. affinis*) with an average GC content of 31.87%. All mitochondrial genomes contained 25 unique PCGs and three rRNA genes. A total of 22 tRNA genes were found among mt genomes, while *trnR-UCU*, *trnS-GCU* and *trnY-GUA* were found only in *P. lanceolata* mt genome. All mt genomes had two copies of *trnM-CAU*, and *G. asiatica* had two copies of *trnC-GCA*.

All organellar genes are listed and divided into several categories according to their functions (Supplementary Table S3). The cp genes were divided into seven categories, genetic system, photosystems, ATP synthesis, metabolism, transport, RNA genes and unknown, and further divided into 26 groups. The mt genes were divided into four categories, oxidative phosphorylation, genetic system, RNA genes and transport, and 10 groups.

Table 1 Basic characteristics of complete organelle genomes of the three species

Organelle	Characteristic	*Grateloupia asiatica*	*Pachymeniopsis lanceolata*	*Polyopes affinis*
Chloroplast	Genome size (bp)	189,389	192,611	196,173
	GC content (%)	30.3	30.2	29.3
	Protein coding genes no.	204	202	199
	tRNA genes no.	30	29	29
	rRNA genes no.	3	3	3
Mitochondria	Genome size (bp)	26,543	32,498	35,863
	GC content (%)	30	31.9	32.4
	Protein coding genes no.	24	25	26
	tRNA genes no.	24	24	24
	rRNA genes no.	2	2	2

SSRs and repeats analysis

Among cp genomes, six (*G. asiatica*), 13 (*P. lanceolata*) and 10 (*P. affinis*) SSRs with a total of nine repeat units were identified. The most abundant repeat sequence was A/T, and G/C repeats were not identified. SSRs of the di-nucleotide type were not found among the three species. Tri-nucleotide repeat types, AAG/CTT, AAT/ATT, AGC/CTG and AGG/CCT, were the most abundant, and a total of 16 repeats were found. One tetra- and three penta-nucleotide repeat units were found only once in cp genomes, AGAT/ATCT and AAATT/AATTT of *P. affinis*, AAAAC/GTTTT of *P. lanceolata*, and AATAT/ATATT of *G. asiatica*. The repeat units in the mt genomes were found only once (A/T and AGG/CCT in *G. asiatica*, and AAT/ATT and ATGCC/ATGGC in *P. lanceolata*), and no repeats were found in the mt genome of *P. affinis* (Fig. 2A).

Although REPuter was set up to detect four types (forward, palindromic, reverse, and complementary) of repeats, only two (forward and palindromic) were found in the organellar genomes of three species (Fig. 2B). Repeats in the range of 30 to 39 bp were identified to be the most abundant repeats in both forward and palindromic repeat types in cp genomes. For most repeats found in mt genomes, one repeat per length ranges was identified, except for 80–89 bp in *P. affinis* and > 100 bp in *P. lanceolata*.

Comparative analysis of organelle genomes

The synteny of genes was identified among species in class Florideophyceae. In class Florideophyceae, the order Halymeniales showed a highly conserved gene order among the species, and *Palmaria palmata* (order Palmariales) and *Sporolithon durum* (Sporolithales) had almost identical gene orders (Fig. 3A). Compared to other species in Florideophyceae, except for *Hildenbrandia rubra* (Hildenbrandiales), *Ahnfeltia plicata* (Ahnfeltiales) and *Chondrus crispus* (Gigartinales) had one inversion (from *rps6* to *ycf46*, a total of 31 genes) in their cp genomes. The mt genomes of species in Florideophyceae also showed highly conserved gene order except for *Hildenbrandia rubra* (Fig. 3B).

We used mVISTA to identify organelle genomic divergence among the three species, with reference sequences of previously reported *G. asiatica* (Fig. 4). These results revealed that the gene-coding regions (purple) were more conserved compared to the non-coding regions (pink, except for rRNA regions around 63–68 k in cp genomes and 0–2.3 k, 22 k, and 34–35.3 k in mt genomes) in both cp and mt genomes. In particular, the mitochondrial *cox1* gene has two introns in *G. asiatica*, one intron in *P. lanceolata*, and no introns in *P. affinis*, as shown in Fig. 4B.

To further analyze the divergence hotspots of the organelle genomes among three species, a sliding window analysis of nucleotide diversity π (Pi) was performed. The Pi value of the three cp genomes ranged from 0.011 to 0.382 with the most divergent (Pi > 0.3) non-coding regions of *ycf80-trnR(CCG)*, *ccsA-psaK*, *ycf22-psaC*, *rrn16-ompR*, *rpl11-moeB*, *clpC-rpl19* and *orf65-ycf45* and the protein-coding regions of *infB*, *ycf92*, *orf621*, and *orf12* (Fig. 5A). We also observed eight hotspots with high sequence variability (Pi > 0.29), including one non-coding region, *cob-trnL(UAG)*, and seven protein-coding regions of *rps3*, *atp4*, *sdh3*, *nad5*, *atp8*, *rpl5*, and *tatC* in the mt genomes (Fig. 5B). The Pi value of the three mt genomes ranged from 0.023 to 0.432.

We estimated the RSCU for all PCGs in the organelle genomes of the three species (Supplementary Table S4) and plotted the results in Fig. 6. In the cp genomes, 204 PCGs consisted of 48,641 codons (*G. asiatica*), 202 PCGs consisted of 48,642 codons (*P. lanceolata*), and 199 PCGs consisted of 47,893 codons (*P. affinis*). Leucine was the predominant amino acid in PCGs with a count of 5067 (10.42% in *G. asiatica*), 5037 (10.36% in *P. lanceolata*), and 4995 (10.43% in *P. affinis*), followed by isoleucine (average counts of 4914, 10.16%), lysine (average counts of 3507, 7.25%), and serine

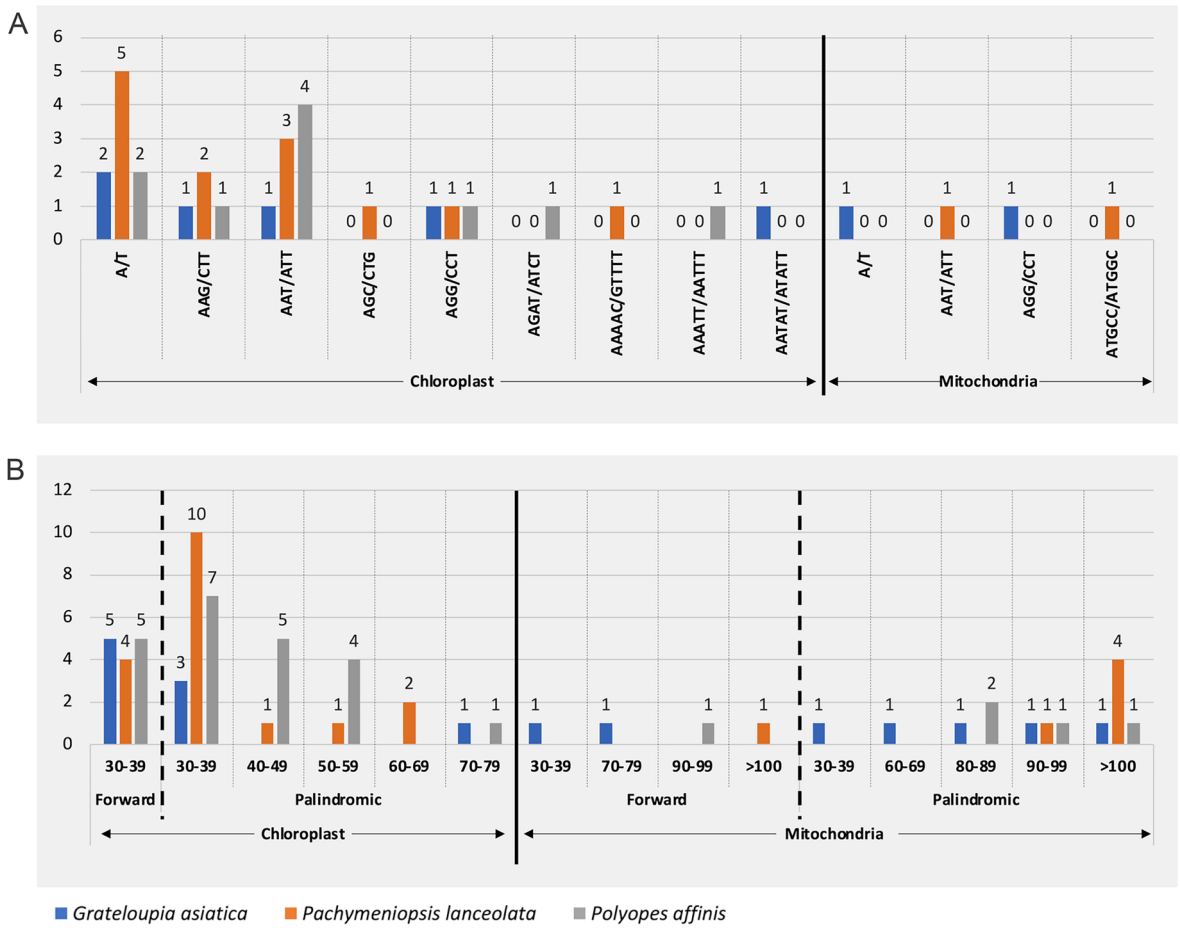


Fig. 2 Comparison of the repeat sequences of the organelle genomes of three species. **A**: units and number of simple sequence repeats (SSRs). X-axis: repeat units; Y-axis: frequency. **B**: Long and interspersed repeats. X-axis: repeat length; Y-axis: copy number

(average counts of 3424, 7.08%), while cysteine (average counts of 486, 1%) was the least used amino acid among the cp genomes of the three species. The two most frequently used codons were AAA (lysine) and UUA (leucine). AAU (asparagine) and AUU (isoleucine) were the third and fourth most frequently used codons in *P. lanceolata* and *P. affinis*, whereas they were the fourth and third most frequently used codons in *G. asiatica*. The codons AGA (arginine) and UUA were found to have the highest RSCU values (average values of 3.35 and 3.24, respectively). The mt genomes of all three species contained 25 PCGs, with 6088 codons (*G. asiatica*), 6079 codons (*P. lanceolata*), and 6066 codons (*P. affinis*). Like the cp genomes, the mt genomes also contained leucine as the most abundant amino acid, followed by isoleucine (average counts of 918 and 601, respectively). Serine was the third most abundant amino acid in the cp genomes but fourth in the mt genomes, with phenylalanine being the third most abundant amino acid (average counts of 548 and 505 for phenylalanine and serine, respectively). UUA was identified as the most frequently used codon and had the highest RSCU value in mt genomes (average counts of 497 and RSCU value of 3.25). In all cases, it was confirmed that A/U bases appeared preferentially at the third position of the codon, rather than C/G, and all codons with RSCU > 1 ended in A or U.

The non-synonymous (Ka) and synonymous (Ks) substitutions and their ratios (Ka/Ks) were calculated for the organelle genomes of seven Halymeniaceae accessions to evaluate natural selection pressure on nucleotides. The three chloroplast genes in the four comparisons (*psbV* in *G.asiatica* [AP018129.2] vs. *G. filicina*, *ycf33* in *G. turuturu* vs. *P. lanceolata*, *orf621* in *G. filicina* vs. *G. turuturu*, and *G. asiatica* [AP018129.2] vs. *G. turuturu*) had Ka/Ks values over 1.0, but were not statistically significant (Fisher's test *P*-value > 0.05), which were consequently excluded from the results. The significant pairwise Ka/Ks values between species are plotted as a heatmap in Supplementary Fig. S1. Two genes (cp gene *ycf92* in *G. turuturu* vs. *P. affinis*, and *G.*

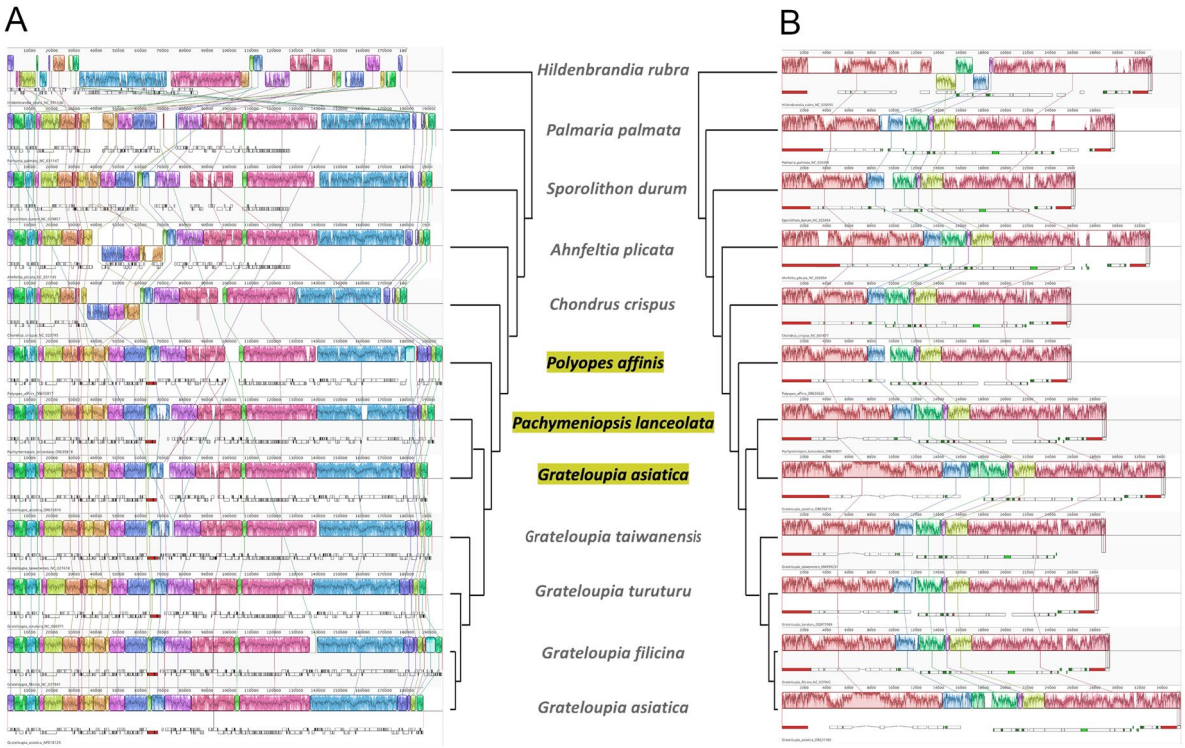


Fig. 3 Comparison of the complete chloroplast genomes **A** and mitochondria genomes **B** of 12 species of Florideophyceae using the Mauve algorithm. Local co-linear blocks indicate syntenic regions

taiwanensis vs. *G. turuturu*, and mt gene *rpl20* in *G. filicina* vs. *P. affinis*) had Ka/Ks values over 0.5, and in all other cases, the Ka/Ks values were less than 0.5.

Phylogenetic analysis

Phylogenetic relationships were inferred from the aligned, trimmed, and concatenated sequences of 194 and 23 PCGs of cp and mt genomes, respectively (Supplementary Table S2). The cp PCGs were obtained from 26 NCBI GenBank accessions and mt PCGs were obtained from 30 accessions. The phylogenetic trees were reconstructed using the concatenated cp PCGs sequence matrix (Supplementary Fig. S2a), mt PCGs sequence matrix (Supplementary Fig. S2b), and combined matrix of the cp and the mt matrices (Supplementary Fig. S2c) with the best-fit models of GTR + F + I + R5, TIM3 + F + I + R7, and GTR + F + R7, respectively. To construct a combined phylogenetic tree, the combined sequence matrix was concatenated by selecting species for which both cp and mt accessions were available (25 species). According to the results of the cp-based phylogenetic tree, species of the order Halymeniales (including four genera: Grateloupia, Halymenia, Pachymeniopsis, and Polyopes) were divided into two clades (*P. affinis* and *H. maculata* were clustered together and separated from other Halymeniales species), but were closely clustered. In the mt-based ML tree, the Halymeniales species were also divided into two clades, with exception of one species (previously reported *P. affinis*, NC_63816). The other two Polyopes species were grouped into the same clade as Halymeniales species, but NC_63816 was grouped into a different clade together with *Chondrus crispus* (Gigartinales). Together with the combined tree results, Polyopes species seems to be evolutionarily distant among Halymeniales species. In all three ML trees, all species belonging to the class Florideophyceae clustered with the order Bangiales (one of the orders of the class Bangiophyceae).

Development of species-specific markers for the identification of three Halymeniaceae species

The PCG sequences of the three cp genomes were aligned to detect the SNPs, and we developed 15 species-specific markers (five markers per species) from six genes (*accD*, *atpB*, *ccsA*, and *gltB* for all the three species, *rbcL* for *G. asiatica* and *P. affinis*, and *ilvH* for *P. lanceolata*) (Fig. 7). The markers were designed with PCR product sizes ranging from 103 to 223 bp, cut-off Ct ranging from 24 to 26 cycles, and melting temperatures ranging from 56 to 61.4 ℃

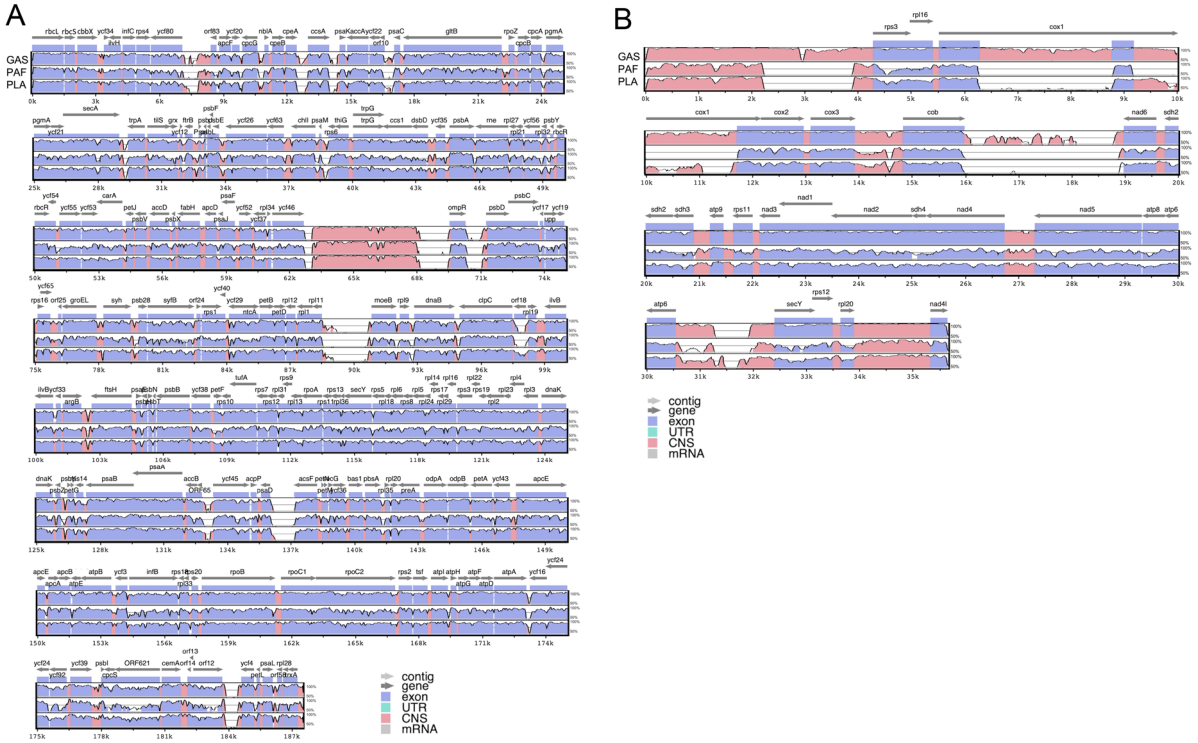


Fig. 4 Sequence alignment and comparison of three species using mVISTA in chloroplast genomes **A** and mitochondrial genomes **B**. Previously reported organelle genomes of *G. asiatica* (AP018129.2 for cp genome and OR221180.1 for mt genome) were used as alignment references. *GAS G. asiatica*, *PAF P. affinis*, *PLA P. lanceolata*

(Supplementary Table S5). We focused on cp genomes for the development of SNP-based markers because cp genomes have more conserved sequences than those of mt genomes.

Discussion

Overview of organelle genome assembly

Organelle genomes play an important role in DNA barcoding, molecular marker development and phylogenomics (Cai et al. 2021; Li et al. 2021; Yang and Kim 2015). In this study, we obtained the complete organelle genomes of *Grateloupia asiatica*, *Pachymeniopsis lanceolata* and *Polyopes affinis*, which included the first cp genome of *P. lanceolata* and *P. affinis* and the first mt genome of *P. lanceolata*. Our results provided insights into the genomic architecture, gene content and structural variations, forming a basis for further studies.

Gene composition, pseudogenization and selection pressure

Gene composition across the three species was highly conserved, particularly in the mitochondrial genomes. The missing PCGs were not found in the mitochondria, but appeared in the cp genomes, and most of them were functionally unknown genes, such as *orf13*/*16*/*18*/*24*/*25*/*48*/ and *ycf21*/*56*. *P. affinis* exhibited pseudogenization of *grx*, leading to truncated functionality. This aligns with the previous finding, pseudogenization commonly occurs in cp genomes (Frailey et al. 2018; Liu et al. 2020; Scobeyeva et al. 2021). These findings shed light on the evolutionary dynamics of the chloroplast genome in Florideophyceae and highlighted the potential genetic marker development for further studies.

Ka/Ks analysis supported that all PCGs are experiencing purifying selection, with no gene showing a significant Ka/Ks > 1. Most genes had Ka/Ks values < 0.1, suggesting the strong evolutionary constraints for maintaining functionality. These findings aligned with those of previous studies emphasizing the evolutionary stability of organelle genomes in Florideophyceae (Wang et al. 2010; Wu et al. 2020; Yu et al. 2019; Zhou et al. 2023).

Codon usage bias analysis revealed a preference for codons ending in A or U, which is consistent with the trend observed among red algae (Lee et al. 2022; Lu et al. 2023; Munyao et al. 2020; Xu et al. 2023). This bias reflects the potential optimization of translational efficiency and adds to our understanding of the evolutionary dynamics influencing

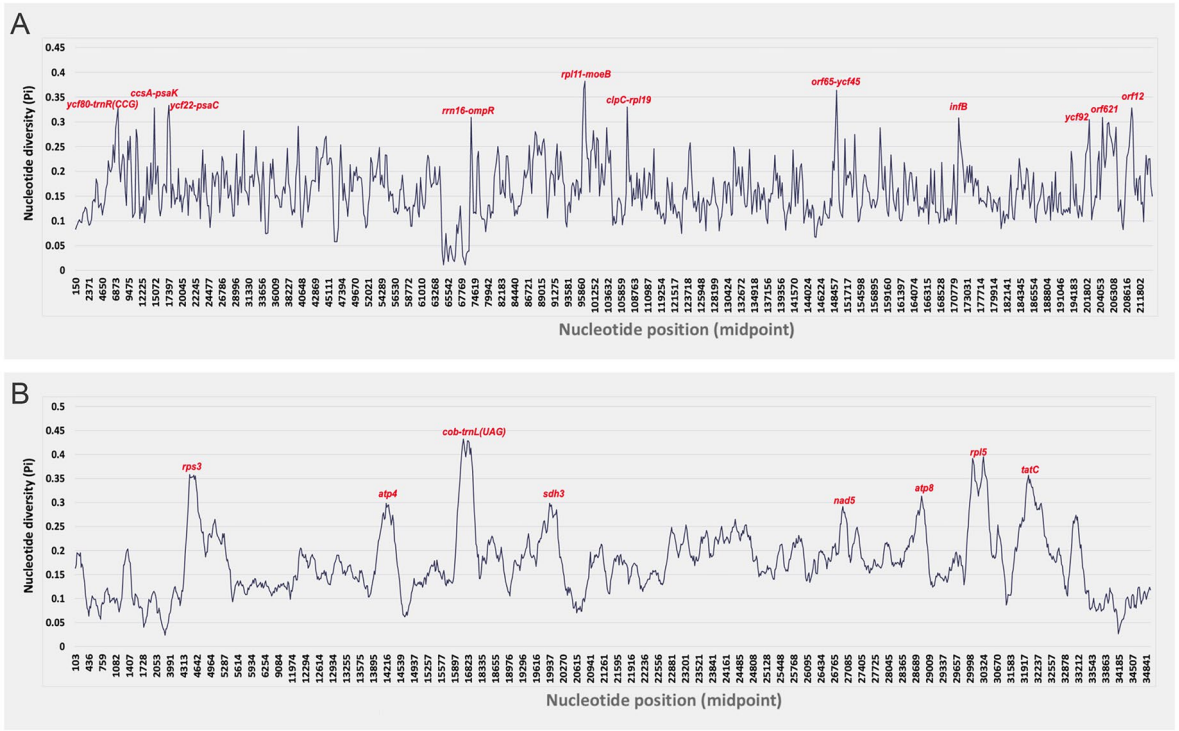


Fig. 5 Sliding window analysis of the complete chloroplast genomes **A** and mitochondria genomes **B** of the three species

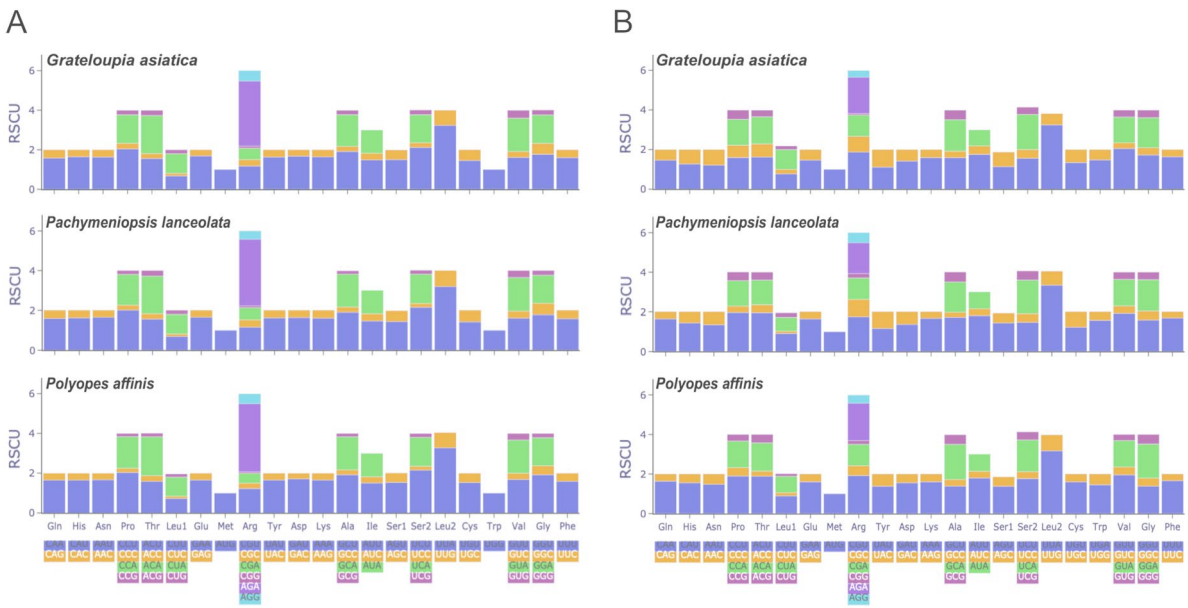


Fig. 6 The relative synonymous codon usage (RSCU) in all protein-coding genes of chloroplast genomes **A** and mitochondrial genomes **B**

organelle genomes (Plotkin and Kudla 2011; Sharp and Matassi 1994; Yang et al. 2023; Yao et al. 2023; Zhou et al. 2016).

Structural variation in organelle genomes

Unexpectedly, chloroplast genomes in Florideophyceae displayed more structural rearrangements than mitochondrial

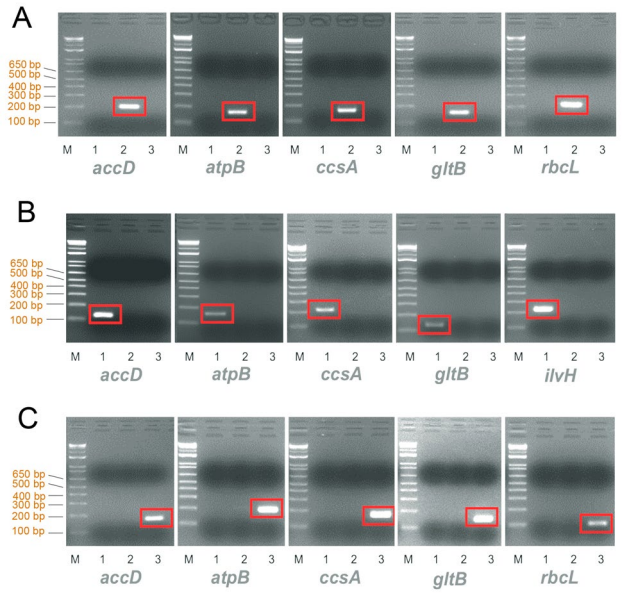


Fig. 7 Development of quantitative PCR based on the single nucleotide polymorphisms (SNPs) of chloroplast PCGs. Target species are **A**: *G. asiatica*, **B**: *P. lanceolata*, and **C**: *P. affinis*. Lane 1: *P. lanceolata*, 2: *G. asiatica*, 3: *P. affinis*, M: DNA ladder

genomes (Fig. 3), in contrast to the typical trends observed in other taxa (Smith and Keeling 2015). In particular, an inverse region of 31 genes, *rps6* to *ycf46*, was identified in *Ahnfeltia plicata* and *Chondrus crispus*. Although there are still many uncertainties about the gene rearrangements, repeat elements are known to be one of the important factors in such rearrangements (Maréchal and Brisson 2010; Ogihara et al. 1988; Tremblay-Belzile et al. 2015; Zampini et al. 2015; Zhang et al. 2012). In the cp genomes of *Palmaria palmata* and *Ahnfeltia plicata*, inverted rRNA gene (rRNA operon) repeats were identified on both sides of the region from *rps6* to *ycf46*. Lee et al. (2016) illustrated the structures of the cp genomes of red algae and seed plants, and divided the cp genomes of Florideophyceae into three types according to their structural characteristics. Consistent with their results, *Palmaria palmata*, *Sporolithon durum*, and seven other species belonging to Halymeniales, including the three species assembled in this study, were the R2-type, and two species with an inversion region (*A. plicata* and *C. crispus*) were the R1-type. The difference between the R1- and R2-type is that the inversion region is identical to the region between the inverted repeat rRNA operons. In both R1- and R2-types of cp genomes, there are species in which one set of rRNA operons was lost during the evolutionary process (such as *S. durum* and Halymeniale in R2-type, and *C. crispus* in R1-type). Many cp genomes found in diaphoretickes (Archaeplastida, by primary endosymbiosis, such as Viridiplantae, and the Sar supergroup, by secondary endosymbiosis, such as brown algae) have a typical quadripartite architecture consisting of a large single copy (LSC), a small single copy (SSC), and inverted repeats (IRs), where the IRs most likely contain an rRNA operon and several genes or pseudogenized sequences. Loss of IRs occasionally occurs in cp genomes; however, information regarding its cause is limited (Bedoya et al. 2019; Jin et al. 2020a, b).

Sequence similarity and evolutionary hotspots

Sequence similarity analysis demonstrated lower identity in mt genomes (73.4%) than in cp genomes (81.1%) across the seven Halymeniales species. The three species used in this study showed average similarities of 74% and 66.54% in the cp and mt genomes, respectively, and the average similarity of all genes (including PCGs, rRNAs, and tRNAs) was 84.8% and 70.1% in the cp and mt genomes. mVISTA and sliding window analyses identified divergent hotspots (Figs. 4 and 5), highlighting regions with high variability that could be valuable for marker development. The lower similarity in mt genomes may reflect a faster evolutionary rate, providing a fine-scale resolution for species-level differentiation (Smith and Keeling 2015).

Phylogenetic insights

The phylogenetic analyses revealed the relationship among the species in Florideophyceae and highlighted the close association of *P. lanceolata* with Grateloupia species. The newly sequenced cp genome of *G. asiatica* (OR635816) was found to be distant from the previously reported cp genome of *G. asiatica* (AP018129), possibly because AP018129 is a "nearly" complete genome. The combined cp and mt genome-based trees closely resembled the cp-based tree, likely due to the higher number of PCGs in chloroplast genomes. These findings highlight the utility of chloroplast genomes in resolving evolutionary relationships within Florideophyceae.

Six species belonging to order Bangiales and class Bangiophyceae (blue boxes in Supplementary Fig. S2) were clustered together with the species of the class Florideophyceae (red boxes in Supplementary Fig. S2), used to reconstruct phylogenetic trees based on both cp and mt genomes. The taxonomic classifications of the species used in this study were according to NCBI taxonomy, but previous studies differed from NCBI's classification (Lee et al. 2016; Park et al. 2023; Yang et al. 2015). In the results of these researchers, six other species (*C. chilensis*, *C. caldarium*, *D. grisea*, *E. coxiae*, *G. sulphuraria*, and *P. aerugineum*) belonging to the order Bangiophyceae (excluding Bangiales) were classified into different orders such as Cyanidiophyceae or Porphyridophyceae. Such inconsistencies reinforce the need for an

integrative approach that combines genomic data with taxonomic frameworks to resolve ambiguities. The inclusion of newly sequenced genomes in this study provides a foundation for future research on the evolutionary divergence and adaptation mechanisms in red algae.

Species-specific marker development

The high AT content and repetitive nature of cp genomes present challenges for marker development. To overcome this limitation, we targeted protein-coding genes (PCGs) in chloroplast genomes to design species-specific PCR markers. These efforts resulted in 15 specific markers, five in each species, that could distinguish the studied species. These markers are important to ensure species identification in food and medicinal applications (Hardouin et al. 2022; Shabaka and Moawad 2021; Wang et al. 2020), particularly for red algal resources that are prone to misuse or misidentification during processing. The successful development of these markers highlights the practical application of genomic data to address the challenges associated with species identification. They also provided a scalable framework for marker development in other red algal species.

Future perspectives

Our findings in this study highlighted the importance of organelle genomics in advancing our understanding of Florideophyceae diversity and resolving their evolutionary relationship. The newly sequenced organelle genomes provided foundational data for filling in the gaps in red algal research. Expanding organelle genomic data across underrepresented taxa could refine our understanding of red algal evolution and facilitate further phylogenetic studies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42995-025-00327-4.

Acknowledgements This research was supported by the Korea Institute of Marine Science & Technology Promotion (KMIST) funded by the Ministry of Oceans and Fisheries, Korea [grant number 20210656].

Author contributions YJL performed assembly, annotation, data analysis and wrote the manuscript. YMK and CMK performed experiments. S-JC performed sample validation. CSJ initiated the project and modified the manuscript. All authors approved the manuscript.

Data availability The NCBI GenBank accession numbers of organellar genomes of the three species are shown below: 1. *Grateloupia asiatica* chloroplast genome: OR635816.1 (and registered in NCBI Reference Sequence: NC_084248.1). 2. *Grateloupia asiatica* mitochondrial genome: OR635819.1. 3. *Pachymeniopsis lanceolata* chloroplast genome: OR635818.1 (and registered in NCBI Reference Sequence: NC_084250.1). 4. *Pachymeniopsis lanceolata* mitochondrial genome: OR635821.1 (and registered in NCBI Reference Sequence: NC_084251.1). 5. *Polyopes affinis* chloroplast genome: OR635817.1 (and registered in NCBI Reference Sequence: NC_084249.1). 6. *Polyopes affinis* mitochondrial genome: OR635820.1. 7. Raw genome sequencing reads of *Grateloupia asiatica*: SRR26313966 in SRA. 8. Raw genome sequencing reads of *Pachymeniopsis lanceolata*: SRR26313964 in SRA. 9. Raw genome sequencing reads of *Polyopes affinis*: SRR26313965 in SRA.

Declarations

Conflict of interest The authors declare no competing interests.

Animal and human rights statement Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Bedoya AM, Ruhfel BR, Philbrick CT, Madriñán S, Bove CP, Mesterházy A, Olmstead RG (2019) Plastid genomes of five species of riverweeds (Podostemaceae): structural organization and comparative analysis in Malpighiales. Front Plant Sci 10:1035

Beier S, Thiel T, Münch T, Scholz U, Mascher M (2017) MISAweb: a web server for microsatellite prediction. Bioinformatics 33:2583–2585

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:e0250968

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973

Chan PP, Lowe TM (2019) tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol 1962:1–14

Chen S, Zhou Y, Chen Y, Gu J (2018) Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890

Cui Y, Liu T, Wang X, Qu J, Jia X (2019) The complete chloroplast genome of *Sargassum horneri* and its phylogenetic analysis. Mitochondr DNA B Resour 4:3312–3313

Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403

Dierckxsens N, Mardulyn P, Smits G (2017) Novoplasty: *de novo* assembly of organelle genomes from whole genome data. Nucleic Acids Res 45:e18

Edgar RC (2022) Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun 13:6968

Frailey DC, Chaluvadi SR, Vaughn JN, Coatney CG, Bennetzen JL (2018) Gene loss and genome rearrangement in the plastids of five hemiparasites in the family Orobanchaceae. BMC Plant Biol 18:30

Greiner S, Sobanski J, Bock R (2015) Why are most organelle genomes transmitted maternally? BioEssays 37:80–94

Greiner S, Lehwark P, Bock R (2019) Organellargenomedraw (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res 47:W59–W64

Guiry MD (2024) How many species of algae are there? A reprise. Four kingdoms, 14 phyla, 63 classes and still growing. J Phycol 60:214–228

Ha Y, Lee W-H, Kim JK, Jeon H-K, Lee J, Kim Y-J (2022) *Polyopes affinis* suppressed IFN-γ- and TNF-α-induced inflammation in human keratinocytes via down-regulation of the NF-κB and STAT1 pathways. Molecules 27:1836

Hardouin K, Bouyer R, Pliego-Cortés H, Le Men T, Cérantola S, Marty C, Douzenel P, Bedoux G, Bourgougnon N (2022) Chemical characterization of the introduced red alga *Polyopes lancifolius* (Halymeniales, Rhodophyta) from the Gulf of Morbihan, France. Phycologia 61:616–627

Hwang EK, Park CS (2020) Seaweed cultivation and utilization of Korea. Algae 35:107–121

Jin D-M, Jin J-J, Yi T-S (2020a) Plastome structural conservation and evolution in the clusioid clade of Malpighiales. Sci Rep 10:9091

Jin D-M, Wicke S, Gan L, Yang J-B, Jin J-J, Yi T-S (2020b) The loss of the inverted repeat in the putranjivoid clade of Malpighiales. Front Plant Sci 11:942

Jin J-J, Yu WB, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z (2020c) Getorganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol 21:241

Kandale A, Meena AK, Rao MM, Panda P, Mangal AK, Reddy G, Babu R (2011) Marine algae: an introduction, food value and medicinal uses. J Pharm Res 4:219–221

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780

Khan MNA, Choi JS, Lee MC, Kim E, Nam TJ, Fujii H, Hong YK (2008) Anti-inflammatory activities of methanol extracts from various seaweed species. J Environ Biol 29:465–469

Kim SY, Boo SM, Yoon HS, Kim MS (2023) Insight into the species identification and distribution of Grateloupiaceae (Halymeniales, Rhodophyta) having *Grateloupia filicina*-like morphology in the Northwest Pacific. Algae 38:23–38

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R (2001) Reputer: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res 29:4633–4642

Lee J, Cho CH, Park SI, Choi JW, Song HS, West JA, Bhattacharya D, Yoon HS (2016) Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. BMC Biol 14:75

Lee YJ, Kim YD, Uh YR, Kim YM, Seo T-H, Choi S-J, Jang CS (2022) Complete organellar genomes of six *Sargassum* species and development of species-specific markers. Sci Rep 12:20981

Lee YJ, Uh YR, Kim YM, Kim CM, Jang CS (2023) Characterization and comparative analysis of the complete organelle genomes of three red macroalgae species (*Neoporphyra dentata*, *Neoporphyra seriata*, and *Neopyropia yezoensis*) and development of molecular makers for their identification. Genes Genomics 46:355–365

Li JR, Ma HX, Yang XQ, Chen SJ, Chen SW, Qi B, Li CS (2019) Nutritional composition analysis and sensory evaluation of *Grateloupia lanceolata* at different growth stages. Food Ferment Indust 45:255–260

Li R, Jia X, Zhang J, Jia S, Liu T, Qu J, Wang X (2021) The complete plastid genomes of seven Sargassaceae species and their phylogenetic analysis. Front Plant Sci 12:747036

Liu L-X, Du Y-X, Folk RA, Wang S-Y, Soltis DE, Shang F-D, Li P (2020) Plastome evolution in Saxifragaceae and multiple plastid capture events involving *Heuchera* and *Tiarella*. Front Plant Sci 11:361

Liu Y-J, Zhang T-Y, Wang Q-Q, Draisma SGA, Hu Z-M (2024) Comparative structure and evolution of the organellar genomes of *Padina usoehtunii* (Dictyotales) with the brown algal crown radiation clade. BMC Genomics 25:747

Lu G, Wang W, Mao J, Li Q, Que Y (2023) Complete mitogenome assembly of *Selenicereus monacanthus* revealed its molecular features, genome evolution, and phylogenetic implications. BMC Plant Biol 23:541

Maréchal A, Brisson N (2010) Recombination and the maintenance of plant organelle genome stability. New Phytol 186:299–317

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534

Munyao JN, Dong X, Yang JX, Mbandi EM, Wanga VO, Oulo MA, Saina JK, Musili PM, Hu GW (2020) Complete chloroplast genomes of *Chlorophytum comosum* and *Chlorophytum gallabatense*: genome structures, comparative and phylogenetic analysis. Plants 9:296

Ogihara Y, Terachi T, Sasakuma T (1988) Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. Proc Natl Acad Sci 85:8573–8577

Park SI, Cho CH, Ciniglia C, Huang T, Liu S, Bustamante DE, Calderon MS, Mansilla A, McDermott T, Andersen RA, Yoon HS (2023) Revised classification of the Cyanidiophyceae based on plastid genome data with descriptions of the Cavernulicolales ord. nov. and Galdieriales ord. nov. (Rhodophyta). J Phycol 59:444–466

Plotkin JB, Kudla G (2011) Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet 12:32–42

Rawal HC, Kumar PM, Bera B, Singh NK, Mondal TK (2020) Decoding and analysis of organelle genomes of Indian tea (*Camellia assamica*) for phylogenetic confirmation. Genomics 112:659–668

Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34:3299–3302

Saunders GW (2005) Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Philos Trans R Soc Lond Ser B Biol Sci 360:1879–1888

Scobeyeva VA, Artyushin IV, Krinitsina AA, Nikitin PA, Antipin MI, Kuptsov SV, Belenikin MS, Omelchenko DO, Logacheva MD, Konorov EA, Samoilov AE, Speranskaya AS (2021) Gene loss, pseudogenization in plastomes of genus *Allium* (Amaryllidaceae), and putative selection for adaptation to environmental conditions. Front Genet 12:674783

Shabaka S, Moawad M (2021) Ecology and biochemical composition of a newly reported non-indigenous red alga, *Grateloupia gibbesii*, in the Mediterranean Sea, with reference to edible red seaweeds. Reg Stud Mar Sci 44:101767

Sharp PM, Matassi G (1994) Codon usage and genome evolution. Curr Opin Genet Dev 4:851–860

Smith DR, Keeling PJ (2015) Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. Proc Natl Acad Sci 112:10177–10184

Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res 34:W609–W612

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S (2017) Geseq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res 45:W6–W11

Tremblay-Belzile S, Lepage É, Zampini É, Brisson N (2015) Short-range inversions: rethinking organelle genome stability. BioEssays 37:1086–1094

Vaidya G, Lohman DJ, Meier R (2011) Sequencematrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27:171–180

Verlaque M, Brannock PM, Komatsu T, Villalard-Bohnsack M, Marston M (2005) The genus *Grateloupia* C. Agardh (Halymeniaceae, Rhodophyta) in the Thau Lagoon (France, Mediterranean): a case study of marine plurispecific introductions. Phycologia 44:477–496

Villanueva M-J, Morcillo M, Tenorio M-D, Mateos-Aparicio I, Andrés V, Redondo-Cuenca A (2014) Health-promoting effects in the gut and influence on lipid metabolism of *Himanthalia elongata* and *Gigartina pistillata* in hypercholesterolaemic Wistar rats. Eur Food Res Technol 238:409–416

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

Wang X, He L, Ma Y, Huan L, Wang Y, Xia B, Wang G (2020) Economically important red algae resources along the Chinese coast: history, status, and prospects for their utilization. Algal Res 46:101817

Wang L, Wang J, Zhu Y, Cui Z, Kong F, Tang X, Mao Y (2021) Development of organelle single nucleotide polymorphism (SNP) markers and their application for the identification of cytoplasmic inheritance patterns in *Pyropia yezoensis* (Bangiales, Rhodophyta). J Ocean Limnol 39:1447–1457

Wu Z, Liao R, Yang T, Dong X, Lan D, Qin R, Liu H (2020) Analysis of six chloroplast genomes provides insight into the evolution of *Chrysosplenium* (Saxifragaceae). BMC Genomics 21:621

Xu S, Teng K, Zhang H, Gao K, Wu J, Duan L, Yue Y, Fan X (2023) Chloroplast genomes of four *Carex* species: long repetitive sequences trigger dramatic changes in chloroplast genome structure. Front Plant Sci 14:1100876

Yang MY, Kim MS (2015) Taxonomy of *Grateloupia* (Halymeniales, Rhodophyta) by DNA barcode marker analysis and a description of *Pachymeniopsis volvita* sp. nov. J Appl Phycol 27:1373–1384

Yang EC, Kim KM, Kim SY, Lee JM, Boo GH, Lee JH, Nelson WA, Yi G, Schmidt WE, Fredericq S, Boo SM, Bhattacharya D, Yoon HS (2015) Highly conserved mitochondrial genomes among multicellular red algae of the Florideophyceae. Genome Biol Evol 7:2394–2406

Yang M, Ma L, Yang X, Li L, Chen S, Qi B, Wang Y, Li C, Wei Y, Zhao Y (2022) Photosynthetic protein-based edible quality formation in various *Porphyra dentata* harvests determined by label-free proteomics analysis. Cells 11:1136

Yang L, Deng S, Zhu Y, Da Q (2023) Comparative chloroplast genomics of 34 species in subtribe Swertiinae (Gentianaceae) with implications for its phylogeny. BMC Plant Biol 23:164

Yao H, Li T, Ma Z, Wang X, Xu L, Zhang Y, Cai Y, Tang Z (2023) Codon usage pattern of the ancestor of green plants revealed through Rhodophyta. BMC Genomics 24:538

Yu X, Zuo L, Lu D, Lu B, Yang M, Wang J (2019) Comparative analysis of chloroplast genomes of five *Robinia* species: genome comparative and evolution analysis. Gene 689:141–151

Zampini É, Lepage É, Tremblay-Belzile S, Truche S, Brisson N (2015) Organelle DNA rearrangement mapping reveals u-turn-like inversions as a major source of genomic instability in *Arabidopsis* and humans. Genome Res 25:645–654

Zhang T, Hu S, Zhang G, Pan L, Zhang X, Al-Mssallem IS, Yu J (2012) The organelle genomes of Hassawi rice (*Oryza sativa* L.) and its hybrid in Saudi Arabia: genome variation, rearrangement, and origins. PLoS ONE 7:e42041

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) Phylosuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour 20:348–355

Zhao X, Liu C, He L, Zeng Z, Zhang A, Li H, Hu Z, Lou S (2022) Structure and phylogeny of chloroplast and mitochondrial genomes of a chlorophycean algae *Pectinodesmus pectinatus* (Scenedesmaceae, Sphaeropleales). Life 12:1912

Zhou Z, Dang Y, Zhou M, Li L, Yu C-H, Fu J, Chen S, Liu Y (2016) Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proc Natl Acad Sci 113:E6117–E6125

Zhou C, Wang P, Zeng Q, Zeng R, Hu W, Sun L, Liu S, Luan F, Zhu Q (2023) Comparative chloroplast genome analysis of seven extant *Citrullus* species insight into genetic variation, phylogenetic relationships, and selective pressure. Sci Rep 13:6779

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.