J. Aquat. Plant Manage. 65(1): 1–13

Comparative chloroplast genomic features of purified and wild *Sargassum fusiforme* from China

SEN YANG, JIAANG CAO, LIJIE LIU, PENG ZHANG, TIEGAN WANG, AND LIDONG LIN

ABSTRACT

Sargassum fusiforme is a perennial, large, edible, and medicinal brown alga endemic to the low intertidal zones of the northwestern Pacific coast. It is primarily cultivated in China and South Korea, with limited cultivation in Japan, and holds significant economic value. Wild *Sargassum fusiforme* (W-*S. fusiforme*) inhabits the low intertidal zone, whereas the cultivated line purified *Sargassum fusiforme* (P-*S. fusiforme*) grows in floating-culture habitats. To uncover the structural characteristics and environmental adaptation features of the chloroplast genome (cpDNA) of *S. fusiforme* across different ecological niches, three types of P-*S. fusiforme* (P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3) and three types of W-*S. fusiforme* (W-*S. fusiforme* 4, W-*S. fusiforme* 5, W-*S. fusiforme* 6) from the coastal regions of China were selected in this study. The research involved sequencing, assembling, annotating, and conducting comparative analyses of the cpDNA. Both genomes consist of a large single-copy region (LSC: 73,437 base pairs [bp]) and a small single-copy region (SSC: 39,960 bp) separated by repeat sequences. The guanine (G) and cytosine (C) content in both types was found to be 30.43%. In *S. fusiforme* 6, the long repeat sequences lack a 38 bp forward repeat, whereas other *S. fusiforme* contain three forward repeats, 31 palindromic repeats, and one reverse repeat. Both P-*S. fusiforme* and W-*S. fusiforme* genomes harbor 225 simple sequence repeat (SSR) sites, with mononucleotide and trinucleotide repeats being the most abundant, accounting for 40.89% and 50.22% of the total SSRs. The remaining types of SSRs account for 8.89% of the total. The cpDNA of W-*S. fusiforme* contains 26 trinucleotide repeats, whereas that of P-*S. fusiforme* contains 25 trinucleotide repeats. Both P-*S. fusiforme* and W-*S. fusiforme* cpDNA contain dinucleotide and tetranucleotide repeats, accounting for 5.33% and 3.11% of the total SSRs. Only one hexanucleotide repeat was identified. The average nucleotide diversity of cpDNA in both P-*S. fusiforme* and W-*S. fusiforme* was calculated to be 0.0176. The results of codon usage frequency analysis indicate that the cpDNA of both P-*S. fusiforme* and W-*S. fusiforme* contains 29 high-frequency codons and 37 low-frequency codons. The results of the phylogenetic evolutionary tree indicate that P-*S. fusiforme* and W-*S. fusiforme* belong to the same branch, with confidence levels exceeding 96%, indicating a high degree of similarity in their chloroplast genomes. This study provides scientific data for research on the phylogenetic evolution, resource conservation, quality improvement, and breeding applications of *S. fusiforme*.

Key words: comparative analysis, complete chloroplast genome, purified, *Sargassum fusiforme*, wild.

INTRODUCTION

Sargassum fusiforme, a perennial warm-temperate seaweed, belongs to the phylum *Phaeophyta*, *Fucales* order, and Sargassaceae family. It is also known by various names such as "Yu Cao" and "Hordeum marinum," among others. Wild *Sargassum fusiforme* (W-*S. fusiforme*) is widely distributed along the coastal regions of China, ranging from the Liaodong Peninsula in the north to the Leizhou Peninsula in the south, with a particularly high concentration along the coast of Zhejiang Province (Lin and Wu 2021; Guiry 2024), and it is also found in countries such as South Korea and Japan. Since the 1990s, large-scale purification of *S. fusiforme* has been undertaken in the coastal areas of Zhejiang. Research on the biochemical components of *S. fusiforme* indicates that its algal body is rich in 18 essential amino acids and 14 trace elements required by the human body, highlighting its significant nutritional value (Ramdin et al. 2024). As a primary producer in the ocean, *S. fusiforme* provides habitats, breeding grounds, and food sources for various marine organisms (Nesic et al. 2024). *Sargassum fusiforme* is capable of efficiently absorbing nutrients such as nitrogen, phosphorus, and potassium from water bodies, as well as adsorbing heavy metal ions like cadmium, chromium, lead, and mercury (Tsaniya et al. 2021; Johnson et al. 2022), thereby effectively purifying the aquatic environment. Furthermore, the biochemical components extracted from *S. fusiforme*, such as polysaccharides, vitamins, amino acids, and sterols, hold significant economic value in the development of food, pharmaceuticals, and health products (Lomartire and Gonçalves 2022). China is one of the major producers of *S. fusiforme*, with an

First, second, third, and sixth authors, School of Life Sciences, Agriculture and Forestry, Qiqihar University, Qiqihar, 161006, China. Third and sixth authors, Dongtou District Marine Economic Science and Technology Innovation Center, Postdoctoral Workstation, Wenzhou, 325700, China. Fourth and fifth authors, Zhejiang Mariculture Research Institute, Zhejiang Key Laboratory of Coastal Biological Germplasm Resources Conservation and Utilization, Wenzhou, 325000, China; Zhejiang Mariculture Research Institute, Sino-Thai Joint Laboratory of Marine Technology, Wenzhou, 325000, China; and Zhejiang Mariculture Research Institute, Wenzhou Key Laboratory of Marine Biological Genetics and Breeding, Wenzhou, 325000, China. Corresponding authors' E-mail: liulijie2001@163.com, rsclinlidong@126.com, and zhangpeng20011918@163.com. Received for publication November 3, 2025 and in revised form July 19, 2026.

DOI: 10.57257/JAPM-D-25-00021

annual cultivation area of 800.4 ha. The country has pioneered innovative technologies in artificial cultivation, sexual reproduction seedling breeding, selection of superior strains, and hybrid breeding of *S. fusiforme*, leading the world in these areas. These advancements have made a significant contribution to the conservation of W-*S. fusiforme* resources and their sustainable development and utilization.

Chloroplasts are vital organelles found in the cells of terrestrial plants, algae, and certain protists. They play a crucial role in converting light energy into chemical energy and synthesizing organic compounds, thereby supplying the energy and organic materials necessary for growth and development. Additionally, chloroplasts are instrumental in energy conversion and carbon fixation within ecosystems, playing a significant role in maintaining carbon balance and energy flow (Obata et al. 2013; Ligrone 2019). cpDNA contains crucial genetic information and additional phylogenetic signals, making it an effective identifier for distinguishing and classifying species and variants. This enables accurate phylogenetic analysis (Zhu et al. 2024a). The highly conserved nature and slow evolutionary rate of chloroplasts, combined with the endosymbiotic theory, make them an ideal tool for studying molecular evolution (Johnston 2024; Raval et al. 2024; Zhang et al. 2025). The structure of the chloroplast genome (cpDNA) is generally divided into four parts: the large single-copy region (LSC), the small single-copy region (SSC), and two inverted repeat regions (IRa and IRb) (Li et al. 2022). In recent years, significant progress has been made in the study of cpDNA, primarily focusing on genome structure, function, evolution, and its applications in biotechnology and phylogenetics. These advancements are attributed to the close integration of structural genomics and functional genomics, which complement each other and collectively promote the widespread application of genomics in fields such as biology, medicine, and agriculture (Yang et al. 2013). With the sequencing of cpDNA from an increasing number of species, the economic and ecological value of various plants has been significantly explored and enhanced (Hagopian et al. 2004; Lemieux et al. 2016).

Currently, there is a wealth of research reports on *S. fusiforme*, covering various physiological and biochemical aspects such as photosynthesis, environmental stress responses, cellular protection mechanisms, active substances in fertilized eggs, and lipid metabolism (Hong et al. 2021; Liu and Lin 2020; Li et al. 2023b; Lin et al. 2019; Zhang et al. 2021). However, progress in the study of the *S. fusiforme* chloroplast genome has been relatively slow; few studies have been reported, and a complete cpDNA genome map of this species has yet to be successfully constructed (Wang et al. 2020), Studies of the chloroplast genome of *S. fusiforme* remain limited, and its functional information is largely unknown. In this study, we systematically analyzed the cpDNA sequences, structural features, and differences between purified *S. fusiforme* (P-*S. fusiforme*) and W-*S. fusiforme* specimens. These analyses provide important insights into the genetic characteristics of the *S. fusiforme* chloroplast and its ecological niche adaptations (Liu et al. 2012).

MATERIALS AND METHODS

Sample collection and processing

To uncover the structural characteristics and environmental adaptation features of the cpDNA of *S. fusiforme* across different ecological niches, three types of P-*S. fusiforme* (P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3) and three types of W-*S. fusiforme* (W-*S. fusiforme* 4, W-*S. fusiforme* 5, W-*S. fusiforme* 6) from the coastal regions of China were selected in this study.

On 1 October 2019, W-*S. fusiforme* 4 was collected from Moye Island, Shandong Province. On 11 December 2019, W-*S. fusiforme* 5 was collected from Nanao Island, Guangdong Province. On 22 December 2019, W-*S. fusiforme* 6 was collected from Chaiyu Island, Pingyang County, Wenzhou City, Zhejiang Province. After collection, these wild strains of *S. fusiforme* were isolated and temporarily cultivated in the experimental field at Xiandieyan Sea Area, Dongtou District, Wenzhou City. Balloon (P-*S. fusiforme* 1; parent: Shandong Mo Island, 2016 to 2019 purified four generations), wide rod sac (P-*S. fusiforme* 2; parent: wild *S. fusiforme* in Nanao Island, Guangdong, 2016 to 2019 purified four generations), and roddies (*P-S. fusiforme* 3; parent: wild *S. fusiforme* in Chaiyu Island, Pingyang County, Wenzhou, 2016 to 2019 purified four generations) were collected in Dongtou District, Wenzhou city on 22 March 2020. The three P-*S. fusiforme* strains and the above three W-*S. fusiforme* strains were 200 g each. (Note: The P-*S. fusiforme* preparation has remained in use to the present day.) After miscellaneous removal, fresh filtered seawater cleaning, and sterile seawater washing, the samples were marked, packed and placed in sealed pockets, kept at 4 C constant temperature, and mailed to Nanjing Jisi Huiyuan Biotechnology Co. within 24 h for entrusted testing and analysis. DNA library construction and sequencing were performed by Nanjing Geneisky Biotech Co., Ltd. (Nanjing, Jiangsu Province, China).

Preparation and sequencing of DNA library

The process was carried out in accordance with the standard protocol provided by Illumina, which included sample quality inspection, library construction, library quality inspection, and library sequencing. After the genomic DNA of the algal samples passed quality inspection, the DNA was fragmented using a mechanical method (ultrasonication). The fragmented DNA was then subjected to purification, end repair, addition of an "A" base to the 3′ end, and ligation of sequencing adapters. The fragments were size-selected using agarose gel electrophoresis, followed by PCR amplification to construct the sequencing library. The constructed library underwent quality control (QC) testing, and libraries that met the QC standards were sequenced on the Illumina Novaseq platform with a read length of PE150. The raw sequencing data were filtered using fastp (version 0.20.0, https://github.com/ OpenGene/fastp). (Sequencing was performed by Nanjing Jisi Huiyuan Biotechnology Co., Nanjing, Jiangsu, China.)

Bioinformatics analysis

The raw data were processed through data filtering to remove adapter sequences and low-quality reads, resulting

in high-quality clean data. We assembled the clean data according to the cpDNA of the reference species to obtain the cpDNA assembly results. We then annotated the gene structure of the cpDNA assembly results, created a cpDNA map, and analyzed basic content such as SSRs in the cpDNA.

Genome assembly

Extraction of cpDNA sequencing sequences. To reduce the complexity of subsequent sequence assembly, the cpDNA database (KU500638.1) was aligned using Bowtie2 (version 2.2.4, http://bowtie-bio. sourceforge. net/bowtie2/index.shtml) in the very-sensitive-local mode. The sequencing reads that successfully aligned to the cpDNA database were identified and extracted as the cpDNA sequences for the project samples.

Sequence assembly. SPAdes V3.10.1 (http://cab.spbu.ru/software/spades/) was utilized by the core assembly module for the construction of the chloroplast genome. The assembly process employed k-mer values of 55, 87, and 121 and was conducted without reliance on a reference genome, ensuring a *de novo* assembly approach. Due to the characteristics of second-generation sequencing, repetitive sequences in the genome, and specific genomic structures, it is not possible to directly obtain a complete circular genome sequence through a single assembly. Therefore, additional methods were employed to acquire the complete circular genome sequence.

Genome analysis

Extraction of cpDNA sequencing sequences. In this study, two distinct methodologies were employed for the annotation of cpDNA to enhance the accuracy of the annotation process. The first approach involved the utilization of Prodigal v2.6.3 (https://www.github.com/hyattpd/Prodigal) for the annotation of chloroplast coding sequences. The prediction of rRNA was conducted using HMMER v3.1b2 (http://www.hmmer.org/), while tRNA prediction was performed with ARAGORN v1.2.38 (http://130.235.244.92/ARAGORN/). The second method entailed extracting gene sequences from closely related species that have been published on NCBI. These sequences were then aligned against the assembled sequences using BLAST v2.6 (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to derive a secondary set of annotations. The annotations from both methods were manually scrutinized to identify discrepancies. Erroneous and redundant annotations were removed, and the boundaries of multiple exons were determined, culminating in the final annotation set.

Chloroplast genome map. The cpDNA map was constructed utilizing OGDRAW (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html).

IRscope analysis. The cpDNA is characterized by a circular structure, featuring four junctions between the inverted repeats (IRs) and the large single copy (LSC) and small single copy (SSC) regions. These junctions are identified as LSC-IRb, IRb-SSC, SSC-IRa, and IRa-LSC. During the evolutionary process of the genome, the boundaries of the IRs may undergo expansion and contraction, leading to the migration of certain genes into the IR regions or the single-copy regions. To visualize these boundary dynamics, the SVG module in Perl was employed to graphically represent the boundary information.

Phylogenetic analysis. The sequences from different species were aligned using MAFFT software (version 7.427, with the –auto option) to perform multiple sequence alignment, and the good aligned data were trimmed with trimAl(v1.4.rev15). Then phylogenetic tree construction was performed using the software RAxML v8.2.10 (https://cme.h-its.org/exelixis/software.html) and selecting the GTR model and hill climbing algorithm, with bootstrap = 1,000.

RESULTS

Structure and characteristics of cpDNA

The chloroplast DNA (cpDNA) of all six strains of *S. fusiforme* was assembled into single circular molecules. The average size of the cpDNA for P-*S. fusiforme* was 124,297 bp, while the mean size for W-*S. fusiforme* was slightly smaller at 124,293 bp (Table 1). Both P-*S. fusiforme* and W-*S. fusiforme* possess chloroplast DNA (cpDNA) comprising a total of 173 genes. This genetic repertoire includes 139 protein-coding genes, 28 transfer RNA (tRNA) genes, and six ribosomal RNA (rRNA) genes. The chloroplast genome exhibits a quadripartite structure, featuring a pair of IRs, each measuring 10,900 bp. These IRs are separated by a LSC region spanning 73,437 bp and a SSC region of 39,960 bp (Fig. 1). The GC content of DNA serves as a crucial indicator of species affinity. In this context, the GC content across the LSC region, the SSC region, the IRs, and the entire genome is uniformly 30.43%. The cpDNA of the six strains of *S. fusiforme* exhibits a high degree of conservation on both the heavy and light strands. This conserved structure includes two inverted repeat regions (IRA and IRB) and one intron. When all genes were functionally categorized, both P-*S. fusiforme* and W-*S. fusiforme* were found to possess 51 genes involved in photosynthesis. The most abundant among these were the genes associated with photosystem I and photosystem II, totaling 27 genes (including *psaA*, *psaB*, etc.). This suite of photosynthetic genes also encompasses those encoding for the cytochrome b6/f complex subunits (such as *petA* and *petB*, totaling nine genes), ATP synthase (*atpA* etc.), the large subunit of Rubisco (*rbcL* etc.), and chlorophyll reductase (*chlB* etc.) (Table 2). Notably, genes encoding NADH dehydrogenase were absent from this photosynthetic gene complement. The remaining genes include 45 genes encoding large and small ribosomal proteins (such as *rpl1* and *rps1*), four RNA polymerase genes, and three rRNA genes. The genome also includes one gene involved in protease synthesis, two genes participating in cytochrome c-type synthesis, 18 genes representing chloroplast open reading frames (ORFs), and an additional 18 genes with various other functions (Table 2). These findings indicate that the gene content and genomic organization of *S. fusiforme* exhibit a high degree of conservation.

Table 1. Sequence characteristics of chloroplast genome in *S. fusiforme*.

Feature name	P-*S. fusiforme* 1	P-*S. fusiforme* 2	P-*S. fusiforme* 3	W-*S. fusiforme* 4	W-*S. fusiforme* 5	W-*S. fusiforme* 6
Genome size (bp, base pairs)	124,297	124,297	124,297	124,298	124,295	124,286
Overall G+C content (%)	30.43	30.43	30.43	30.43	30.43	30.43
LSC length	73,437	73,437	73,437	73,438	73,435	73,436
SSC length	39,960	39,960	39,960	39,960	39,960	39,960
No. of genes	173	173	173	173	173	173
No. of protein-coding genes	139	139	139	139	139	139
No. of tRNA genes	28	28	28	28	28	28
No. of rRNA genes	6	6	6	6	6	6
Length of IRB	5,450	5,450	5,450	5,450	5,450	5,450

Condon usage frequency

Owing to the degeneracy of the genetic code, each amino acid is encoded by at least one codon. There is significant variation in codon usage frequency among different species and organisms. This unequal use of synonymous codons, where certain codons are preferred over others that encode the same amino acid, is known as codon usage bias or relative synonymous codon usage (RSCU). In the cpDNA of both P-*S. fusiforme* and W-*S. fusiforme*, the RSCU values range from 0.0064 to 3.974 (Supplemental Table 1, Fig. 2). Within this range, 29 codons exhibit RSCU values greater than 1, indicating a high preference for these codons in protein synthesis. Conversely, 37 codons have RSCU values less than 1, indicating a low preference for these codons in protein synthesis. Additionally, no codon usage bias was detected for tryptophan, as indicated by an RSCU value of 1 (Supplemental Table 1). All high-usage codons (RSCU > 1) predominantly feature G and C nucleotides at the third codon position (Fig. 2). It is evident that the cpDNA of both W-*S. fusiforme* and P-*S. fusiforme* exhibit convergent patterns in codon usage frequency.

Long repeat sequence analysis and simple repeat sequence analysis

A total of 35 long repeat sequences were identified in the cpDNA of three purified strains of *S. fusiforme* (P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3) and three wild-type strains

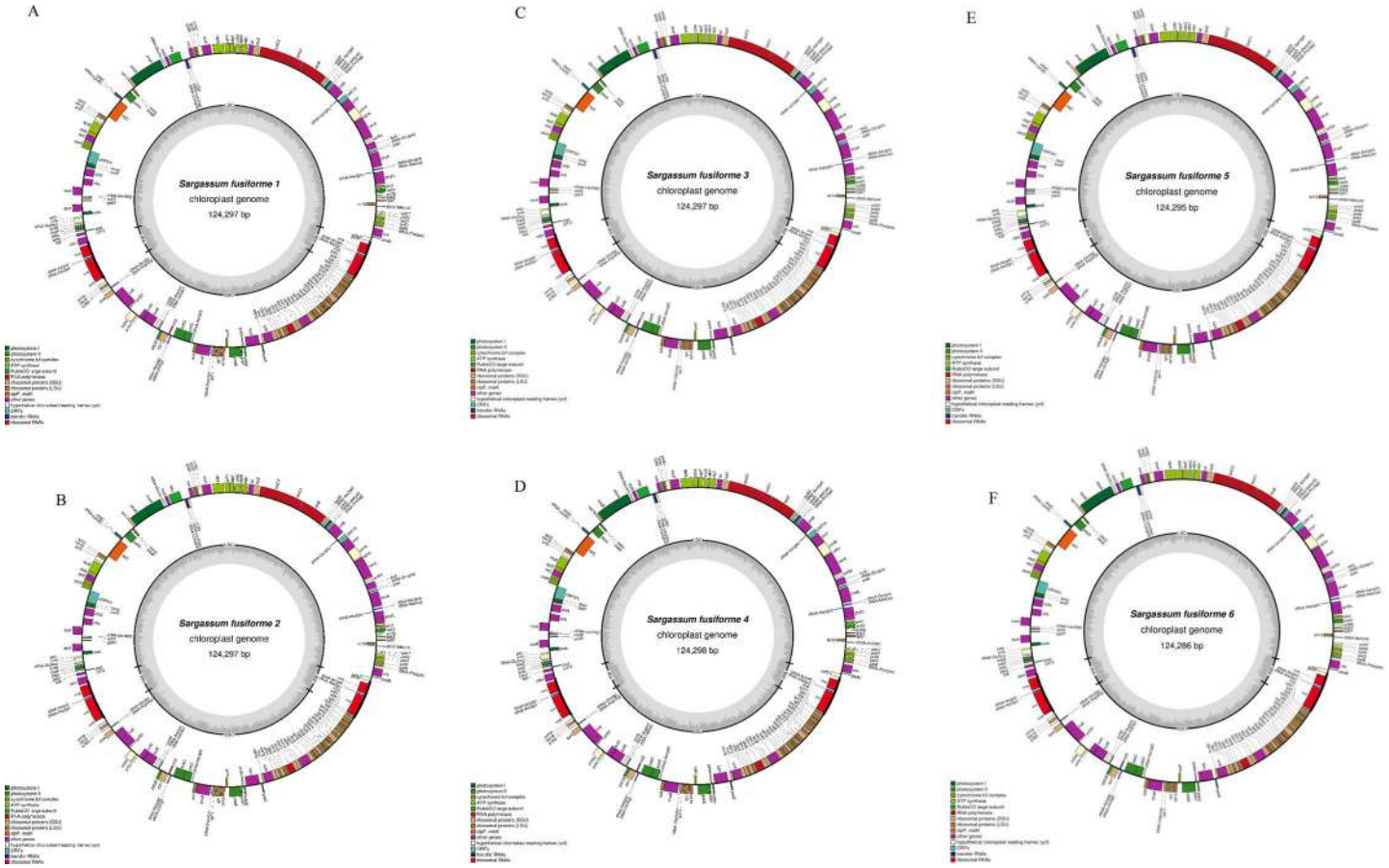


Figure 1. Chloroplast genome atlas. P-*S. fusiforme* 1 (A), P-*S. fusiforme* 2 (B), P-*S. fusiforme* 3 (C), W-*S. fusiforme* 4 (D), W-*S. fusiforme* 5 (E), and W-*S. fusiforme* 6 (F). The preceding genes are located on the outer side of the circle, while the anticoding genes are located on the inner side of the circle. The internal gray circle represents the GC content.

Table 2. Statistical table of functional classification of chloroplast genes.

Category	Gene group	Gene name[1]
Photosynthesis	Subunits of photosystem I	*psaA, psaB, psaC, p saD, p saE, psaF, psaI, psaJ, psaL, psaM*
	Subunits of photosystem II	*psb28, psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbN, psbT, psbV, psbX, psbY*
	Subunits of NADH dehydrogenase	—
	Subunits of cytochrome b/f complex	*petA, petB, petD, petF, petG, petJ, petL, petM, petN*
	Subunits of ATP synthase	*atpA, atpB, atpD, atpE, atpF, atpG, atpH, atpI*
	Large subunit of Rubisco	*rbcL, rbcR, rbcS*
	Subunits of photochlorophyllide reductase	*chlB, chlI, chlL, chlN*
Self-replication	Proteins of large ribosomal subunit	*rpl1, rpl11, rpl12, rpl13, rpl14, rpl16, rpl18, rpl19, rpl2, rpl20, rpl21, rpl22, rpl23, rpl24, rpl27, rpl29, rpl3, rpl31, rpl32, rpl33, rpl34, rpl35, rpl36, rpl4, rpl5, rpl6, rpl9*
	Proteins of small ribosomal subunit	*rps1, rps10, rps11, rps12, rps13, rps14, rps16, rps17, rps18, rps19, rps2, rps20, rps3, rps4, rps5, rps7, rps8, rps9*
	Subunits of RNA polymerase	*rpoA, rpoB, rpoC1, rpoC2*
	Ribosomal RNAs	*rrn5(2), rrnl(2), rrns(2)*
	Transfer RNAs	*tRNA-Ala(tgc)(2), tRNA-Arg(acg), tRNA-Arg(tct), tRNA-Asn(gtt), tRNA-Asp (gtc), tRNA-Cys(gca), tRNA-Gln(ttg), tRNA-Glu(ttc), tRNA-Gly(gcc), tRNA-Gly(tcc), tRNA-His(gtg), tRNA-Ile(gat)(2), tRNA-Leu(taa),[2] t NA-Leu(tag), tRNA-Lys(ttt), tRNA-Met(cat)(3), tRNA-Phe(gaa), tRNA-Pro(tgg), tRNA-Ser (gct), tRNA-Ser(tga), tRNA-Uhr(tgt), tRNA-Urp(cca), tRNA-Uyr(gta), tRNA-Val(tac)*
Genes of unknown function	Conserved hypothetical chloroplast ORF	*ORF501, ORF76, ycf12, ycf19, ycf3, ycf33, ycf34, ycf35, ycf37, ycf39, ycf4, ycf41, ycf42, ycf46, ycf47, ycf54, ycf65, ycf66*
Other genes	Maturase	—
	Protease	*clpC*
	Envelope membrane protein	—
	Acetyl-CoA carboxylase	—
	c-type cytochrome synthesis gene	*ccs1, ccsA*
	Translation initiation factor	—
	Other	*acsF, cbbx, dnaB, dnaK, ftrB, ftsH, groEL, ilvB, ilvH, secA, secY, sufB, sufC, tatC, thiG, thiS, tsf, tufA*

[1]Gene(no.): Genes with a copy number greater than 1, with their copy numbers indicated in parentheses.
[2]Contains an intron.

(W-*S. fusiforme* 4, W-*S. fusiforme* 5, W-*S. fusiforme* 6). Among these repeats, three were forward repeats, 31 were palindromic repeats, and one was a reverse repeat. Notably, a 38 bp forward repeat was absent only in *S. fusiforme* 6. In both P-*S. fusiforme* and W-*S. fusiforme*, a total of 225 SSRs were identified (Supplemental Table 2). The most abundant SSRs were mononucleotide and trinucleotide repeats, accounting for 40.89% and 50.22% of the total SSRs, respectively. In contrast, dinucleotide and tetranucleotide repeats were less frequent, comprising 5.33% and 3.12% of the total, In addition, a hexanucleotide was detected at 0.44%. (Fig. 3). W-*S. fusiforme* strains contain 26 trinucleotide repeats, whereas P-*S. fusiforme* strains have only 25. Among the 225 SSRs detected in both P-*S. fusiforme* and W-*S. fusiforme* 117 (52%) were A/T repeats, representing the highest frequency of all SSR types. These A/T repeats were followed by AT/AT repeats, which accounted for 29.3% of the total SSRs.

IR shrinkage and expansion analysis

Inverted repeats are a crucial structural feature of chloroplast genomes, playing a pivotal role in maintaining genomic stability, increasing gene redundancy, and facilitating DNA repair processes. The boundaries between the *IRs/SSC* region, as well as between the *IRs/LSC* region, along with their adjacent genes, were compared across the cpDNA of P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3, W-*S. fusiforme* 4, W-*S. fusiforme* 5, and W-*S. fusiforme* 6 (Fig. 4). The junctions between *LSC/IRb*, *IRb/SSC*, SSC/IRa, and *IRa/LSC* are highly conserved across all six Cp genomes. In P-*S. fusiforme* 1, P-*S. fusiforme* 2, and P-*S. fusiforme* 3, the length of the LSC region is consistently 73,437 bp. W-*S. fusiforme* 4 has an LSC length of 73,438 bp, W-*S. fusiforme* 5 measures 73,435 bp, and W-*S. fusiforme* 6 is 73,426 bp, resulting in an average LSC length of 73,433 bp across the wild-type strains. Both W-*S. fusiforme* and P-*S. fusiforme* exhibit a consistent pattern at the *LSC/IRb* junction, where the gene *cbbx* extends 718 bp from IRb into the LSC region and 188 bp from the LSC into IRb. At the *IRb/SSC* boundary, the gene *rrn5* is located 181 bp into the SSC region, while *ycf19* is positioned further from the *IRb/SSC* boundary, extending 396 bp into the SSC. At the *SSC/IRa* boundary, the gene *rpl21* extends 89 bp into the IRa region while retaining 229 bp within the SSC region. Additionally, *rrn5* is located 181 bp from the *SSC/IRa* boundary, and *ycf37* extends 7 bp into the LSC region from the *IRa/LSC* boundary.

Nucleotide diversity (Pi) analysis

The MAFFT software (in auto mode) was employed to perform a global alignment of homologous sequences from different species, and following the alignment, vcftools was utilized to calculate the Pi value (nucleotide diversity) for each gene. The analysis revealed that the average Pi value between P-*S. fusiforme* and W-*S. fusiforme* is 0.0176 (Supplemental Table 3). However, there is considerable variation in Pi values across different regions of the cpDNA. In the LSC region, the Pi values range from 0.002 to 0.0369, with an average value of

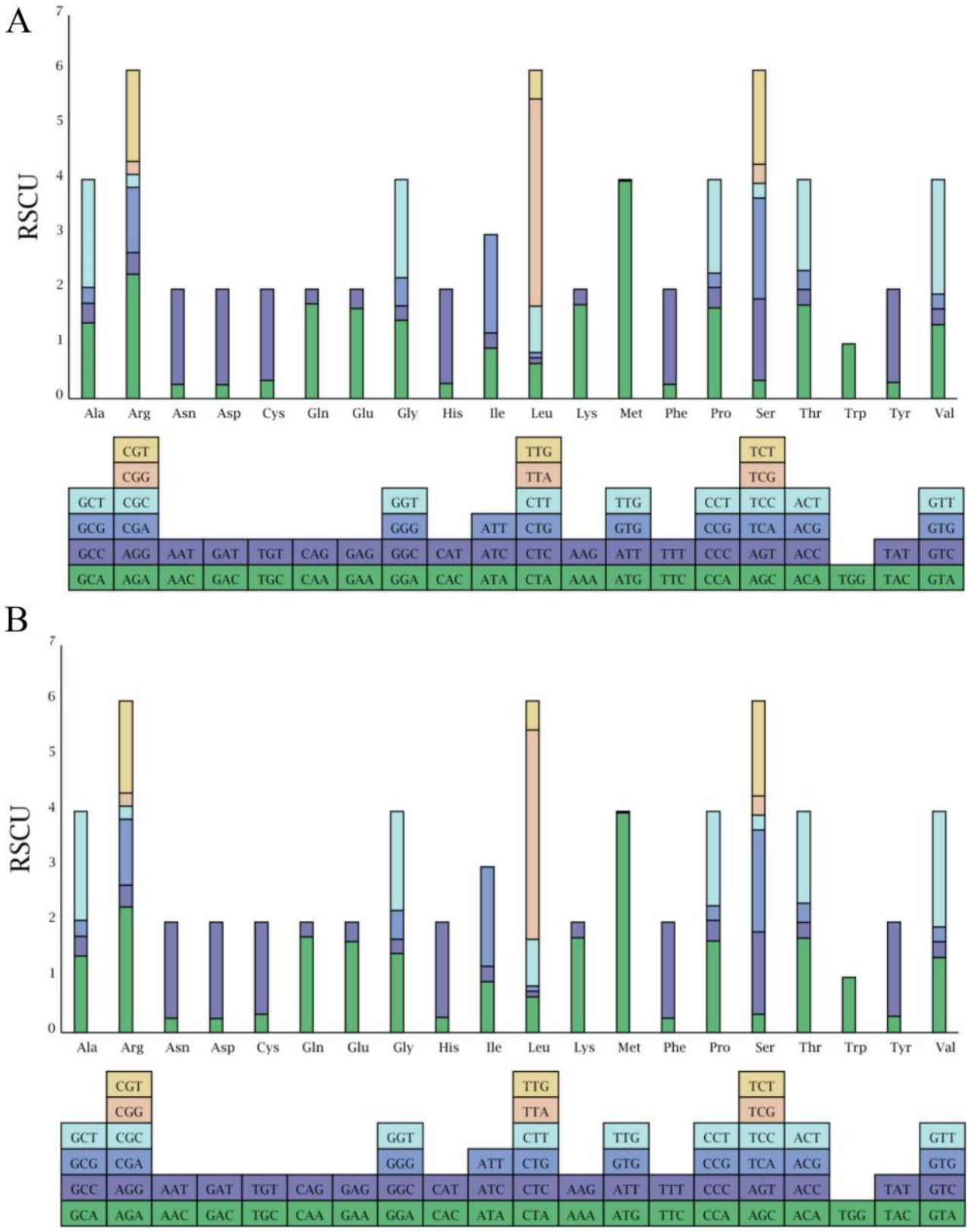


Figure 2. Codon usage bias in all protein-coding genes of the chloroplast genome. P- *S. fusiforme* (A); W-*S. fusiforme* (B). Vertical axis unit: %; horizontal axis: all codons encoding each amino acid.

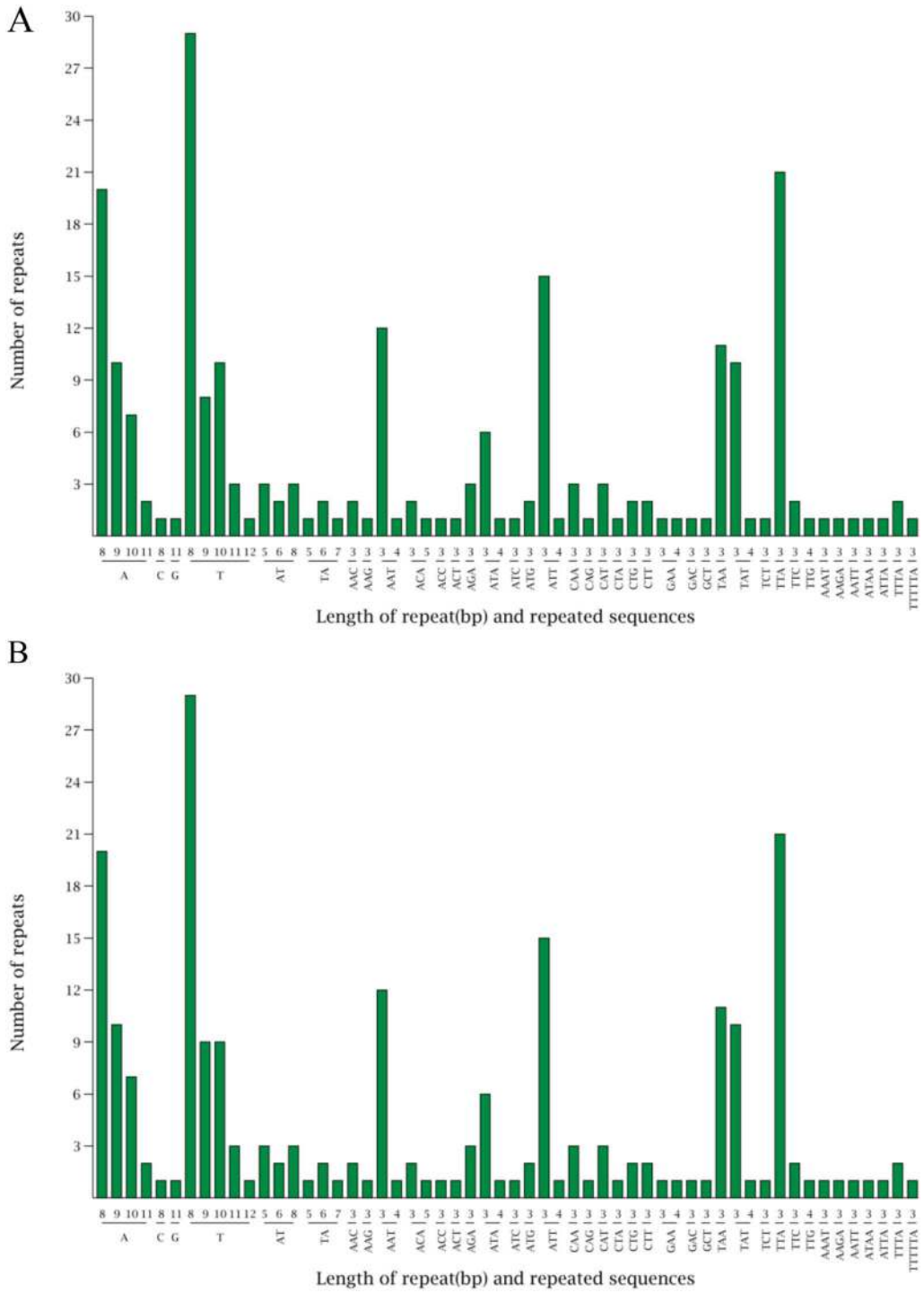


Figure 3. Statistical chart of quantity of SSR types. The horizontal axis represents SSR repeat units (units), and the vertical axis represents the number of SSR. (A) P-*S. fusiforme* from Dongtou, Zhejiang; (B) W-*S. fusiforme*.

0.0172. In the SSC region, the Pi values range from 0.008 to 0.0388, with an average value of 0.0182, the IR regions exhibit a much lower range of Pi values, from 0 to 0.004, with an average of 0.002. P-*S. fusiforme* and W-*S. fusiforme* exhibit a high degree of similarity, with only minor differences in regions where the Pi value exceeds 0.03. These regions exhibit significant variability within the genes *ycf33*, *ORF501*, *petJ*, *psaJ*, and *ycf39* (Fig. 5). In the cpDNA analysis of P-*S. fusiforme*, regions with Pi values greater than 0.03 include *psaJ*, *petJ*, *ORF501*, *thiS*, and *psaE* (Supplemental Table 3).

Phylogenetic analysis

A phylogenetic analysis was conducted using the cpDNA sequences of nine species: P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3, W-*S. fusiforme* 4, W-*S. fusiforme* 5, W-*S. fusiforme* 6,

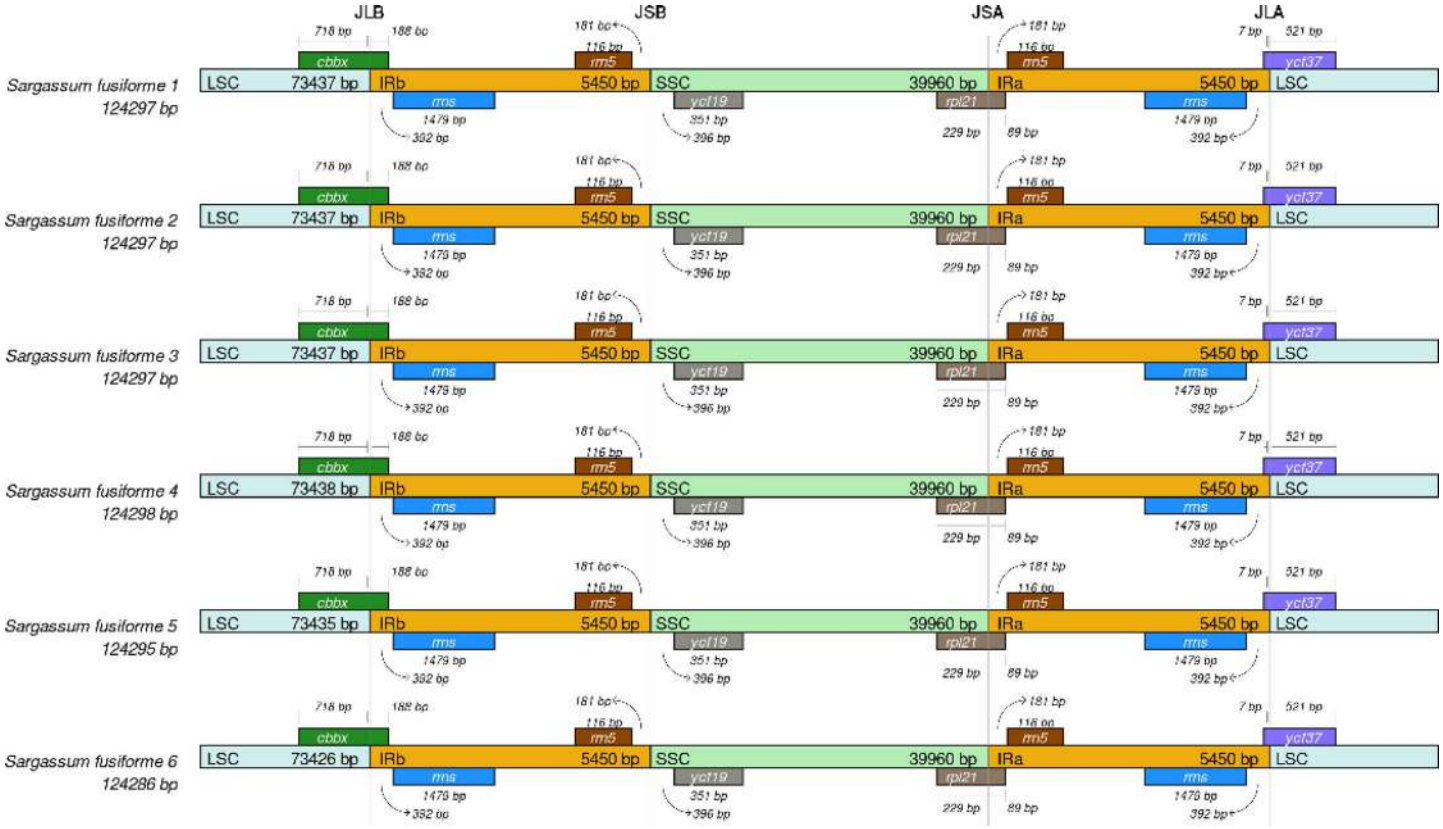


Figure 4. Analysis of chloroplast IR boundary changes. The thin lines represent the connection points of each region, and the information of genes near the connection points is presented on the graph.

Sargassum confusum, *Sargassum thunbergii*, and *Sargassum horneri*. These species represent four genera within the Sargassaceae family. The phylogenetic tree was constructed using the GTR model with the hill-climbing algorithm and bootstrap support of 1000 replicates. The nine *Sargassum* species were divided into four distinct branches in the phylogenetic tree. *Sargassum thunbergii* and *S. fusiforme* exhibited a close phylogenetic relationship. *Sargassum horneri* showed some differentiation from P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3, W-*S. fusiforme* 4, W-*S. fusiforme* 5, and W-*S. fusiforme* 6, but they still clustered within the same genus. P-*S. fusiforme* 1, P-*S. fusiforme* 2, P-*S. fusiforme* 3, W-*S. fusiforme* 4, W-*S. fusiforme* 5, and W-*S. fusiforme* 6 all clustered within a single branch with bootstrap support values exceeding 96%; this high level of confidence indicates a strong genetic similarity among the six strains of *S. fusiforme*, underscoring their close evolutionary relationship and genetic coherence. *Sargassum confusum* formed a distinct branch on its own, indicating a more distant phylogenetic relationship with *S. fusiforme* compared to the other species analyzed (Fig. 6).

DISCUSSION

Structure and characteristics of cpDNA

The cpDNA, characterized by its relatively low evolutionary rate and high degree of conservation, serves as an ideal tool for studying phylogenetic relationships and species evolution (Furan et al. 2024). Despite the high species diversity and ecological significance of *S. fusiforme* (Tian et al. 2023), there is a notable scarcity of genomic information available for this species. We comprehensively evaluated the phylogenetic structure of *S. fusiforme* through cpDNA sequencing, assembly, annotation, and comparative analyses of P-*S. fusiforme* and W-*S. fusiforme*. The results showed that the mean cpDNA length was 124,297 bp in P-*S. fusiforme* and 124,293 bp in W-*S. fusiforme*. Although W-*S. fusiforme* had a slightly smaller cpDNA genome than P-*S. fusiforme*, no associated gene loss was detected in the nuclear genome. Additionally, the variation in genome size among the strains follows the following order: W-*S. fusiforme* 4 > P-*S. fusiforme* 1, P-*S. fusiforme* 2, and P-*S. fusiforme* 3 > W-*S. fusiforme* 5 > W-*S. fusiforme* 6. This size variation is associated with the expansion and contraction of the IR regions (Li et al. 2023a). In both ecotypes of *S. fusiforme*, a total of 173 shared genes were annotated in the cpDNA, comprising 28 tRNA genes, 6 rRNA genes, and 139 mRNA genes. The cpDNA sequences of the two ecotypes of *S. fusiforme* exhibit a sequence similarity of 99.99%, indicating a very close genetic relationship between P-*S. fusiforme* and W-*S. fusiforme*.

Functional genomics aims to characterize the complete genetic information encoded in the genome and its functions, thereby elucidating the relationships among gene expression, regulation, and function, as well as the molecular mechanisms underlying biological processes such as growth and development (Przybyla and Gilbert. 2022). Chloroplasts originated from cyanobacterial ancestors (Falcón et al. 2010), but were progressively remodeled during the

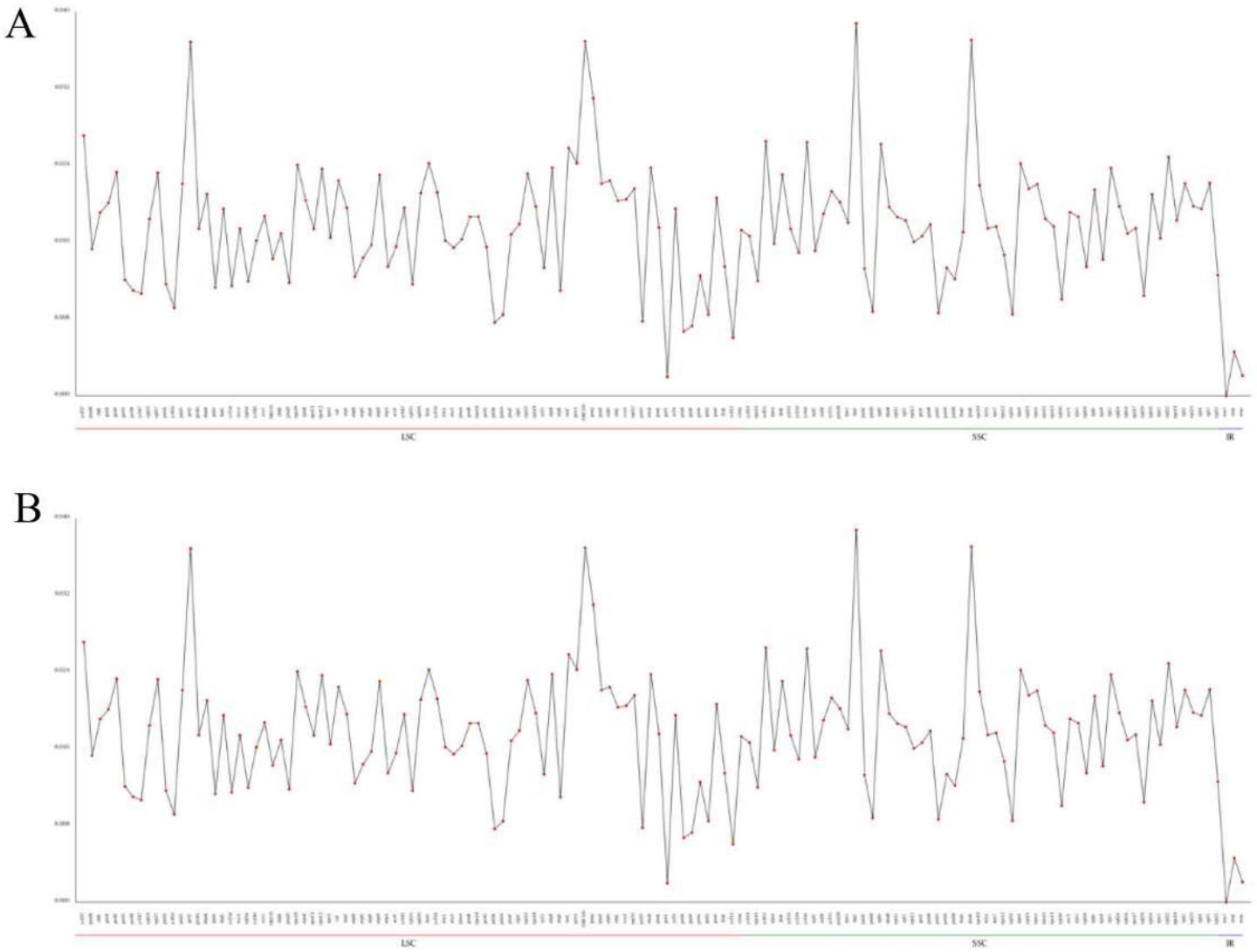


Figure 5. Line graph of gene Pi value. 5-A: Zhejiang Dongtou; P-*S. fusiforme*; 5-B: W-*S. fusiforme*. The horizontal axis represents gene names, and the vertical axis represents Pi values.

evolution of algae and plants (Raval et al. 2024). In the present study, the photosynthesis-related genes identified were mainly associated with core functional modules, including photosystem I, photosystem II, the cytochrome b6/f complex, ATP synthase, the large subunit of Rubisco, and chlorophyll reductase. These findings indicate that the chloroplast genome of *S. fusiforme* retains the core genes required to sustain the light reactions and part of carbon assimilation. Notably, this pattern, namely, the integrity and stability of core photosynthetic genes, is not unique to *S. fusiforme* but appears to be common among brown algae. The chloroplast genome of *Undaria pinnatifida* likewise contains 139 protein-coding genes, 28 tRNA genes, and 6 rRNA genes, and its gene content is highly similar to that of *Saccharina japonica* (Wang et al. 2013; Zhang et al. 2015b). Similarly, the chloroplast genome of *Costaria costata* retains a basic set of functional genes comparable to those of *S. japonica* and *U. pinnatifida* (Zhang et al. 2015a). In this study, P-*S. fusiforme* and W-*S. fusiforme* were identical in both the number and categories of photosynthesis-related genes, suggesting that no substantial divergence has occurred in the core photosynthetic functions of the chloroplast genome between the two lineages. No genes encoding NADH dehydrogenase were detected in the chloroplast genome of *S. fusiforme*. Previous reviews have shown that NADH dehydrogenase genes are absent from plastid DNA in the vast majority of algae, whereas they are more commonly retained in land plants. In addition, the NDH complex is thought to function primarily in photosynthetic responses to fluctuating light and stress conditions in terrestrial environments (Liu et al. 2024). Compared with the photosynthesis-related modules, genes involved in self-replication were likewise highly conserved. In the present study, 75 genes associated with self-replication, including numerous tRNA genes, ribosomal protein genes, RNA polymerase genes, and rRNA genes, together constituted the chloroplast genetic system. Previous studies have shown that this class of genes is typically among the most stable core components

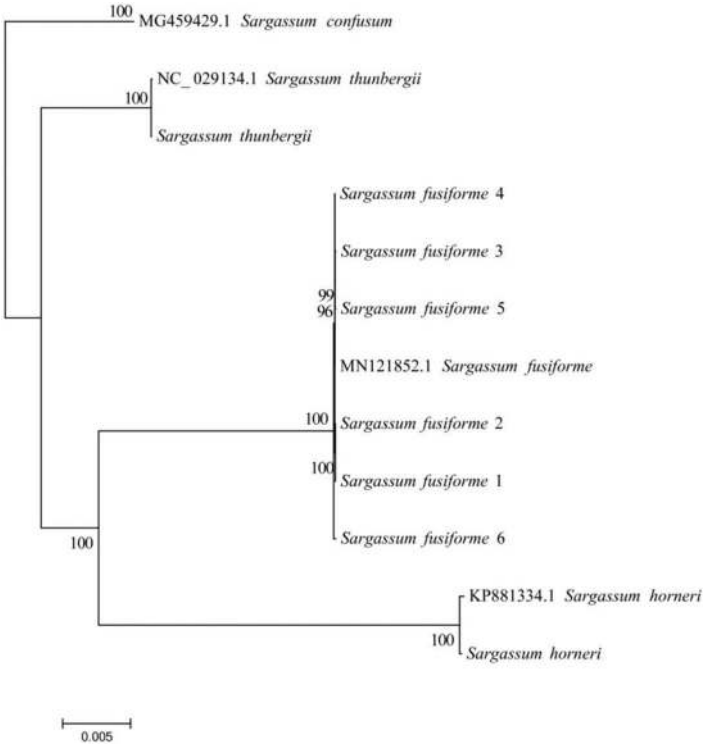


Figure 6. Evolutionary analysis of chloroplast system. The numbers displayed on the branch indicate confidence level.

of brown algal plastids. For example, in a comparative analysis of six *Sargassum* species, coding regions were also highly conserved, such that only a small number of SNPs, rather than large-scale gene differences, could be used to develop species-specific molecular markers. Together, these findings indicate that the chloroplast transcription–translation system in brown algae is subject to strong purifying selection (Martín et al. 2010). Overall, the chloroplast genomes of different *S. fusiforme* lineages exhibited marked functional conservation, with no evident divergence in either the core photosynthetic modules or the self-replication system.

Long repeat sequence analysis and simple repeat sequence analysis

The identification of 35 long repeat sequences (including forward repeats, palindromic repeats, and reverse repeats) in the cpDNA of *S. fusiforme* highlights their significant roles in genome stability, gene regulation, and evolutionary processes. However, the absence of a 38-bp forward repeat in *S. fusiforme* 6 suggests the presence of intraspecific genetic diversity, which may potentially affect local genome stability. Compared to previous studies, the repeat sequence patterns in the cpDNA of *S. fusiforme* show similarities with those of other *Sargassum* species, indicating a high degree of conservation in these sequences during the evolution of *Sargassum* (Yan et al. 2022). Repeat sequences, including SSRs, have been widely utilized in taxonomic analysis and phylogenetic studies, as well as serving as molecular markers in comparative genomics (Liu and Pang 2016). In this study, five types of repeat sequences were identified in the cpDNA. The most abundant repeat sequences identified were single-nucleotide SSRs composed of adenine (A) or thymine (T). Additionally, only a small number of long repeat sequences were detected, and the majority of these repeat sequences exhibited a high AT content. The higher incidence of A/T repeats may be attributed to the fact that A/T strands separate more easily than G/C strands, which increases the likelihood of slipped-strand mispairing (Liu et al. 2022). Furthermore, it was observed that the majority of repeat sequences are located in noncoding regions, a finding that aligns with patterns seen in higher plants and other algal species (Lin et al. 2022).

Codon usage frequency

Codon usage plays a significant role in gene expression levels and translation efficiency (Parvathy et al. 2022). Codon usage provides a valuable mechanism for evolution through molecular-level regulation and mutation (Wang et al. 2022). In chloroplast genomes, AAA is the most frequently used synonymous codon across species (Chen et al. 2023). Additionally, the codons UUA for leucine (Leu), GUU for valine (Val), and AGA for arginine (Arg) exhibit higher RSCU values (> 2), indicating that these synonymous codons are used more frequently than expected. The synonymous codon usage patterns in the cpDNA of *S. fusiforme* are relatively stable, consistent with findings from cpDNA studies of *Arabidopsis thaliana*, *Oryza sativa*, and other plants. This stability reflects a conservation of codon usage during evolution (Chiapello et al. 1998; Chakraborty et al. 2020). These stable codon usage patterns can be instrumental in future studies on the evolution of the *Sargassum* family, providing a crucial basis for further research into gene expression regulation and evolutionary dynamics.

IR shrinkage and expansion analysis

In this study subtle boundary variation was observed. Changes in chloroplast genome length are typically first manifested as expansions and contractions at the IR/SC boundaries, which are considered important mechanisms underlying plastome size variation and local sequence divergence. At the same time, what is commonly observed within the same genus is not large-scale rearrangement, but rather slight positional shifts of genes adjacent to the boundaries (Mu et al. 2023). In this context, the uniform LSC length in P-*S. fusiforme* and the slight variation in LSC length in W-*S. fusiforme* suggest that the purified and wild types may have undergone different selective processes acting on chloroplast genes. The purified type appears to have fixed a more stable plastome configuration during long-term artificial seed selection and cultivation, whereas the wild type has retained more subtle boundary variation across heterogeneous intertidal environments (Waheeda et al. 2023). The gene *cbbX*, located near the LSC/IRb boundary, is a red-type Rubisco activase-associated factor involved in Rubisco activation and the regulation of carbon fixation (Yang et al. 2015). Another boundary-proximal gene, *ycf37*, is associated with the assembly and stability of photosystem I (Wilde et al. 2001; Wang et al. 2024). Therefore, although

the six *S. fusiforme* samples showed slight shifts in boundary coordinates, the composition and orientation of genes adjacent to these boundaries remained consistent, indicating that these regions exhibit a degree of structural variation while still being constrained by the conservation of core functions, including carbon assimilation, chloroplast translation, and photosystem stability. Niche differentiation may maintain microstructural diversity in wild populations, whereas purification under cultivation is more likely to select for and fix boundary configurations and local gene sequences better suited to artificial environments (Amit et al. 2002). These features indicate that the cpDNA of P-*S. fusiforme* and W-*S. fusiforme* is highly conserved overall.

Nucleotide diversity (Pi) analysis

Nucleotide diversity reflects sequence variation in studies of genetic diversity (Zhang et al. 2023). The mean Pi value between the cpDNAs of P-*S. fusiforme* and W-*S. fusiforme* was 0.0176, indicating a high level of conservation between the two taxa. The mean Pi values for the LSC, SSC and IR regions were 0.0173, 0.0182 and 0.002, respectively, indicating that the IR region is more conserved than the other two regions, consistent with Pi patterns reported for a range of plants and algae (Cauz-Santos et al. 2025). These results suggest that the two ecological forms of *S. fusiforme* share a high degree of genetic similarity at the genomic level. Five regions with Pi values greater than 0.03 were further identified. In the cpDNA of P-*S. fusiforme*, these regions included *psaJ*, *petJ*, *ORF501*, *thiS* and *psaE*, whereas in W-*S. fusiforme* they included *ycf33*, *ORF501*, *petJ*, *psaJ*, and *ycf39*. The lineage-specific hypervariability of *thiS* and *psaE* in P-*S. fusiforme* may be related to their relatively less central functions compared with genes forming the core photosynthetic backbone of the chloroplast. Specifically, *psaE* encodes a peripheral small subunit of photosystem I that is mainly involved in acceptor-side stabilization and ferredoxin docking. By contrast, *thiS* encodes a sulfur carrier protein involved in thiazole-ring biosynthesis during thiamine synthesis and is therefore more closely associated with auxiliary metabolic functions. In addition, because both *thiS* and *psaE* are relatively short genes, even a small number of lineage-specific SNPs could substantially elevate their Pi values. Their apparent hypervariability may therefore more likely reflect the accumulation of local microvariation in purified materials (Dorrestein et al. 2004). These genes participate in metabolic pathways that may affect the quality and yield of *S. fusiforme* (Zhu et al. 2024b). By contrast, the functions of *ycf33*, *ORF501*, and *ycf39* in the cpDNA of W-*S. fusiforme* remain unclear, although their high mutation levels may be influenced by environmental factors. These highly variable regions not only provide clues to the evolutionary potential of *S. fusiforme*, but also represent valuable resources for interspecific phylogenetic studies.

Phylogenetic analysis

With continued advances in sequencing technologies, cpDNA has become increasingly important in phylogenetic and taxonomic studies at lower taxonomic levels. To date, thousands of chloroplast genomes have been sequenced, greatly reducing the long-standing reliance of phylogenetic studies on short sequence fragments (Parks et al. 2009).

Complete chloroplast genome sequences in plants provide richer information for phylogenetic and population-based studies, thereby improving the efficiency of species identification. Phylogenomics has become increasingly favored and has recently been widely discussed and applied. The variation contained in cpDNA provides abundant phylogenetic information for a deeper resolution of evolutionary relationships among species. At present, no study has reported the phylogenetic relationship between W-*S. fusiforme* and P-*S. fusiforme*. In this study, analysis of the available cpDNA data from nine *Sargassum* taxa resolved four clades. Among them, *S. thunbergii* showed a close relationship with *S. fusiforme*, consistent with the findings of Huang et al. (Huang et al. 2017). *Sargassum thunbergii* and *S. fusiforme* exhibit high similarity at both morphological and molecular levels. In contrast, *S. confusum* formed a distinct lineage with substantial divergence from *S. fusiforme*, although it remained within the same genus. This result is consistent with the study of Li et al. (2021), indicating that *S. confusum* occupies a unique evolutionary position within the *Sargassaceae*. The six *S. fusiforme* lineages formed a highly supported clade, indicating a high degree of genetic similarity at the genomic level. These findings provide an important basis for further understanding the evolutionary relationships and species divergence within the *Sargassaceae*.

CONCLUSION

In this study, second-generation sequencing technology was used to assemble and annotate the cpDNAs of six *S. fusiforme* lineages. Comparative genomic analysis showed that the cpDNA of W-*S. fusiforme* was slightly smaller than that of P-*S. fusiforme*, but no nuclear gene loss was detected, and the difference in genome size was associated with expansion and contraction of the IR regions. A total of 173 shared genes were annotated in the cpDNAs of the two ecological forms of *S. fusiforme*, with sequence similarity reaching 99.99%, indicating a close phylogenetic relationship and no evident fragment loss.

Analysis of simple sequence repeats (SSRs) and codon usage revealed that A/T mononucleotide SSRs were the most abundant and were predominantly located in noncoding regions, a pattern similar to that reported in higher plants and in some species of *Chlorophyta*, *Fucales* (Phaeophyceae), and *Halymeniales* (Rhodophyta). Differences between P-*S. fusiforme* and W-*S. fusiforme* in the IR region were mainly reflected in the positioning of the *cbbX* gene at the LSC/IRb boundary, whereas the other regions remained highly similar. Nucleotide diversity (Pi) analysis of the cpDNAs of P-*S. fusiforme* and W-*S. fusiforme* showed a high degree of genetic similarity between them at the genomic level. The IR region was the most conserved, whereas the relatively high Pi values in the LSC and SSC regions, together with highly variable loci such as *ycf33*, *ORF501*, *petJ*, *psaJ*, *ycf39*, *thiS*, and *psaE*, may reflect the effects of environmental adaptation and artificial selection. These findings provide a useful reference for evolutionary studies of the *Sargassaceae*. Differences in gene structure and noncoding regions

between the P-*S. fusiforme* and W-*S. fusiforme* lineages may also provide additional information for genetic breeding. Sequencing results indicated a high degree of convergence in the cpDNAs of P-*S. fusiforme* and W-*S. fusiforme*.

This study provides a foundation for understanding the genetic architecture, evolutionary features, and environmental adaptation of the cpDNAs of W-*S. fusiforme* and P-*S. fusiforme*, and offers a useful resource for future research on *S. fusiforme* germplasm identification, genetic stability assessment of new varieties and molecular breeding.

ACKNOWLEDGEMENTS

This work was supported by the algae industry technology system of Ministry of Agriculture and Rural Affairs of the People's Republic of China (CARS-50); Science and Technology Innovation Environment Construction Plan Project of Dongtou District, Wenzhou City N2023Y04; Science and Technology Planning Project of Dongtou District, Wenzhou City (N2024Y08); and Heilongjiang Provincial Undergraduate Colleges and Universities Fundamental Research Business Funding Project (145309512).

LITERATURE CITED

Amit L, Lilya V, Rachel N. 2002. The stable assembly of newly synthesized PsaE into the photosystem I complex occurring via the exchange mechanism is facilitated by electrostatic interactions. Biochemistry 41:11192–11199.

Cauz-Santos LA, da Costa ZP, Sader MA, van den Berg C, Vieira MLC. 2025. Chloroplast genomic insights into adaptive evolution and rapid radiation in the genus *Passiflora* (Passifloraceae). BMC Plant Biol. 25:192–211.

Chakraborty S, Yengkhom S, Uddin A. 2020. Analysis of codon usage bias of chloroplast genes in *Oryza* species: Codon usage of chloroplast genes in *Oryza* species. Planta 252:1–20.

Chen J, Ma W, Hu X, Zhou K. 2023. Synonymous codon usage bias in the chloroplast genomes of 13 oil-tea *Camellia* samples from south China. Forests 14:794–805.

Chiapello H, Lisacek F, Caboche M, Hénaut A. 1998. Codon usage and gene function are related in sequences of *Arabidopsis thaliana*. Gene 209:GC1–GC38.

Dorrestein P, Zhai H, Mclafferty F, Begley T. 2004. The biosynthesis of the thiazole phosphate moiety of thiaminthe sulfur transfer mediated by the sulfur carrier protein this. Chem. Biol. 11:1373–1381.

Falcón LI, Magallón S, Castillo A. 2010. Dating the cyanobacterial ancestor of the chloroplast. ISME J. 4:777–783.

Furan M, Genli G. 2024. An examination of the significance of chloroplast genome analysis in the study of plant evolution and phylogenetics. IJANSER 8:188–192.

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.

Hagopian JC, Reis M, Kitajima JP, Bhattacharya D, De Oliveira MC. 2004. 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. J. Mol. Evol. 59:464–477.

Hong M, Ma Z, Wang X, Shen Y, Wu M, Chen B, Zhang T. 2021. Effects of light intensity and ammonium stress on photosynthesis in *Sargassum fusiforme* seedlings. Chemosphere 273:128605.

Huang C, Sun Z, Gao D, Yao J, Hu Z, Li Y, Wang Y, Xu K, Chen W. 2017. Molecular analysis of *Sargassum* from the northern China seas. Phytotaxa 319:71–83.

Johnson I, Kumar M. 2022. Algal-based biomaterials for environmental remediation of heavy metals, pp. 157–184. In: H. H. Ngo, W. Guo, A. Pandey, J.-S. Chang, and D.-J. Lee (eds.). Algae-Based Biomaterials for Sustainable Development: Biomedical, Environmental Remediation and Sustainability Assessment. Elsevier, Netherlands.

Johnston IG. 2024. The nitroplast and its relatives support a universal model of features predicting gene retention in endosymbiont and organelle genomes. GBE 16:132–145.

Lemieux C, Otis C, Turmel M. 2016. Comparative chloroplast genome analyses of streptophyte green algae uncover major structural alterations in the Klebsormidiophyceae, Coleochaetophyceae and Zygnematophyceae. Front. Plant Sci. 7:697–721.

Li H, Guo Q, Xu L, et al. 2023a. CPJSdraw: Analysis and visualization of junction sites of chloroplast genomes. PeerJ 11:e15326.

Li J, Wang Z, Chen J, et al. 2023b. Protective effect of *Sargassum fusiforme* polysaccharides on H_2O_2-induced injury in LO2 cells. Front. Mar. Sci. 9:1098758.

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.

Li Y, Zhang LN, Wang TX, Zhang C, Wang R, Zhang D, Xie Y, Zhou N, Wang W, Zhang H, Hu B, Li W, Zhao Q, Wang L, Wu X. 2022. The complete chloroplast genome sequences of three lilies: Genome structure, comparative genomic and phylogenetic analyses. J. Plant Res. 135:723–737.

Ligrone R. 2019. The chloroplast and photosynthetic eukaryotes. pp. 269–310. In: Biological innovations that built the world: A four-billion-year journey through life and earth history, Vol. 1. Springer, New York.

Lin LD, Shang TG, Chen BB, Zhang X, Ma ZL. 2019. Analysis of bioactive substances and mineral elements in fertilized eggs of economic seaweed *Sargassum fusiforme*. J. Oceanol. Limnol. 50:429–436.

Lin LD, Wu M. 2021. Screening of *Sargassum fusiforme* cultivars in Dongtou, Zhejiang Province based on clustering and simplified genome joint analysis. J. Oceanol. Limnol. 52:1224–1235.

Lin P, Yin H, Wang K, Gao H, Liu L. 2022. Comparative genomic analysis uncovers the chloroplast genome variation and phylogenetic relationships of *Camellia* species. Biomolecules 12:1474.

Liu B, Ge J, Zhu H, Li S, Dong X, Liu G, Hu Z. 2022. Complete chloroplast genome of the green alga *Chaetophora lobata* (Chlorophyceae, Chlorophyta): morphological features and phylogenetic and comparative analysis. Eur. J. Phycol. 57:181–192.

Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genomics 13:715.

Liu F, Pang S. 2016. Chloroplast genome of *Sargassum horneri* (Sargassaceae, Phaeophyceae): Comparative chloroplast genomics of brown algae. J. Appl. Phycol. 28:1419–1426.

Liu LJ, Lin LD. 2020. Effect of heat stress on *Sargassum fusiforme* leaf metabolome. J. Plant Biol. 63:229–241.

Liu YJ, Zhang TY, Wang QQ, Draisma SGA, Hu ZM. 2024. Comparative structure and evolution of the organellar genomes of *Padina usoehtunii* (Dictyotales) with the brown algal crown radiation clade. BMC Genomics 25:1–15.

Lomartire S, Gonçalves AMM. 2022. An overview of potential seaweed-derived bioactive compounds for pharmaceutical applications. Mar Drugs 20:141.

Martín M, Sabater B. 2010. Plastid ndh genes in plant evolution. Plant Physiol. Biochem. 48:636–645.

Mu Z, Zhang Y, Zhang B, Cheng Y, Shang F, Wang H. 2023. Intraspecific chloroplast genome variation and domestication origins of major cultivars of *Styphnolobium japonicum*. Genes 14:1156.

Nesic A, Meseldzija S, Benavides S, Figueroa FA, Cabrera-Barjas G. 2024. Seaweed as a valuable and sustainable resource for food packaging materials. Foods 13:3212–3234.

Obata T, Fernie AR, Nunes-Nesi A. 2013. The central carbon and energy metabolism of marine diatoms. Metabolites 3:325–346.

Parks M, Cronn R, Liston A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 7:1–17.

Parvathy ST, Udayasuriyan V, Bhadana V. 2022. Codon usage bias. Mol. Biol. Rep. 49:539–565.

Przybyla L, Gilbert LA. 2022. A new era in functional genomics screens. Nat. Rev. Genet. 23:89–103.

Ramdin TM, Ali O, Ramsubhag A. 2024. *Sargassum* spp. and *Ascophyllum nodosum* seaweed extract formulations improve plant growth and alter the bacterial dynamics in hydroponically cultivated lettuce. J. Appl. Phycol. 36:3769–3782.

Raval PK, MacLeod AI, Gould SB. 2024. Complete organellar genomes of six *Sargassum* organellar remodeling during plant evolutionary transitions from algae to angiosperms. PLOS Biol. 22:e3002608.

Tian S, Chen B, Wu M, Cao C, Gu Z. 2023. Are there environmental benefits derived from coastal aquaculture of *Sargassum fusiforme*? Aquaculture 563:738909.

Tsaniya AR, Dewi EN, Anggo AD. 2021. Characteristics of liquid organic fertilizer from different composition types of seaweed between *Gracilaria* sp. and *Sargassum* sp. JPCS IOP 1943:012–071.

Waheeda K, Kitchel H, Wang Q, Chiu PL. 2023. Molecular mechanism of Rubisco activase: Dynamic assembly and Rubisco remodeling. Front. Mol. Biosci. 10:1125922.

Wang J, Xu K, Tang L, Wang Z, Yu X, Wang S, et al. 2024. Insights into the genetic structure and domestication patterns in cultivated populations of *Neopyropia yezoensis*. Aquaculture 592:741229.

Wang Q, Lyu X, Cheng J, Lin Y, Abdoulaye AH. 2022. Codon usage provides insights into the adaptive evolution of mycoviruses in their associated fungi host. Int. J. Mol. Sci. 23:7441–7453.

Wang S, Lin L, Shi Y, Qian W, Li N. 2020. First draft genome assembly of the seaweed *Sargassum fusiforme*. Front. Genet. 11:590065.

Wang X, Shao Z, Fu W, Yao J, Hu Q, Duan D. 2013. Chloroplast genome of one brown seaweed, *Saccharina japonica* (Laminariales, Phaeophyta): Its structural features and phylogenetic analyses with other photosynthetic plastids. Mar. Genomics 10:1–9.

Wilde A, Lünser K, Ossenbühl F, Nickelsen J, Börner T. 2001. Characterization of the cyanobacterial ycf37: Mutation decreases the photosystem I content. Biochem. J. 357:211–216.

Yan L, Wang H, Huang X, Li Y, Yue Y, Wang Z, Tang S. 2022. Chloroplast genomes of genus *Tilia*: Comparative genomics and molecular evolution. Front. Genet. 13:925726.

Yang JB, Tang M, Li HT, Zhang ZR, Li DZ. 2013. Complete chloroplast genome of the genus *Cymbidium*: Lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evol. Biol. 13:1–12.

Yang H, Liu J, Wen X, Lu C. 2015. Molecular mechanism of photosystem I assembly in oxygenic organisms. BBA-Bioenergetics 1847:838–848.

Zhang G, Chen C, Lu W, Li J, Fang T, Yang K. 2023. Genetic diversity and phylogeography of *Taenioides cirratus* in five geographical populations based on mitochondrial COI and Cytb gene sequences. J. Appl. Ichthyol. 2023:4459823.

Zhang L, Wang X, Liu T, Wang H, Wang G, Chi S, et al. 2015a. Complete plastid genome of the brown alga *Costaria costata* (Laminariales, Phaeophyceae). PLoS ONE 10:e0140144.

Zhang L, Wang X, Liu T, Wang G, Chi S, Liu C, Wang H. 2015b. Complete plastid genome sequence of the brown alga *Undaria pinnatifida*. PLoS ONE 10:e0139366.

Zhang Y, Liu J, Mao G, Zuo J. 2021. *Sargassum fusiforme* fucoidan alleviates diet-induced insulin resistance by inhibiting colon-derived ceramide biosynthesis. Food Funct. 12:8440–8453.

Zhang Z, Shi W, Hu S, Mou L, Shi C, Zhu B, Yang J. 2025. Comprehensive chloroplast genome analysis of four *Callitriche* (Plantaginaceae) species for phylogenetic and conservation insights. Horticulturae 11:66–79.

Zhu J, Huang Y, Chai W. 2024a. Decoding the chloroplast genome of *Tetrastigma* (Vitaceae): variations and phylogenetic selection insights. Int. J. Mol. Sci. 25:8290–8316.

Zhu S, Sun S, Zhao W, Yang X, Mao H, Sheng L. 2024b. Utilizing transcriptomics and proteomics to unravel key genes and proteins of *Oryza sativa* seedlings mediated by selenium in response to cadmium stress. BMC Plant Biol. 24:360–382.