MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
KYIV NATIONAL UNIVERSITY OF TECHNOLOGIES AND DESIGN
Faculty of Chemical and Biopharmaceutical Technologies
Department of Biotechnology, Leather and Fur

QUALIFICATION THESIS

on the topic Mitochondrial genomes reveal the differential mechanism between *Gracilariopsis* and *Gracilaria*
First (Bachelor's) level of higher education
Specialty 162 "Biotechnology and Bioengineering"
Educational and professional program "Biotechnology"

Completed: student of group BEBT-21
Deng Li

Scientific supervisor
Tetiana Shcherbatiuk,
Dr. Sc., Professor

Reviewer
Ihor Hretskyi,
Ph.D., Associate Professor

Kyiv 2025

KYIV NATIONAL UNIVERSITY OF TECHNOLOGIES AND DESIGN

Faculty: Chemical and Biopharmaceutical Technologies
Department: Biotechnology, Leather and Fur
First (Bachelor's) level of higher education
Specialty: 162 Biotechnology and Bioengineering
Educational and professional program Biotechnology

APPROVE
Head of Biotechnology, Leather and Fur Department, Professor,
Dr. Sc., Prof.
_____Olena MOKROUSOVA
«___»_____2025

ASSIGNMENTS
FOR THE QUALIFICATION THESIS
Deng Li

1. Thesis topic **Mitochondrial genomes reveal the differential mechanism between *Gracilariopsis* and *Gracilaria***

Scientific supervisor Dr. Sc., Prof. Tetiana Shcherbatiuk

approved by the order of KNUTD "05" March 2025, № 50-уч

2. Initial data for work: assignments for qualification thesis, scientific literature on the topic of qualification thesis, materials of Pre-graduation practice

3. Content of the thesis (list of questions to be developed): literature review; object, purpose, and methods of the study; experimental part; conclusions

4. Date of issuance of the assignments 05.03.2025

WORK CALENDAR

№	The name of the stages of the qualification thesis	Terms of performance of stage	Note on performance
1	Introduction	until 11 April 2025	
2	Chapter 1. Literature review	until 20 April 2025	
3	Chapter 2. Object, purpose, and methods of the study	until 30 April 2025	
4	Chapter 3. Experimental part	until 11 May 2025	
5	Conclusions	until 15 May 2025	
6	Draw up a bachelor's thesis (final version)	until 25 May 2025	
7	Submission of qualification work to the supervisor for feedback	until 27 May 2025	
8	Submission of bachelor's thesis to the department for review (14 days before the defense)	28 May 2025	
9	Checking the bachelor's thesis for signs of plagiarism (10 days before the defense)	01 June 2025	Similarity coefficient ____% Citation rate ____%
10	Submission of bachelor's thesis for approval by the head of the department (from 7 days before the defense)	04 June 2025	

I am familiar with the task:

Student _____ Deng Li

Scientific supervisor _____ Tetiana SHCHERBATIUK

Abstract

Deng Li. Mitochondrial genomes reveal the differential mechanism between *Gracilariopsis* and *Gracilaria.* Manuscript.

Qualification thesis, specialty 162 "Biotechnology and Bioengineering". Kyiv national university of technologies and design, Kyiv, 2025.

Studies have shown that all five species of the *Gracilaria* genus have 25 protein-coding genes. The number of proteins encoded by the mitochondrial genomes of different species in the genus *Gracilariopsis* is 23 to 27. The differences among them lie in the quantity or type of orf. Regarding the number of coding genes, species of the *Gracilaria* genus are more conserved compared to those of the *Gracilariopsis* genus. The mitochondrial genomes of the 10 species selected in this study contain most of the two rRNA genes (*LSU, SSU*). The mitochondrial genome of only *Gp. lemaneiformis* in the genus *Gracilariopsis* contains three rRNA genes, and an additional *rrn*5 gene is encoded.

The collinearity analysis performed using Mauve indicates that tRNA deletions, duplexes, or rearrangements often occur in the *trn*N-*trn*A rich region. The *trn*H gene in the mitochondrial genome of the genus *Gracilariopsis* has undergone rearrangement, which can serve as evidence to distinguish *Gracilaria* from the genus *Gracilariopsis*.

There are differences in the relevant parameters of mitochondrial genomic codons between the two. The ENC range and average value of the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* are different. The average values of CBI and Fop are different, and there are also differences in the RSCU values. The preference for the use of mitochondrial genome codons in both is relatively weak, and compared with the genus *Gracilariopsis*, the preference for codons in the mitochondrial genome of the genus *Gracilaria* is even weaker.

The phylogenetic trees of *Gracilaria* and *Gracilariopsis* were reconstructed using two methods: Bayesian Inference (BI) and maximum likelihood (ML). The research shows that *Gracilaria* and *Gracilariopsis* are divided into two subbranches, with

Gracilariopsis being the first to separate. *Gracilaria* has a higher evolutionary status, supporting *Gracilariopsis* to become an independent genus.

Key words: Gracilaria; Gracilariopsis; Mitochondrial genome; Comparative Genomics; phylogenetic analysis

TABLE OF CONTENTS

INTRODUCTION

Red algae play a significant role in Marine ecosystems and the bioeconomy. *Gracilaria* and *Gracilariopsis*, as important groups of red algae, the study of their mitochondrial genomes is of crucial significance for revealing species evolutionary relationships, genetic mechanisms, and systematic classification.

Using bioinformatics methods, this study on genome structure, gene composition and codon unrael the in-depth analysis, and construct phylogenetic tree to explore its evolutionary relationships. The research found that there are differences in the structure and genetic composition of the mitochondrial genomes of the two types of algae. The codon preference is affected by multiple factors, and the phylogenetic tree clearly shows their genetic relationship.

This paper in the introduction of research background and significance, then in order to carry out the mitochondrial genome comparison, codon preferences research, system development analysis, finally summarizes achievements, aimed at the red algae provide theoretical basis for related research.

CHAPTER I
LITERATURE REVIEW

1.1 OVERVIEW OF RED ALGAE

Rhodophyta, as a phylum of algae, is not only one of the oldest groups of algae but also one of the groups with the largest number of species. According to the latest classification system, the Red Algae phylum is divided into seven classes. It includes the Florideophyceae, Bangiophyceae, Compsopogonophyceae, Porphyridiophyceae, Rhodellophyceae[1] and Myriophyllum Stylonematophyceae and Cyanidiphyceae [2].The vast majority of red algae grow in the ocean, distributed from polar regions to tropical waters. Most of them grow in intertidal zones and shallow sea areas, while some can survive in deeper waters. About 200 species are found in freshwater and mostly live in environments such as rapids. The reproductive methods include asexual reproduction and sexual reproduction. Asexual reproduction occurs through various flagelleless conidia, such as monospores and tetrasidia[3]. Sexual reproduction is dioecious. The male organ is the spermatosac, which produces immotile sperm without flagella. The female organ is the fruiting cell, which has fertilization filaments. After fertilization, some fruiting cells immediately undergo meiosis to produce fruiting spores that develop into gametophytes, while others do not undergo meiosis to develop fruiting sporophytes, which parasitize on gametophytes. Red algae have significant edible and economic value. Nori is rich in protein. China is the country with the largest nori cultivation scale in the world and exports a large amount to Japan. In addition, some red algae such as Gelidium and *Gracilaria* can be used as raw materials for making Agar. AGAR is edible, but its main application is in medicine or scientific research as a culture medium for microorganisms, etc [4].

1.2 BIOLOGICAL CHARACTERISTICS OF *Gracilaria* AND *Gracilariopsis*

1.2.1 DISTRIBUTION AND MORPHOLOGICAL CHARACTERISTICS OF *Gracilaria* AND *Gracilariopsis*

The genera *Gracilaria* and *Gracilariopsis* belong to the phylum Rhodophyta, order *Gracilaria*les and family *Gracilaria*ceae in the natural classification system. The *Gracilaria* species of the genus *Gracilaria* have diverse morphologies, ranging from solitary to clustered. Their color is mostly light brown to dark brown, but there are also light purplish-brown or yellowish-green ones, and they are nearly soft and bony. It is linear and cylindrical in shape, usually with a main trunk reaching the top. The branches are irregularly alternate or deviated, and the base is slightly contracted. The fixator is disc-shaped. It grows on rocks from the mid-tidal zone to the subtidal zone, and thrives in calm inner bays and fertile water areas. In artificial cultivation, it is usually found on attachments such as wood and bamboo. It is widely distributed around the world, including along the coast of China.

Gracilariopsis is reddish brown algae body, water loss after dark, RuanGuZhi, cylindrical, branch is more, generally for three, has four branches, more mature algal branch elongation and upper branches often naked into a whip. It is a temperate seaweed that prefers to grow in inner bays with fresh water inflow and fertile water quality. It thrives especially in bays where the wind and waves are relatively calm, the water flow is unobstructed, the terrain is flat, and the water quality is relatively clear. It grows in low-tide zones that are sunny, have clean seawater, and have sandy bottoms. It is mainly distributed in the northern regions of our country. With the successful breeding of this kind of strain with high temperature resistance, high yield and high AGAR content, its breeding range has gradually expanded to Fujian, Guangdong and Hainan regions, becoming the most important AGAR farmed seaweed in China.

1.2.2 LIFE HISTORY AND ECONOMIC VALUE OF *Gracilaria* AND *Gracilariopsis*

In the life history of *Gracilaria*, there is a phenomenon of generational alternation. Generally, there are three planting objects in the life history: sporophyte, gametophyte and fruiting sporophyte, and all of them can live independently. After the sporophyte matures, cortical cells will continuously develop into sporangia. Observation reveals that the sporangia contain spores. After the spores in the sporangia are released, they gradually develop into gametophytes. The gametophyte includes two types: the female gametophyte and the male gametophyte. When the male gametophyte develops into a sperm and is further released, and eventually combines and fertilizes with the female gametophyte, a new sporophyte will develop.

Mature male gametophyte of the genus *Gracilariopsis* produce antheridium, release sperm. The female gametophyte produces fruiting cells. When the sperm reach the female gametophyte with the water flow and combine with the fruiting cells, the fertilization process is completed. After fertilization, the fruiting cell develops into a fruiting sporophyte, which parasitizes on the female gametophyte. It is generally a tiny spherical or oval structure, has no independent photosynthesis and nutrient absorption capacity, and relies on the female gametophyte for nutrition. After the fruit sporophyte matures, it undergoes meiosis to produce fruit spores. After the release of fruit spores, they germinate in a suitable environment to form diploid sporophytes, which are somewhat similar in appearance to gametophytes, but the number of chromosomes in the cells is different, and there may also be differences in cell structure, etc. After the sporophyte matures, quadrangle sporangia form on its surface, and each quadrangle sporangia undergoes meiosis to produce four quadrangle spores. Tetrad spores are released into seawater and spread with the current. When they encounter suitable substrates and environmental conditions, they germinate to form new male and female gametophytes, thus completing the cycle of their life history.

Gracilaria and *Gracilariopsis*, as important members of the Rhodophyta phylum, play an indispensable role in Marine ecosystems. This type of large economic seaweed not only plays a key role in the ecosystem but also attracts much attention due to its economic value. As primary producers in the ocean, they fix a large amount of carbon

through photosynthesis, providing a rich energy source for the Marine food chain. The *Gracilaria* genus of seaweed is rich in species, widely distributed in China, with a short growth cycle and mature artificial cultivation techniques. It is an important economic algae in China and one of the main sources of AGAR extraction in China. *Gracilariopsis* lemaneiformis is the second most produced cultivated seaweed in China after kelp. *Gracilariopsis* lemaneiformis is not only an important source for the production of AGAR, but also can be used as feed for aquatic animals such as abalones. Meanwhile, it has a strong ability to absorb nitrogen and phosphorus, which can alleviate the problem of eutrophication of seawater. It has relatively high economic and ecological value. In addition, the seaweeds of the *Gracilaria* and *Gracilariopsis* genera have also been extensively studied and developed due to their unique bioactive substances. These substances include polysaccharides, proteins, fatty acids, pigments and other secondary metabolites, and have broad application prospects, covering everything from the food industry to the pharmaceutical field.

1.3 RESEARCH PROGRESS ON SYSTEMATIC CLASSIFICATION OF JIANGLIACEAE

The *Gracilaria*ceae family belongs to the phylum Redophyta, class Euredophyta, and order Gracilariformes in biological taxonomy. It is a type of algae that prefers warm waters. Its distribution range covers tropical, subtropical and temperate sea areas all over the world. Numerous studies have shown that *Gracilaria*ceae seaweeds have a wide range of uses. This type of seaweed plays a significant role in fields such as aquaculture, pharmaceutical research and development, food processing, and Marine ecological restoration. It is an extremely important large economic algae in China.

Researchers such as Zhao Xiabo analyzed the molecular sequences of the 18S rRNA gene, cox2-3 septum and RUBISCO septum of several common *Gracilaria* seagrasses along the coast of China, and conducted a study on the molecular phylogenetic relationship in combination with the existing relevant data of GenBank,

providing new evidence for the phylogenetic evolution and taxonomic status of *Gracilaria*. Currently, the morphological taxonomy of the family *Gracilaria*ceae is relatively well-studied. However, most species of this family are highly influenced by their growth environment, leading to significant morphological variations. Additionally, the lack of reproductive organs in some species poses numerous challenges for traditional morphological studies. Furthermore, there is still controversy regarding the relationships between the genera of the *Gracilaria*ceae family, particularly the relationship between *Gracilaria* and the genera Eustacyria (also known as the genus *Gracilariopsis*) and Polykaryon.

A total of 86 species and varieties of *Gracilaria* have been recorded in China, including 55 endemic species. Among them, 42 species and varieties have been described in detail, and the vast majority of the work is based on the results of morphological studies. This figure is quite prominent internationally, and the existence of many species is worth further verification.

In 2023, the research team of Algal Physiology and Precision Breeding at the Institute of Oceanology, Chinese Academy of Sciences, made new progress in the study of species diversity in the family Hydractiniidae. They discovered new species such as Zeng's Hydractinius and, based on morphological and molecular evidence, including rbcL/coi gene sequences, constructed a DNA sequence dataset for Hydractiniidae species in China. This work revealed the phylogenetic relationships of species in the genera Hydractinius and Longshuca, clarifying the systematic position of Chinese Hydractiniidae species within the global Hydractiniidae group.

In short, there are still many problems in the classification of the family Clavulaceae. Although molecular biology has made some progress, most studies are limited to sequences such as ITS and rbcL, and the complex classification relationship between algae is still unclear. Therefore, a new classification method is needed for research.

1.4 MITOCHONDRIA AND MITOCHONDRIAL GENOME

1.4.1 STRUCTURE AND FUNCTION OF MITOCHONDRIA

Mitochondria are semi-autonomous organelles distributed in eukaryotic cells. In cells, mitochondria undertake the function of oxidative phosphorylation to synthesize ATP. All kinds of life activities within cells cannot do without mitochondria, which are known as the energy factories of cells. Mitochondria are important organelles that mainly play a role in the production of biological energy. The mitochondrial genome has been widely applied in numerous research fields such as insect phylogenetics, molecular evolution, population genetics and systematic geography. The mitochondrial genome is a double-stranded cyclic DNA. As a molecular marker, it has the following advantages: relatively stable gene composition, high homology, easy analysis, small genome size, low sequencing cost and easy access/

Mitochondria (Mitochondrion) membrane by parallel to each other, both inside and outside two layer units nested form of closed cystic structure. From the inside out, it can be divided into four functional areas: mitochondrial matrix, inner mitochondrial membrane, mitochondrial intermembrane space, and outer mitochondrial membrane. A large number of respiratory-related enzymes, proteins and lipids exist in the mitochondrial matrix and are involved in metabolic pathways and energy conversions such as the tricarboxylic acid cycle, fatty acid oxidation and amino acid degradation, etc. The inner mitochondrial membrane is a unit membrane that encloses plastids and is located on the inner side of the outer mitochondrial membrane. The cristae is a structure formed by the inward folding of the inner mitochondrial membrane, which greatly increases the surface area of the inner membrane. Enzymes and protein molecules distributed on the membrane synthesize ATP during oxidative phosphorylation and play an important role in processes such as mitochondrial division regulation. The inner surface of the cristae has uniformly distributed thylakoids and thylakoids containing ATPase. ATP can be synthesized using the energy produced by the respiratory chain. The membrane space is located between the two membranes of mitochondria and is filled with amorphous liquid, containing many substrate molecules, soluble enzymes

and cofactors for biochemical reactions, etc. The outer mitochondrial membrane is smooth and rich in pore proteins, with strong permeability, and plays a major role in the organelle boundary membrane and signal transduction processes. Mitochondria are involved in processes such as cell information transmission, cell differentiation, and cell apoptosis in addition to cell energy supply, and can regulate cell growth and the cell cycle.

1.4.2 ORIGIN OF MITOCHONDRIA

The endosymbiotic origin theory holds that mitochondria and chloroplasts originated respectively from aerobic respiration bacteria that coexisted with primitive eukaryotic cells and photosynthetic autotrophic cyanobacteria. Mitochondria as breathing organelles of eukaryotic, its evolution history is closely combined with the evolution of eukaryotes. Mitochondria oxidize and decompose organic matter into energy for cells through respiration, and this process occurs in the cristae structure formed by the folding of the inner mitochondrial membrane. The presence of cristae greatly increases the inner membrane surface area, providing attachment sites for enzymes and proteins related to respiration, thereby enhancing respiratory efficiency. However, the evolutionary origin of the mitochondrial cristae remains unclear to this day.

However, recent research results to revive about mitochondrial crest endosymbiont before probably originated from point of view. Studies have shown that the mitochondrial cristae may be homologous to the cytoplasmic inner membrane (ICM) used for energy utilization by various α -Proteobacteria. From the perspective of phylogenetics, the mitochondrial contact sites and cristae tissue system (MICOS) that control cristae development evolved from alphaMICOS, which exists only in the Alpha-Proteobacteria phylum. Based on this, ICM is very likely to gradually transform into cristae during the symbiotic origin process within mitochondria[1]. Therefore, based on the existing evidence, it can be inferred that some prokaryotes "phagocytize" α -

Proteobacteria cells and the transfer of most of the genetic material of α -proteobacteria to the nucleus occurred. At the same time, a symbiotic relationship was established among them and they continuously evolved into today's mitochondria.

1.4.3 RESEARCH PROGRESS OF RED ALGAE MITOCHONDRIAL GENOME

The mitochondrial genome of red algae is relatively conservative. The mitochondrial genome of almost all red algae is circular, with compact gene structure, few spacer regions, and multiple overlapping regions. In addition, the coding sequence also shows a considerable degree of conservation [30].

In terms of the genetic code, most red algae use UGA to encode tryptophan. During biological evolution, most red algae possess three mitochondrial genome-encoded succinate dehydrogenase genes (sdh2, sdh3, sdh4). However, some red algae, especially unicellular red algae, may lose one or more of these succinate dehydrogenase genes during evolution. Compared to other algae, red algae have fewer encoded genes, with most lacking the 5S rRNA gene. In terms of A+T content, the mitochondrial genomes of red algae generally have a higher proportion. Currently, the mitochondrial genomes of red algae are widely used in evolutionary biology and related research fields, providing important evidence for further exploring the biological characteristics and genetic information of red algae.

1.5 RESEARCH PROGRESS ON MITOCHONDRIAL GENOME MEASUREMENT METHODS

Mitochondrial genome sequencing methods have been continuously innovated, driving advancements in life sciences research. In the early days, Sanger sequencing was primarily used, which relies on the principle of double-deoxyribonucleotides to terminate the process. The fragments are separated by electrophoresis, allowing for extremely accurate sequence reading with over 99.99% accuracy. This method was first

used to complete the sequencing of the human mitochondrial genome. However, it has low throughput, high costs, and complex operations.

As technology advances, second-generation sequencing (NGS) has emerged, such as Illumina sequencing technology, which fragments DNA for large-scale parallel sequencing, producing massive short-read sequences in one go, and then stitching them together to obtain results. It boasts high throughput and low cost, allowing for the rapid determination of mitochondrial genomes from large numbers of samples, but it faces challenges with short reads and complex regions during assembly.

Now, third-generation sequencing (TGS), like PacBioRS and Nanopore sequencing, can sequence single molecules in real time, read long and long, effectively solve complex structure and variation problems, but the error rate is slightly higher and the cost is relatively expensive.

1.5.1 DOUBLE DEOXY TERMINATION SEQUENCING METHOD

The dideoxy termination sequencing method, also known as Sanger sequencing, is a DNA sequence determination technique invented by Frederick Sanger in 1977. The principle involves using dideoxy nucleotides (ddNTPs) to terminate DNA strand elongation, followed by electrophoresis to separate DNA fragments of different lengths, and reading the DNA sequence through autoradiography. The procedure involves extracting mitochondrial DNA, performing PCR amplification to obtain sufficient target fragments, mixing the amplified products with primers, DNA polymerase, dNTPs, and ddNTPs for sequencing reactions, and then analyzing the sequences through electrophoresis and autoradiography, either manually or using software. This method is commonly used for sequencing specific genes or regions in the mitochondrial genome, such as analyzing mutation sites related to mitochondrial diseases.

1.5.2 DIRECT SEQUENCING OF PCR AMPLIFICATION PRODUCTS

The amplification of mitochondrial genomes using PCR technology mainly involves the following steps: First, sample processing, where samples such as blood or tissue are collected and mitochondrial DNA is extracted using the phenol-chloroform method or reagent kit. Next, based on known mitochondrial genome sequences, primers are designed using specialized software and synthesized by a company. Then, the reaction mixture is prepared on ice, combining template DNA, primers, dNTPs, DNA polymerase, and buffer in the correct proportions. Afterward, PCR reaction conditions are set, starting with high-temperature pre-denaturation, followed by multiple rounds of denaturation, annealing, and extension cycles to allow for extensive DNA amplification. Finally, the product is detected through agarose gel electrophoresis to determine if the amplification was successful.

1.5.3 HIGH-THROUGHPUT SEQUENCING METHODS

High-throughput sequencing, also known as second-generation sequencing technology, is used to determine mitochondrial genomes and primarily consists of four steps: sample processing, library construction, sequencing, and data analysis. First, obtain samples rich in mitochondria, such as animal liver tissue. Use physical, chemical, or enzymatic methods to disrupt the cells, then purify the mitochondria using differential centrifugation. Finally, use a DNA extraction kit to obtain high-purity mitochondrial DNA. Then, the extracted mitochondrial DNA is fragmented by ultrasound or enzymatic digestion, with specific connectors attached to both ends of the fragments to complete the construction of the DNA library. This step enables the DNA fragments to adapt to the sequencing platform. The library is then loaded onto the sequencer. On the sequencing chip, DNA fragments are amplified to form DNA clusters. During sequencing, dNTP is added one by one to the newly synthesized strands. Each time one is added, an optical or electrical signal is released, which is captured and recorded by the sequencer to identify the bases. Finally, the sequencer generates a vast amount of raw data. Quality control is carried out first to filter out low-

quality data. Then, the high-quality data were compared with the reference mitochondrial genome sequences using alignment software, spliced and assembled to obtain the complete mitochondrial genome sequences, and its characteristics and variations were analyzed.

High throughput sequencing method has many advantages, including high flux, a parallel to hundreds of thousands to millions of DNA molecule sequence determined, at the same time test of a large sample or a large number of loci, such as whole genome sequencing, whole transcriptome sequencing, etc.; It is fast. Compared with traditional sequencing technology, it can generate a large amount of data in a short time, shortening the research and diagnosis cycle. For example, the metagenomic detection of pathogenic microorganisms usually yields results within 1-2 days. High accuracy, capable of sequencing a large number of DNA molecules, improving detection accuracy through deep coverage, able to detect low-frequency mutations, and also allowing for multiple rounds of verification and comparison of sequencing results. Rich in information, it can not only obtain DNA sequence information, but also provide information such as gene structure, function, expression level, and methylation status, which is used to study the relationship between gene regulation, genetic variation and diseases.

1.6 THE RESEARCH PURPOSE AND SIGNIFICANCE OF THIS SUBJECT

The *Gracilaria* and *Gracilariopsis* genera of algae hold important positions in the Red Algae phylum. By comparing their mitochondrial genomes, one can gain an in-depth understanding of the phylogenetic relationships within red algae. For a long time, morphological characteristics have been used to classify species and phylogenetic processes in order to understand the evolution of species. However, identifying species and genetic relationships only based on external characteristics faces huge challenges[30].

This research through the *Gracilaria* genera and *Gracilariopsis* mitochondrial genome evolution analysis, comparative analysis and system two genera taxonomic

research is dedicated to provide you with the new basis of molecular level. Through the method of comparative genomics, the structural differences and gene arrangement characteristics of the mitochondrial genomes of the two genera were analyzed. By conducting phylogenetic analysis, a phylogenetic tree based on mitochondrial genomic data was constructed to clarify the phylogenetic relationship between the two genera and provide evolutionary evidence for the classification of the two genera. The research results can supplement the deficiencies of traditional morphological classification, provide key molecular support for the precise classification, resource identification and systematic evolution research of *Gracilaria* and *Gracilariopsis*, and promote the refined development of algal taxonomy.

Summary of the chapter I

1. The background knowledge of red algae research was systematically sorted out. First, introduce the overall situation of red algae, and then focus on the *Gracilaria* and *Gracilariopsis* algae, elaborating in detail on their distribution, morphology, life history and economic value, highlighting their importance.
2. In the section on the research progress of the systematic classification of *Gracilaria*ceae, the existing research achievements and development trends in this field are presented.
3. The mitochondria and mitochondrial genome, mitochondrial deep insight into the structure, function, origin, and red algae and the research status of mitochondrial genome. Meanwhile, the development history of mitochondrial genome determination methods is summarized, ranging from the traditional dideoxy termination sequencing method and direct sequencing of PCR amplification products to advanced high-throughput sequencing methods.
4. Clear the purpose of this topic research is to reveal *Gracilaria* and *Gracilariopsis* mitochondrial genome characteristics and evolutionary relationships of algae research

significance is to provide theory support for the research on algae, classification, evolution, lay a solid theoretical basis for study of subsequent chapters.

Chapter Ⅱ
OBJECT, PURPOSE, AND METHODS OF THE STUDY

2.1 DATA SOURCES

This study selected partial species data of the genus *Gracilaria* from the GenBank database:*Gracilaria chilensis*, *Gracilaria vermiculophylla*, *Gracilaria changii*, *Gracilaria edulis*, and *Gracilaria salicornia*; and partial species data of the genus *Gracilariopsis* from the mitochondrial whole genome:*Gracilariopsis lemaneiformis*, *Gracilariopsis andersonis*, *Gracilariopsis oryzoides*, and *Gracilariopsis chorda*. For specific information, see Table 2-1.

Table 2-1 Data source information

Algal names	Length (bp)	accession number
Gracilaria chilensis	26,898	NC_026831
Gracilaria vermiculophylla	25,971	NC_027064
Gracilaria changii	25,729	NC_034681
Gracilaria edulis	25,708	NC_037889
Gracilaria salicornia	25,272	NC_023784
Gracilariopsis lemaneiformis	25,883	JQ071938
Gracilariopsis andersonis	27,036	NC_014772
Gracilariopsis oryzoides	25,161	NC_014771
Gracilariopsis chorda	26,534	NC_023251

2.2 EXPERIMENTAL METHOD

In this study, Geneious was used to conduct a comparative analysis of the characteristics of the genomic coding genes (tRNA genes, rRNA genes, and protein coding genes) of the selected algal mitochondrial genomes. The specific method is as follows: Download the relevant sequences of Gracilia and *Gracilariopsis* species from the NCBI database, import the sequence data into the Geneious software, and use Geneious' gene prediction tools, such as orf prediction, to determine the open reading box of protein-coding genes, as well as the quantity and location of tRNA genes and rRNA genes.

The visual genomic map of the entire mitochondrial genome was drawn using the online tool OGDRAW. The mitochondrial genome of the above species was analyzed for collinearity using the Mauve plugin in Geneious.

2.3 RESULTS AND ANALYSIS

2.3.1 ANALYSIS OF MITOCHONDRIAL GENOME STRUCTURE CHARACTERISTICS OF JIANGLI AND LONGXU SPECIES

The full length of the mitochondrial genome of *G. chilensis* among the *Gracilaria* species obtained in this study is 26,898 bp (GenBank registration number: NC_026831). The mitochondrial genome encodes 52 genes, including 25 protein-coding genes, 2 rRNA genes and 25 tRNA genes, with a GC content of 27.6%. The full length of the mitochondrial genome of *G. vermiculophylla* is 25,971 bp (GenBank registry number: NC_027064). The mitochondrial genome encodes 48 genes, including 25 protein-coding genes, 2 rRNA genes and 21 tRNA genes, with a GC content of 28.3%. The full length of the mitochondrial genome of *G. changii* is 25,729 bp (GenBank registration number: NC_034681). The mitochondrial genome encodes 50 genes, including 25 protein-coding genes, 2 rRNA genes and 23 tRNA genes, with a GC content of 27.7%. The full length of the mitochondrial genome of *G. edulis* is 25,708 bp (GenBank registry number: NC_037889). The mitochondrial genome encodes 50 genes, including 25 protein-coding genes, 2 rRNA genes and 23 tRNA genes, with a GC content of

25.5%. The mitochondrial genome of *G. salicornia* is 25,272bp in length (GenBank registration number: NC_023784). The mitochondrial genome encodes 46 genes, including 25 protein-coding genes, 2 rRNA genes and 20 tRNA genes, with a GC content of 28.4%. The full length of the mitochondrial genome of *Gp. lemaneiformis*, a species of the genus Lemaneiformis, is 25,883 bp (GenBank registry number: JQ071938). The mitochondrial genome encodes 50 genes, including 26 protein-coding genes, 3 rRNA genes and 21 tRNA genes, with a GC content of 27.5%. The full length of the mitochondrial genome of *Gp. andersonii* is 27,036 bp (GenBank registry number: NC_014772). The mitochondrial genome encodes 47 genes, including 27 protein-coding genes, 2 rRNA genes and 18 tRNA genes, with a GC content of 28.0%. The full length of the mitochondrial genome of *Gp. oryzoides* is 25,161 bp (GenBank registry number: NC_014771). The mitochondrial genome encodes 44 genes, including 23 protein-coding genes, 2 rRNA genes and 19 tRNA genes, with a GC content of 28.1%. The full length of the mitochondrial genome of *Gp. chorda* is 26,534 bp (GenBank registry number: NC_023251). The mitochondrial genome encodes 51 genes, including 26 protein-coding genes, 2 rRNA genes and 23 tRNA genes, with a GC content of 27.6%.

Table 2-2 Source Information of Experimental Samples

sample name	Length (bp)	GC content (%)	protein coding gene	tRNA gene	rRNA gene	orf	Number of introns	Take ATG as the starting codon	Take ATT as the starting codon	Use TAA as the stop codon	Use TAG as the stop codon	Genbank accession number
G. chilensis	26,898	27.6	25	25	2	1	0	24	*rps*11	22	3 (*rps*3 *atp*8 *orf*148)	NC_026831
G. vermiculophylla	25,971	28.3	25	21	2	1	0	25	0	22	3 (*atp*8 *nad*4 *rpl*16)	NC_027064
G. changii	25,729	27.7	25	23	2	1	0	25	0	22	3 (*rpl*20	NC_034681

											*rps*11 *sdh*C)	
G. edulis	25,708	25.5	25	23	2	1	0	25	0	24	1 (*rpl*20)	NC_037889
G. salicornia	25,272	28.4	25	20	2	1	0	25	0	20	5 (*nad2* *nad*4 *rpl*16 *sdh*B *ymf*39)	NC_023784
Gp. lemaneiformis	25,883	27.5	26	21	3	2	1	26	0	22	4 (*rpl*20 *rps*12 *nad*6 *nad*4)	JQ071938
Gp. andersonii	27,036	28.0	27	18	2	3	0	27	0	22	5 (*rpl*16 *rps*12	NC_014772

											*orf*87 *orf*143 *orf*95)	
Gp. oryzoides	25,161	28.1	23	19	2	1	0	23	0	20	3 (*rps*12 *nad*4L *nad*2)	NC_014771
Gp. chorda	26,534	27.6	26	23	2	3	1	26	0	23	3 (*rpl*20 *nad*6 *nad*4)	NC_023251

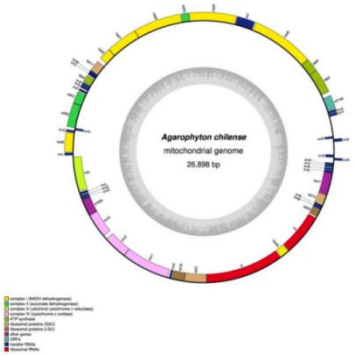


Figure 2-1 Complete mitochondrial genome map of *Gracilaria chilensis*

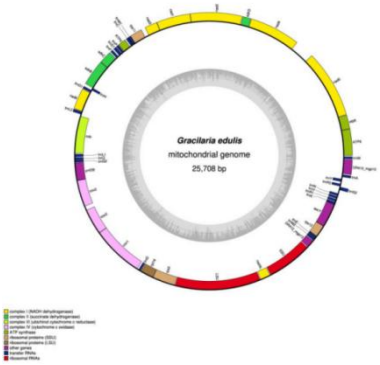


Figure 2-2 Complete mitochondrial genome map of *Gracilaria edulis*

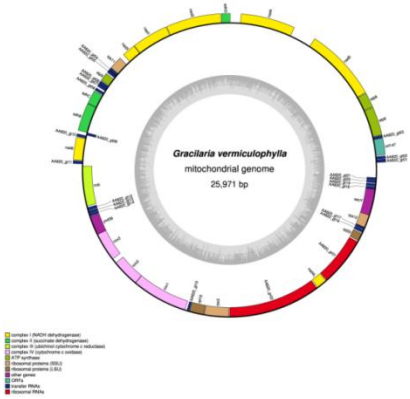


Figure 2-3 Complete mitochondrial genome map of *Gracilaria vermiculophylla*

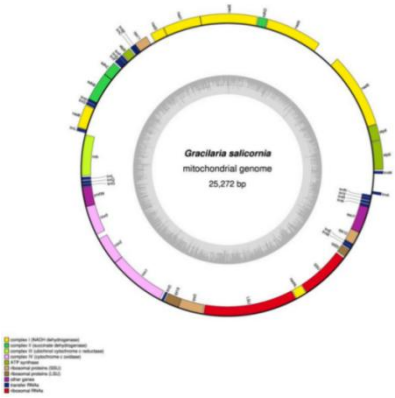


Figure 2-4 Complete mitochondrial genome map of *Gracilaria salicornia*

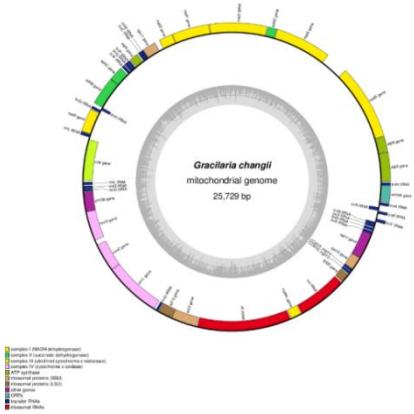


Figure 2-5 Complete mitochondrial genome map of *Gracilaria changii*

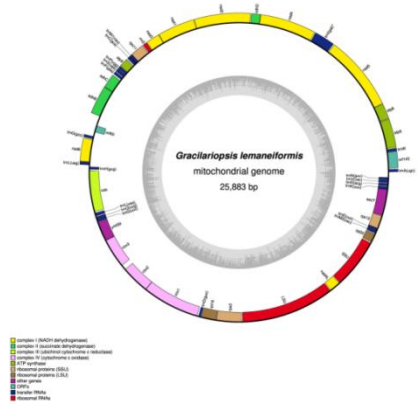


Figure 2-6 Complete mitochondrial genome map of *Gracilariopsis lemaneiformis*

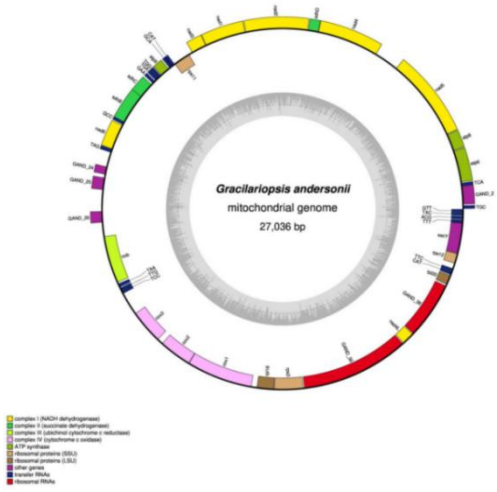


Figure 2-7 Complete mitochondrial genome map of *Gracilariopsis andersonis*

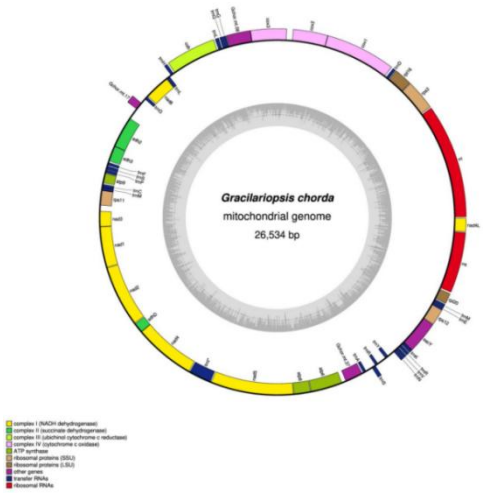


Figure 2-8 Complete mitochondrial genome map of *Gracilariopsis chorda*

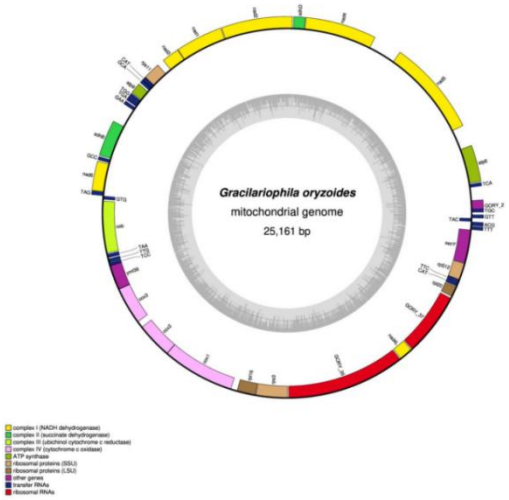


Figure 2-9 Complete mitochondrial genome map of *Gracilariopsis oryzoides*

2.3.2 PROTEIN CODING GENES

In this study, the protein-coding genes of the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* were investigated respectively. The results indicated that each of the five species of *Gracilaria* had 25 protein-coding genes (Table 2-3), mainly including three cytochrome oxidase subunits, namely cox1, cox2, and cox3. One cytochrome b, namely cob; The seven coenzyme dehydrogenase subunits include nad1, nad2, nad3, nad4, nad5, nad6 and nad4L; Five ribosomal proteins, namely rps3, rps11, rps12, rpl16 and rpl20; Four ATP synthase subunits, including atp6, atp8, atp9, ymf39; Three succinate dehydrogenase complexes, namely sdhB, sdhC and sdhD; One membrane protein secY and one orf (orf148 or orf147). Among the genus *Gracilariopsis*,*Gp. lemaneiformis* has 26 protein-coding genes, *Gp. andersonii* has 27, and *Gp. oryzoides* has 23. *Gp. chorda* has 26 protein-coding genes. There are mainly three cytochrome oxidase subunits, namely cox1, cox2 and cox3; One cytochrome b, namely cob; The seven coenzyme dehydrogenase subunits include nad1, nad2, nad3, nad4, nad5, nad6 and nad4L; Five ribosomal proteins, namely rps3, rps11, rps12, rpl16 and rpl20; Four ATP synthase subunits, including atp6, atp8, atp9, ymf39; Three succinate dehydrogenase complexes, namely sdhB, sdhC and sdhD; One membrane protein, secY and orf. Therefore, they can be distinguished based on the differences in the number and types of protein-coding genes between the two, especially the differences in the coding genes of the orf protein. The start codons of protein-coding genes in *Gracilaria* and *Gracilariopsis* are mostly ATG, and only a very small number of genes are ATT. The stop codon is TAA, and in very few genes it is TAG. The lengths of the same gene are similar in different species.

Table 2-3 Characteristics of mitochondrial genomic protein-coding genes of *Gracilaria* and *Gracilariopsis*

gene product	Gene types	*G. chilensis*	*G. vermiculophylla*	*G. changii*	*G. edulis*	*G. salicornia*	*Gp. lemaneiformis*	*Gp. andersonii*	*Gp. oryzoides*	*Gp. chorda*
Cytochrome oxidase	*cox*1	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1593[c]	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1596[c]	ATG[a] TAA[b] 1596[c]
	*cox*2	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]	ATG[a] TAA[b] 792[c]
	*cox*3	ATG[a] TAA[b] 819[c]	ATG[a] TAA[b] 819[c]	ATG[a] TAA[b] 819[c]	ATG[a] TAA[b] 819[c]	ATG[a] TAAb 819[c]	ATG[a] TAA[b] 819[c]	ATG[a] TAA[b] 819[c]	ATG[a] TAA[b] 819[c]	ATG[a] TAA[b] 816[c]
Apocytochrome b	*cob*	ATG[a] TAA[b] 1143[c]	ATG[a] TAA[b] 1140[c]	ATG[a] TAA[b] 1143[c]	ATG[a] TAA[b] 1143[c]	ATG[a] TAA[b] 1140[c]	ATG[a] TAA[b] 1143[c]	ATG[a] TAA[b] 1140[c]	ATG[a] TAA[b] 1140[c]	ATG[a] TAA[b] 1143[c]
NADH	*nad*1	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]

dehydrogenase		TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]	TAA[b] 981[c]
	*nad*2	ATG[a] TAA[b] 1497[c]	ATG[a] TAA[b] 1476[c]	ATG[a] TAA[b] 1497[c]	ATG[a] TAA[b] 1497[c]	ATG[a] TAG[b] 1497[c]	ATG[a] TAA[b] 1476[c]	ATG[a] TAA[b] 1479[c]	ATG[a] TAG[b] 1473[c]	ATG[a] TAA[b] 1476[c]
	*nad*3	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]	ATG[a] TAA[b] 366[c]
	*nad*4	ATG[a] TAA[b] 1476[c]	ATG[a] TAA[b] 1461[c]	ATG[a] TAA[b] 1476[c]	ATG[a] TAA[b] 1476[c]	ATG[a] TAG[b] 1476[c]	ATG[a] TAG[b] 1476[c]	ATG[a] TAA[b] 1476[c]	ATG[a] TAA[b] 1476[c]	ATG[a] TAG[b] 1476[c]
	*nad*5	ATG[a] TAA[b] 1977[c]	ATG[a] TAA[b] 1977[c]	ATG[a] TAA[b] 1977[c]	ATG[a] TAA[b] 1977[c]	ATG[a] TAA[b] 1977[c]	ATG[a] TAA[b] 1992[c]	ATG[a] TAA[b] 1992[c]	ATG[a] TAA[b] 1992[c]	ATG[a] TAA[b] 1992[c]
	*nad*6	ATG[a] TAA[b] 609[c]	ATG[a] TAA[b] 609[c]	ATG[a] TAA[b] 609[c]	ATG[a] TAA[b] 609[c]	ATG[a] TAA[b] 609[c]	ATG[a] TAG[b] 609[c]	ATG[a] TAA[b] 609[c]	ATG[a] TAA[b] 609[c]	ATG[a] TAG[b] 609[c]
	*nad*4L	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]

		TAA[b] 306[c]	TAA[b] 306[c]	TAA[b] 306[c]	TAA[b] 306[c]	TAA[b] 306[c]	TAA[b] 306[c]	TAA[b] 306[c]	TAG[b] 306[c]	TAA[b] 306[c]
Ribosomal proteins	*rps*3	ATG[a] TAG[b] 696[c]	ATG[a] TAA[b] 696[c]	ATG[a] TAA[b] 717[c]	ATG[a] TAA[b] 696[c]	ATG[a] TAA[b] 696[c]	ATG[a] TAA[b] 696[c]	ATG[a] TAA[b] 696[c]	ATG[a] TAA[b] 696[c]	ATG[a] TAA[b] 696[c]
	*rps*11	ATT[a] TAA[b] 357[c]	ATG[a] TAA[b] 354[c]	ATG[a] TAG[b] 360[c]	ATG[a] TAA[b] 360[c]	ATG[a] TAA[b] 357[c]	ATG[a] TAA[b] 360[c]	ATG[a] TAA[b] 405[c]	ATG[a] TAA[b] 360[c]	ATG[a] TAA[b] 360[c]
	*rps*12	ATG[a] TAA[b] 369[c]	ATG[a] TAA[b] 369[c]	ATG[a] TAA[b] 369[c]	ATG[a] TAA[b] 372[c]	ATG[a] TAA[b] 369[c]	ATG[a] TAG[b] 369[c]	ATG[a] TAG[b] 240[c]	ATG[a] TAG[b] 369[c]	ATG[a] TAA[b] 384[c]
	*rpl*16	ATG[a] TAA[b] 414[c]	ATG[a] TAG[b] 414[c]	ATG[a] TAA[b] 405[c]	ATG[a] TAA[b] 405[c]	ATG[a] TAG[b] 405[c]	ATG[a] TAA[b] 411[c]	ATG[a] TAG[b] 411[c]	ATG[a] TAA[b] 411[c]	ATG[a] TAA[b] 411[c]
	*rpl*20	ATG[a] TAA[b] 234[c]	ATG[a] TAA[b] 234[c]	ATG[a] TAG[b] 252[c]	ATG[a] TAG[b] 243[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAG[b] 243[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAG[b] 243[c]
ATPase subunits	*atp*6	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]

		TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]	TAA[b] 762[c]
	*atp*8	ATG[a] TAG[b] 408[c]	ATG[a] TAG[b] 408[c]	ATG[a] TAA[b] 408[c]	ATG[a] TAA[b] 408[c]	ATG[a] TAA[b] 408[c]	ATG[a] TAA[b] 405[c]	ATG[a] TAA[b] 408[c]	_	ATG[a] TAA[b] 405[c]
	*atp*9	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]	ATG[a] TAA[b] 231[c]
	*ymf*39	ATG[a] TAA[b] 542[c]	ATG[a] TAA[b] 543[c]	ATG[a] TAA[b] 546[c]	ATG[a] TAA[b] 543[c]	ATG[a] TAG[b] 543[c]	ATG[a] TAA[b] 543[c]	_	ATG[a] TAA[b] 543[c]	_
Succinate dehydrogenase	*sdhB*	ATG[a] TAA[b] 753[c]	ATG[a] TAA[b] 753[c]	ATG[a] TAA[b] 753[c]	ATG[a] TAA[b] 753[c]	ATG[a] TAG[b] 753[c]	ATG[a] TAA[b] 753[c]	ATG[a] TAA[b] 777[c]	ATG[a] TAA[b] 753[c]	ATG[a] TAA[b] 753[c]
	*sdh*C	ATG[a] TAA[b] 378[c]	ATG[a] TAA[b] 375[c]	ATG[a] TAG[b] 378[c]	ATG[a] TAA[b] 375[c]	ATG[a] TAA[b] 378[c]	ATG[a] TAA[b] 372[c]	ATG[a] TAA[b] 381[c]	_	ATG[a] TAA[b] 372[c]
	*sdh*D	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]	ATG[a]

		TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]	TAA[b] 240[c]
Integral membrane	*sec*Y	ATG[a] TAA[b] 735[c]	ATG[a] TAA[b] 768[c]	ATG[a] TAA[b] 747[c]	ATG[a] TAA[b] 741[c]	ATG[a] TAA[b] 741[c]	ATG[a] TAA[b] 735[c]	ATG[a] TAA[b] 735[c]	ATG[a] TAA[b] 735[c]	ATG[a] TAA[b] 735[c]
Opening	*orf*148	ATG[a] TAG[b] 444[c]	–	ATG[a] TAA[b] 444[c]	ATG[a] TAA[b] 435[c]	ATG[a] TAA[b] 444[c]	–	–	–	–
	*orf*147	–	ATG[a] TAA[b] 444[c]	–	–	–	–	–	–	–
	*orf*143	–	–	–	–	–	ATG[a] TAA[b] 432[c]	ATG[a] TAG[b] 429[c]	–	–
	orf60	–	–	–	–	–	ATG[a] TAA[b] 183[c]	–	–	–
	*orf*80	–	–	–	–	–	–	–	ATG[a]	–

									TAA[b] 240[c]	
	*orf*87	–	–	–	–	–	–	ATG[a] TAG[b] 261[c]	–	–
	*orf*95	–	–	–	–	–	–	ATG[a] TAG[b] 285[c]	–	–
	*orf*61	–	–	–	–	–	–	ATG[a] TAA[b] 183[c]	–	–
	*orf*37	–	–	–	–	–	–	–	–	ATG[a] TAA[b] 432[c]
	*orf*17	–	–	–	–	–	–	–	–	ATG[a] TAA[b] 183[c]
	*orf*08	–	–	–	–	–	–	–	–	ATG[a]

										TAA[b] 543[c]

Note: a: Start codon; b: Stop codon c: Length

2.3.3 RRNA GENE

The mitochondrial genomes of the five species of the genus *Gracilaria* and those of *Gp. andersonii*, *Gp. oryzoides*, and *Gp. chorda* in the genus *Gracilariopsis* all contain only two rRNA genes, namely the LSU gene and the SSU gene. The mitochondrial genome of *Gp. lemaneiformis* additionally encodes the rrn5 gene (Table 2-4).

Table 2-4 rRNA genes of mitochondrial genomes of *Gracilaria* and *Gracilariopsis*

Gene	*LSU*	*SSU*	*rrn*5
G. chilensis	+	+	_
G. vermiculophylla	+	+	_
G. changii	+	+	_
G. edulis	+	+	_
G. salicornia	+	+	_
Gp. lemaneiformis	+	+	+
Gp. andersonii	+	+	_
Gp. oryzoides	+	+	_
Gp. chorda	+	+	_

2.3.4 CO-LINEARITY ANALYSIS

The results of the execution using Mauve in the Geneious software are shown in Figure 2-10. The mitochondrial genomes of *Gracilaria* and *Gracilariopsis* are relatively conserved and have good collinearity. Compared with *Gracilaria* genera, of the *Gracilariopsis* trnH gene rearrangement, can be used as evidence to distinguish the two genera. For example: In the *Gracilaria* genus, there are trnG genes, trnH genes and trnL genes between the sdhB genes and cob genes of each species, and the sequence is: sdhB-trnH-trnG-trnL-cob; In the genus *Gracilariopsis*, trnG genes and trnL genes exist between the sdhB genes and cob genes in each species. Except for the *Gp. andersonii*, whose sequence is: sdhB-trnG-trnL-cob, the other three are: sdhB-trnG-trnL-trnH-cob(Table 2-5).

Table 2-5 The sequence of tRNA genes in *Gracilaria* and *Gracilariopsis*

species	Genes and sequences						
G. chilensis	*sdh*2	*trn*H	*trn*G	*nad*6	*trn*L	_	*cob*
G. vermiculophylla	*sdh*2	*trn*H	*trn*G	*nad*6	*trn*L	_	*cob*
G. changii	*sdh*2	*trn*H	*trn*G	*nad*6	*trn*L	_	*cob*
G. edulis	*sdh*2	*trn*H	*trn*G	*nad*6	*trn*L	_	*cob*
G. salicornia	*sdh*2	*trn*H	*trn*G	*nad*6	*trn*L	_	*cob*
Gp. lemaneiformis	*sdh*2	_	*trn*G	*nad*6	*trn*L	*trn*H	*cob*
Gp. andersonii	*sdh*2	_	*trn*G	*nad*6	*trn*L	*trn*H	*cob*
Gp. oryzoides	*sdh*2	_	*trn*G	*nad*6	*trn*L	*trn*H	*cob*
Gp. chorda	*sdh*2	_	*trn*G	*nad*6	*trn*L	*trn*H	*cob*

In the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* species, in the trnN-trnA region, which is rich in tRNA, gene deletions and rearrangements often

occur. For example: In the mitochondrial genomes of *G. salicornia*, *Gp. lemaneiformis*, *Gp. andersonii* and *Gp. oryzoides*, only trnN and trnA exist in the trnN-trnA region, and the sequence is: trnN-trnA; The sequence of *G. changii* and *Gp. chorda* is as follows: trnN-trnY-trnR-trnS-trnA; The gene sequences in the mitochondrial genomes of *G. edulis*, *G. chilensis* and *G. vermiculophylla* are trnN-trnS-trnR-trnY-trnA, trnN-trnR-trnS-trnS-trnR-trnY-trnA and trnN-trnY-trnA, respectively.

Figure 2-10 Co-linearity analysis of mitochondrial genomes of *Gracilaria* and *Gracilariopsis*

Summary of the chapter Ⅱ

1.The research basis was clarified from the data sources, and experimental methods such as bioinformatics were adopted to conduct a comprehensive analysis of the mitochondrial genomes of the two types of algae. The results show that there are

differences between the two in terms of genomic structural characteristics, including genomic size, gene arrangement sequence, etc. Protein-coding genes, rRNA genes in quantity, sequence characteristics also have different performance.

2.A total of linear analysis to further reveal the relationships between the homology and evolution of genome. These results provide data support for understanding the uniqueness and commonality of the mitochondrial genomes of the two types of algae, which is helpful for recognizing their genetic characteristics at the molecular level and also provides an important genomic data basis for subsequent codon preference and phylogenetic analysis.

Chapter Ⅲ
EXPERIMENTAL PART

3.1 EXPERIMENTAL METHOD

Codon bias refers to the phenomenon in organisms where different codons that encode the same amino acid are used with varying frequencies. For example, leucine can be encoded by six different codons, but in a particular organism, some codons (such as UUA) may be used much more frequently than others (such as CUU). Comparing codon biases across different species can help understand their phylogenetic relationships and evolutionary history. Changes in codon bias can reflect selective pressures and functional evolution during gene evolution. Therefore, the use preferences of codons can provide theoretical support for further research on the mitochondrial genomes of the genera Nannochloris and Nannochloropsis. The specific research methods are as follows:

3.1.1 EXPERIMENTAL METHOD OF PASSWORD COMPOSITION

First, use the CodonW software to integrate the 20 CDS sequences screened out from the mitochondrial genomes of *Gracilaria* and *Gracilaria* into a ".fasta "format file through TBtools. Then, enter".\ codonw.exe "in the CodonW software and drag the file into the CodonW folder. Then enter "-all_indices -nomenu" and click the Enter key to run. After the operation is completed, the corresponding ".out "and".blk "files will appear in the same folder. The ".out "file shows the values of some related parameters, such as codon Fit Index (CAI), Optimal codon usage frequency (Fop), codon Preference index (CBI), etc. The".blk "file shows the relative usage degree of codons, that is, the

RSCU value [13]. In emboss website (https://www.bioinformatics.nl/emboss-explorer/) import CDS ". Fasta format file "ENC actual value, GC and GC3 value.

3.1.2 ENC-PLOT ANALYSIS AND PLOTTING

According to the measured GC_3 values, the expected value of ENC can be calculated using the formula: Expected ENC =2 + GC_3 + 29 / [GC_3^2 + (1-GC_3)^2]^{34,}. Then, with GC_3 as the x-axis and the expected value of ENC as the y-axis use Origin software to plot a standard curve. Next, plot scatter diagrams with the actual GC_3 and ENC values on the x-axis and y-axis respectively. Finally, merge these two plots using Origin drawing software to obtain the ENC-plot.

3.1.3 PR2-PLOT ANALYSIS AND PLOTTING

The PR2-plot is a scatter plot with $G_3/(G_3+C_3)$ as the horizontal coordinate and $A_3/(A_3+T_3)$ as the vertical coordinate. The relevant values measured by CodonW are calculated according to the horizontal and vertical coordinates and plotted in excel. The center point of the PR2-plot is the value when A=T and C=G, which means that the point represents the codon with no preference.

3.2 RESULTS AND ANALYSIS

3.2.1 ANALYSIS OF PASSWORD COMPOSITION

The parameters related to the codons of the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* were analyzed (Table 3-1). The results showed that the GC contents at different positions in the codons were different, but the GC contents of *Gracilaria* and *Gracilariopsis* were not much different. The average ENC of the codon usage frequency of the mitochondrial genome of *Gracilaria* was 36.9, and most of them exceeded 35. The average ENC value of the genomic codon of *Gracilariopsis* was 37.2,

and most of them were also higher than 35. This indicates that the mitochondrial genomes of both show a relatively weak preference in the use of codons. In addition, the CAI (Codon Adaptation Index) and CBI (Codon Preference Index) concepts are similar, both used to measure the preference for the use of codons in genes, reflecting the degree of a species' "preference" for a specific codon. The higher the CAI value is, the higher the gene expression efficiency may be. Genes with high CBI values usually use codons that frequently occur in host cells, which is conducive to efficient translation and is commonly seen in highly expressed genes. The average CAI values of the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* were 0.153 and 0.154 respectively, indicating that the gene expression levels were basically the same. The average values of mitochondrial genomic CBI of the two were -0.197 and -0.184 respectively, indicating that the codon preference index was similar, but the *Gracilariopsis* was slightly more significant than the *Gracilaria*. The average values of Fop (Optimal Codon Usage Frequency) were 0.285 and 0.295 respectively; The Gravy average values were both 0.551 and 0.572, which indicates that the encoded protein may be hydrophobic.

In the analysis of codon preference, when the RSCU value is 1, it indicates that the codon has no significant preference. When the RSCU value is greater than 1, it indicates that the codon has a relatively high usage preference. When RSCU is less than 1, it indicates that the preference for the use of codons is relatively weak. The analysis of the results obtained in this study shows that for codons ending with adenine (A) or uracil (U), the RSCU values are mostly greater than 1, indicating that this codon has a strong preference. The RSCU value of some codons ending with guanine (G) is equal to 1, and there is no preference for codons. Most codons with RSCU values less than 1 end with G or cytosine (C). This result indicates that in the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* codons ending with A or U are preferred, and those ending with G or C are rarely used. Among the codons with RSCU>1, the RSCU value of UUA encoding leucine is the largest, mostly greater than 4. This indicates that the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* have a stronger preference for UUA.

Table 3-1 Main Parameters of mitochondrial genomic codons in *Gracilaria* and *Gracilariopsis*

Gene	CAI	CBI	Fop	Gravy
*atp*6	0.239[a1]/0.129[a2]/0.128[a3]/0.135[a4]/0.1[a5]/0.140[b1]/0.142[b2]/0.142[b3]/0.128[b4]	-0.004[a1]/-0.279[a2]/-0.245[a3]/-0.237[a4]/-0.327[a5]/-0.169[b1]/-0.19[b2]/-0.163[b3]/-0.244[b4]	0.448[a1]/0.218[a2]/0.236[a3]/0.243[a4]/0.185[a5]/0.291[b1]/0.275[b2]/0.291[b3]/0.242[b4]	-0.894[a1]/1.175[a2]/1.234[a3]/1.19[a4]/1.177[a5]/1.125[b1]/1.158[b2]/1.157[b3]/1.157[b4]
*nad*5	0.196[a1]/0.144[a2]/0.131[a3]/0.142[a4]/0.142[a5]/0.148[b1]/0.148[b2]/0.136[b3]/0.139[b4]	0.002[a1]/-0.186[a2]/-0.251[a3]/-0.283[a4]/-0.224[a5]/-0.212[b1]/-0.212[b2]/-0.229[b3]/-0.251[b4]	0.451[a1]/0.285[a2]/0.247[a3]/0.23[a4]/0.[262a5]/0.270[b1]/0.27[b2]/0.261[b3]/0.246[b4]	-0.580[a1]/0.929[a2]/0.890[a3]/1.013[a4]/0.975[a5]/0.939[b1]/0.94[b2]/0.918[b3]/0.162[b4]
*nad*4	0.216[a1]/o.126[a2]/0.127[a3]/0.123[a4]/0.13[a5]/0.122[b1]/0.131[b2]/0.141[b3]/0.119[b4]	0.01[a1]/-0.26[a2]/-0.241[a3]/-0.293[a4]/-0.26[a5]/-0.247[b1]/-0.235[b2]/-	0.451[a1]/0.234[a2]/0.239[a3]/0.211[a4]/0.23[a5]/0.239[b1]/0.237[b2]/0.267[b3]/-.227[b4]	-0.868[a1]/1.104[a2]/1.122[a3]/1.17[a4]/1.11[a5]/1.110[b1]/1.118[b2]/1.111[b3]/0.163[b4]

		0.202[b3]/-0.272[b4]		
*nad*2	0.197[a1]/0.[139a2]/0.116[a3]/0.09[a4]/0.129[a5]/0.219[b1]/0.136[b2]/0.141[b3]/0.139[b4]	-0.023[a1]/-0.27[a2]/-0.303[a3]/-0.283[a4]/-0.273[a5]/0.003[b1]/-0.241[b2]/-0.24[b3]/-0.27[b4]	0.438[a1]/0.235[a2]/0.217[a3]/0.222[a4]/0.220[a5]/0.454[b1]/0.248[b2]/0.256[b3]/0.235[b4]	-0.665[a1]/1.174[a2]/1.066[a3]/1.176[a4]/1.225[a5]/-0.787[b1]/1.215[b2]/1.13[b3]/0.174[b4]
*nad*1	0.179[a1]/0.119[a2]/0.123[a3]/0.148[a4]/0.138[a5]/0.190[b1]/0.125[b2]/0.125[b3]/0.116[b4]	-0.043[a1]/-0.175[a2]/-0.167[a3]/-0.162[a4]/-0.161[a5]/-0.010[b1]/-0.173[b2]/-0.2[b3]/-0.203[b4]	0.429[a1]/0.276[a2]/0.279[a3]/0.283[a4]/0.284[a5]/0.447[b1]/0.28[b2]/0.263[b3]/0.257[b4]	-0.697[a1]/1.115[a2]/1.085[a3]/1.114[a4]/1.049[a5]/-0.790[b1]/1.117[b2]/1.16[b3]/0.116[b4]
*nad*3	0.176[a1]/0.125[a2]/0.133[a3]/0.163[a4]/0.162[a5]/0.216[b1]/0.138[b2]/0.148[b3]/0.124[b4]	0.001[a1]/-0.289[a2]/-0.225[a3]/-0.291[a4]/-0.217[a5]/0.057[b1]/-0.239[b2]/-0.248[b3]/-0.271[b4]	0.443[a1]/0.235[a2]/0.27[a3]/0.226[a4]/0.27[a5]/0.472[b1]/0.252[b2]/0.246[b3]/0.235[b4]	-0.885[a1]/1.112[a2]/1.07[a3]/1.10[a4]/1.04[a5]/-0.917[b1]/1.126[b2]/1.107[b3]/0.188[b4]
*rps*11	0.147[a1]/0.11[a2]/0.127[a3]/0.146[a4]/0.152[a5]/0.1	-0.111[a1]/-0.158[a2]/-0.174[a3]/-	0.342[a1]/0.325[a2]/0.310[a3]/0.319[a4]/0.333[a5]/0.	-0.018[a1]/-0.075[a2]/-0.087[a3]/0.086[a4]/-0.057[a5]/-

	88^{b1}/0.171^{b2}/0.125^{b3}/0.162^{b4}	0.185^{a4}/-0.132^{a5}/-0.140^{b1}/-0.242^{b2}/-0.175^{b3}/-0.083^{b4}	365^{b1}/0.252^{b2}/0.31^{b3}/0.357^{b4}	0.391^{b1}/0.756^{b2}/0.074^{b3}/0.076^{b4}
*nad*6	0.231^{a1}/0.155^{a2}/0.125^{a3}/0.111^{a4}/0.135^{a5}/0.140^{b1}/0.14^{b2}/0.143^{b3}/0.141^{b4}	-0.068^{a1}/-0.076^{a2}/-0.233^{a3}/-0.213^{a4}/-0.235^{a5}/-0.194^{b1}/-0.194^{b2}/-0.213^{b3}/-0.203^{b4}	0.407^{a1}/0.333^{a2}/0.244^{a3}/0.254^{a4}/0.244^{a5}/0.269^{b1}/0.269^{b2}/0.259^{b3}/0.269^{b4}	-0.657^{a1}/1.191^{a2}/1.1735^{a3}/1.213^{a4}/1.202^{a5}/1.092^{b1}/1.092^{b2}/1.0545^{b3}/0.14^{b4}
cob	0.148^{a1}/0.154^{a2}/0.137^{a3}/0.125^{a4}/0.154^{a5}/0.129^{b1}/0.119^{b2}/0.127^{b3}/0.144^{b4}	-0.229^{a1}/-0.185^{a2}/-0.239^{a3}/-0.344^{a4}/-0.194^{a5}/-0.214^{b1}/-0.301^{b2}/-0.251^{b3}/-0.183^{b4}	0.251^{a1}/0.275^{a2}/0.244^{a3}0.186^{a4}/0.272^{a5}/0.261^{b1}/0.208^{b2}/0.239^{b3}/0.277^{b4}	0.876^{a1}/0.892^{a2}/0.866^{a3}/0.873^{a4}/0.871^{a5}/0.853^{b1}/0.86^{b2}/0.859^{b3}/0.15^{b4}
*cox*3	0.168^{a1}/0.156^{a2}/0.158^{a3}/0.141^{a4}/0.162^{a5}/0.219^{b1}/0.139^{b2}/	-0.203^{a1}/-0.233^{a2}/-0.278^{a3}/-0.245^{a4}/-	0.286^{a1}/0.269^{a2}/0.241^{a3}/0.263^{a4}/0.245^{a5}/0.469^{b1}/0.231^{b2}/	0.784^{a1}/0.755^{a2}/0.805^{a3}/0.789^{a4}/0.79^{a5}/-0.561^{b1}/0.726^{b2}/0.7

	0.147^{b3}/0.154^{b4}	0.28^{a5}/0.039^{b1}/-0.297^{b2}/-0.263^{b3}/-0.234^{b4}	0.252^{b3}/0.265^{b4}	78^{b3}/0.146^{b4}
*cox*2	0.154^{a1}/0.168^{a2}/0.162^{a3}/0.131^{a4}/0.147^{a5}/0.162^{b1}/0.162^{b2}/0.18^{b3}/0.163^{b4}	-0.221^{a1}/-0.133^{a2}/-0.225^{a3}/-0.327^{a4}/-0.242^{a5}/-0.132^{b1}/-0.216^{b2}/-0.204^{b3}/-0.24^{b4}	0.273^{a1}/0.333^{a2}/0.272^{a3}/0.211^{a4}/0.262^{a5}/0.364^{b1}/0.276^{b2}/0.282^{b3}/0.264^{b4}	0.381^{a1}/0.345^{a2}/0.414^{a3}/0.388^{a4}/0.466^{a5}/-0.182^{b1}/0.397^{b2}/0.413^{b3}/0.109^{b4}
*cox*1	0.162^{a1}/0.163^{a2}/0.145^{a3}/0.138^{a4}/0.164^{a5}/0.172^{b1}/0.159^{b2}/0.147^{b3}/0.109^{b4}	-0.232^{a1}/-0.216^{a2}/-0,294^{a3}/-0.305^{a4}/-0.271^{a5}/-0.186^{b1}/-0.272^{b2}/-0.292^{b3}/-0.212^{b4}	0.270^{a1}/0.281^{a2}/0.232^{a3}/0.227^{a4}/0.245^{a5}/0.297^{b1}/0.249^{b2}/0.234^{b3}/0.282^{b4}	0.759^{a1}/0.729^{a2}/0.760^{a3}/0.77^{a4}/0.755^{a5}/0.757^{b1}/0.721^{b2}/0.746^{b3}/0.143^{b4}
*rpl*16	0.196^{a1}/0.193^{a2}/0.129^{a3}/0.185^{a4}/0.149^{a5}/0.131^{b1}/0.193^{b2}/0.161^{b3}/0.146^{b4}	-0.119^{a1}/-0.054^{a2}/-0.163^{a3}/-0.151^{a4}/-0.108^{a5}/-0.203^{b1}/-0.016^{b2}/-	0.369^{a1}/0.389^{a2}/0.328^{a3}/0.344^{a4}/0.354^{a5}/0.295^{b1}/0.409^{b2}/0.362^{b3}/0.368^{b4}	-0.284^{a1}/-0.26^{a2}/-0.412^{a3}/-0.37^{a4}/-0.386^{a5}/-0.354^{b1}/-0.422^{b2}/-0.341^{b3}/0.126^{b4}

		0.098[b3]/-0.168[b4]		
*rps*3	0.131[a1]/0.124[a2]/0.138[a3]/0.141[a4]/0.134[a5]/0.205[b1]/0.151[b2]/0.158[b3]/0.137[b4]	-0.299[a1]/-0.287[a2]/-0.242[a3]/-0.241[a4]/-0.233[a5]/-0.069[b1]/-0.157[b2]/-0.163[b3]/-0.218[b4]	0.232[a1]/0.247[a2]/0.270[a3]/0.265[a4]/0.269[a5]/0.385[b1]/0.312[b2]/0.314[b3]/0.271[b4]	0.045[a1]/0.012[a2]/-0.153[a3]/0.096[a4]/-0.059[a5]/0.578[b1]/-0.021[b2]/-0.012[b3]/0.134[b4]
*nad*4L	0.140[a1]/0.128[a2]/0.146[a3]/0.107[a4]/0.126[a5]/0.211[b1]/0.151[b2]/0.139[b3]/0.152[b4]	-0.205[a1]/-0.243[a2]/-0.136[a3]/-0.292[a4]/-0.233[a5]/-0.079[b1]/-0.159[b2]/-0.172[b3]/-0.141[b4]	0.255[a1]/0.234[a2]/0.298[a3]/0.191[a4]/0.239[a5]/0.412[b1]/0.287[b2]/0.277[b3]/0.298[b4]	1.277[a1]/1.265[a2]/1.319[a3]/1.44[a4]/1.316[a5]/-0.753[b1]/1.2[b2]/1.25[b3]/0.089[b4]
*rps*12	0.136[a1]/0.165[a2]/0.145[a3]/0.174[a4]/0.179[a5]/0.159[b1]/0.151[b2]/0.158[b3]/0.155[b4]	-0.090[a1]/-0.083[a2]/-0.216[a3]/-0.11[a4]/-0.078[a5]/-0.101[b1]/-0.089[b2]/-0.11[b3]/-0.105[b4]	0.358[a1]/0.367[a2]/0.292[a3]/0.342[a4]/0.367[a5]/0.350[b1]/0.368[b2]/0.35[b3]/0.352[b4]	-0.716[a1]/-0.69[a2]/-0.786[a3]/-0.547[a4]/-0.6344[a5]/-0.651[b1]/-0.926[b2]/-0.636[b3]/0.055[b4]

*sec*Y	0.099^{a1}/0.15^{a2}/0.156^{a3}/0.127^{a4}/0.149^{a5}/0.217^{b1}/0.142^{b2}/0.138^{b3}/0.135^{b4}	-0.304^{a1}/-0.211^{a2}/-0.311^{a3}/-0.382^{a4}/-0.248^{a5}/-0.021^{b1}/-0.304^{b2}/-0.327^{b3}/-0.289^{b4}	0.222^{a1}/0.268^{a2}/0.225^{a3}/0.169^{a4}/0.244^{a5}/0.445^{b1}/0.22^{b2}/0.213^{b3}/0.228^{b4}	0.618^{a1}/1.037^{a2}/1.088^{a3}/1.167^{a4}/1.27^{a5}/-0.846^{b1}/1.152^{b2}/1.015^{b3}/0.275^{b4}
*rpl*20	0.175^{a1}/0.204^{a2}/0.176^{a3}/0.15^{a4}/0.133^{a5}/0.232^{b1}/0.213^{b2}/0.169^{b3}/0.174^{b4}	0.023^{a1}/-0.017^{a2}/-0.180^{a3}/-0.158^{a4}/-0.291^{a5}/-0.119^{b1}/0^{b2}/-0.193^{b3}/-0.233^{b4}	0.426^{a1}/0.413^{a2}/0.329^{a3}/0.319^{a4}/0.23^{a5}/0.362^{b1}/0.431^{b2}/0.315^{b3}/0.316^{b4}	-0.426^{a1}/-0.216^{a2}/-0.145^{a3}/-0.529^{a4}/0.088^{a5}/0.546^{b1}/-0.521^{b2}/-0.417^{b3}/0.126^{b4}
*sdh*D	0.258^{a1}/0.075^{a2}/0.111^{a3}/0.095^{a4}/0.138^{a5}/0.229^{b1}/0.076^{b2}/0.076^{b3}/0.184^{b4}	-0.012^{a1}/-0.288^{a2}/-0.189^{a3}/-0.249^{a4}/-0.166^{a5}/-0.249^{b1}/-0.32^{b2}/-0.273^{b3}/-0.231^{b4}	0.455^{a1}/0.171^{a2}/0.263^{a3}/0.197^{a4}/0.267^{a5}/0.339^{b1}/0.156^{b2}/0.182^{b3}/0.211^{b4}	-0.806^{a1}/1.486^{a2}/1.387^{a3}/1.514^{a4}/1.198^{a5}/-0.761^{b1}/1.494^{b2}/1.59^{b3}/0.115^{b4}

*atp*9	0.281^{a1}/0.178^{a2}/0.190^{a3}/0.158^{a4}/0.168^{a5}/0.180^{b1}/0.231^{b2}/0.236^{b3}/0.185^{b4}	0.172^{a1}/- 0.065^{a2}/- 0.136^{a3}/- 0.136^{a4}/- 0.231^{a5}/- 0.089^{b1}/- 0.065^{b2}/- 0.065^{b3}/- 0.136^{b4}	0.556^{a1}/0.366^{a2}/0.324^{a3}/0.324^{a4}/0.268^{a5}/0.352^{b1}/0.366^{b2}/0.336^{b3}/0.234^{b4}	- 0.413^{a1}/1.201^{a2}/1.201316^{a3}/1.201^{a4}/1.201^{a5}/1.201^{b1}/1.201^{b2}/1.2^{b3}/0.105^{b4}

Note: a1 represents *Gracilaria chilensis*; a2 represents *Gracilaria vermiculophylla*; a3 represents *Gracilaria changii*; a4 represents *Gracilaria edulis*; a5 represents *Gracilaria salicornia*; b1 represents *Gracilariopsis lemaneiformis*; b2 represents *Gracilariopsis andersonii*; b3 represents *Gracilariopsis oryzoides*; b4 represents *Gracilariopsis chorda*

3.2.2 ENC-PLOT ANALYSIS

The ENC value is usually measured by 35 as the standard to assess the strength of the preference of a codon. When the ENC value exceeds 35, it indicates that the preference of this codon is relatively weak. On the contrary, when the ENC value is lower than 35, it indicates that the preference of this codon is relatively strong.. It can be seen from Table 3-2 that the ENC of the mitochondrial genomic codons of *Gracilaria* and *Gracilariopsis* is mostly greater than 35, indicating that the codon preferences of both are relatively weak.

Use ENC expectations formula calculation, the results are shown in table 3-2.

Table 3-2 Analysis of codon preference in mitochondrial genome of *Gracilaria* and *Gracilariopsis*

Gene	ENC actual value	ENC desired	GC_3

		value	
*atp*6	48.53^{a1}/34.2^{a2}/35.4^{a3}/31.84^{a4}/35.46^{a5}/39.98^{b1}/37.06^{b2}/38.72^{b3}/38.42^{b4}	37.44^{a1}/18.20^{a2}/16.25^{a3}/12.77^{a4}/15.48^{a5}/20.54^{b1}/18.98^{b2}/20.54^{b3}/18.6^{b4}	35.43^{a}/16.14^{a2}/14.17^{a3}/10.63^{a4}/13.39^{a5}/18.5^{b}/16.93^{b2}/18.5^{b3}/16.54^{b4}
*nad*5	46.56^{a1}/39.692^{a2}/38.13^{a3}/33.26^{a4}/38/27^{a5}/34.74^{b1}/38.19^{b2}/38.36^{b3}/35.97^{b4}	34.69^{a}/20.41^{a2}/19.05^{a3}/12.76^{a4}/19.05^{a5}/19.81^{b1}/19.82^{b2}/19.68^{b3}/15.94^{b4}	32.68^{a}/18.36^{a2}/17^{a3}/10.62^{a4}/17^{a5}/17.77^{b}/17.77^{b2}/17.62^{b3}/13.86^{b4}
*nad*4	47.06^{a1}/41.82^{a2}/36.09^{a3}/33.05^{a4}/37.75^{a5}/38.62^{b1}/38.74^{b2}/40.5^{b3}/40.03^{b4}	33.92^{a}/21.14^{a2}/14.51^{a3}/11.93^{a4}/19.33^{a5}/20.32^{b1}/19.53^{b2}/21.15^{b3}/19.33^{b4}	31.91^{a}/19.1^{a2}/13.41^{a3}/9.76^{a4}/17.28^{a5}/18.28^{b}/17.48^{b2}/19.11^{b3}/17.28^{b4}
*nad*2	38.63^{a1}/36.58^{a2}/35.5^{a3}/40^{a4}/41.75^{a5}/37.87^{b1}/36.36^{b2}/37.04^{b3}/36.58^{b4}	32.29^{a}/16.31^{a2}/15.52^{a3}/13.54^{a4}/18.69^{a5}/31.54^{b1}/18.69^{b2}/18.15^{b3}/16.31^{b4}	30.28^{a}/14.23^{a2}/13.43^{a3}/11.42^{a4}/16.63^{a5}/29.53^{b}/16.63^{b2}/16.09^{b3}/14.23^{b4}
*nad*1	37.05^{a1}/41.25^{a2}/37.79^{a3}/31.62^{a4}/37.03^{a5}/52.17^{b1}/40.82^{b2}/38.26^{b3}/40.15^{b4}	39.32^{a}/22.22^{a2}/22.22^{a3}/13.14^{a4}/20.4^{a5}/41.14^{b1}/19.79^{b2}/18.57^{b3}/20.4^{b4}	37.31^{a}/20.18^{a2}/16.82^{a3}/11.01^{a4}/18.35^{a5}/39.14^{b}/17.74^{b2}/16.51^{b3}/18.35^{b4}
*nad*3	36.73^{a1}/28.35^{a2}/39.98^{a3}/32.06^{a4}/31.72^{a5}/38.02^{b1}/33.61^{b2}/37.07^{b3}/32.58^{b4}	35.62^{a}/11.22^{a2}/22.53^{a3}/11.22^{a4}/13.6^{a5}/12^{b2}/15.2^{b3}/13.6^{b4}/35.62^{b1}	33.61^{a}/9.02^{a2}/20.49^{a3}/9.02^{a4}/11.48^{a5}/33.61^{b}/9.84^{b2}/13.11^{b3}/11.48^{b4}
*rps*11	40.29^{a1}/34.22^{a2}/37.72^{a3}/34.55^{a4}/45.48^{a5}/32.99^{b1}/39.38^{b2}/32.36^{b3}/29.47^{b4}	21.86^{a}/14.81^{a2}/17.9^{a3}/9.8^{a4}/18.86^{a5}/17.44^{b1}/18.36^{b2}/15.42^{b3}/13.79^{b4}	19.83^{a}/12.71^{a2}/15.83^{a3}/7.5^{a4}/16.81^{a5}/15.38^{b}/16.3^{b2}/13.33^{b3}/11.67^{b4}
*nad*6	36.17^{a1}/40.88^{a2}/33.18^{a3}	35.02^{a}/21.25^{a2}/16.37^{a3}/14.42^{a4}/19.3^{a}	33^{a}/19.21^{a2}/14.29^{a3}/12.32^{a}

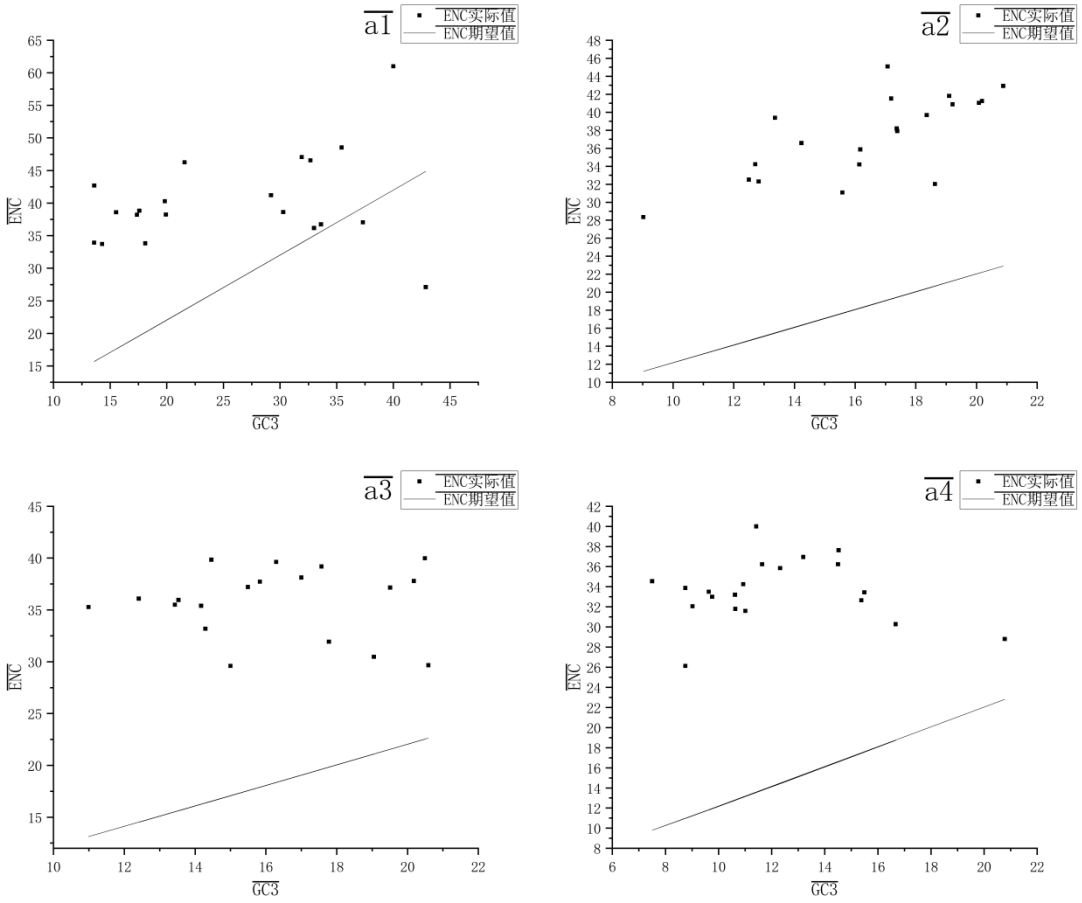
	/35.85^{a4}/36.91^{a5}/33.81^{b1}/36.69^{b2}/38.07^{b3}/38.19^{b4}	^{5}/15.87^{b1}/15.88^{b2}/16.86^{b3}/16.3^{b4}	^{4}/17.24^{a5}/13.79^{b}/13.79^{b2}/14.78^{b3}/14.29^{b4}
cob	38.21^{a1}/38.19^{a2}/37.21^{a3}/33.44^{a4}/41.27^{a5}/38.38^{b1}/37.4^{b2}/37.29^{b3}/37.66^{b4}	19.42^{a}/19.42^{a2}/17.55^{a3}/17.55^{a4}/21^{a5}/18.85^{b1}/15.25^{b2}/15.76^{b3}/16.25^{b4}	17.37^{a}/17.37^{a2}/15.49^{a3}/10.76^{a4}/18.95^{a5}/16.8^{b}/13.16^{b2}/13.68^{b3}/14.17^{b4}
*cox*3	38.83^{a1}/42.92^{a2}/35.28^{a3}/36.96^{a4}/36.73^{a5}/50.64^{b1}/39.32^{b2}/35.63^{b3}/35.12^{b4}	19.63^{a}/22.91^{a2}/13.12^{a3}/15.28^{a4}/15.28^{a5}/43.39^{b1}/20.4^{b2}/17.83^{b3}/16.42^{b4}	17.58^{a}/20.88^{a2}/10.99^{a3}/13.19^{a4}/13.19^{a5}/41.39^{b}/18.32^{b2}/15.75^{b3}/14.34^{b4}
*cox*2	38.6^{a1}/41.04^{a2}/39.63^{a3}/33.88^{a4}/39.48^{a5}/36.03^{b1}/35.12^{b2}/34.54^{b3}/36.83^{b4}	17.59$^{a/}$22.12^{a2}/18.35^{a3}/11^{a4}/18.35^{a5}/40.27^{b1}/16.84^{b2}/16.47^{b3}/17.22^{b4}	15.53^{a}/20.08^{a2}/16.29^{a3}/8.71^{a4}/16.29^{a5}/38.26^{b}/14.77^{b2}/14.39^{b3}/15.15^{b4}
*cox*1	38.24^{a1}/35.87^{a2}/35.96^{a3}/36.24^{a4}/36.07^{a5}/35.99^{b1}/35.06^{b2}/36.54^{b3}/38.98^{b4}	21.96^{a}/18.23^{a2}/15.62^{a3}/16.57^{a4}/19.91^{a5}/20.61^{b1}/17.66^{b2}/18.41^{b3}/20.28^{b4}	19.92^{a}/16.17^{a2}/13.53^{a3}/14.5^{a4}/17.86^{a5}/18.57^{b}/15.6^{b2}/16.35^{b3}/18.23^{b4}
*rpl*16	41.21^{a1}/37.9^{a2}/31.93^{a3}/33.5^{a4}/37^{a5}/33.74^{b1}/33.74^{b2}/38.32^{b3}/36.67^{b4}	31.22^{a}/19.44^{a2}/19.83^{a3}/11.8^{a4}/23.51^{a5}/21.75^{b1}/25.39^{b2}/18.12^{b3}/19.57^{b4}	29.2^{a}/17.39^{a2}/17.78^{a3}/9.63^{a4}/21.48^{a5}/19.72^{b}/23.36^{b2}/16.06^{b3}/17.52^{b4}
*rps*3	33.83^{a1}/39.38^{a2}/39.19^{a3}/36.24^{a4}/39.44^{a5}/43.68^{b1}/40.47^{b2}/40.10^{b3}/32.24^{b4}	20.15^{a}/15.45^{a2}/19.62^{a3}/13.76^{a4}/18.86^{a5}/45.00^{b1}/16.3^{b2}/18.44^{b3}/11.25^{b4}	18.1^{a}/13.36^{a2}/17.57^{a3}/11.64^{a4}/16.81^{a5}/42.99^{b}/14.22^{b2}/16.38^{b3}/9.05^{b4}
*nad*4L	46.26^{a1}/32.03^{a2}/29.66^{a3}/30.28^{a4}/42.54^{a5}/47.24^{b}	23.60^{a}/20.67^{a2}/22.63^{a3}/18.73^{a4}/21.65^{a5}/33.39^{b1}/17.75^{b2}/	21.57^{a}/18.63^{a2}/20.59^{a3}/16.67^{a4}/19.61^{a5}/31.37^{b}/15.69^{b}

	^{1}/36.12^{b2}/41.34^{b3}/29.51^{b4}	20.67^{b3}/17.75^{b4}	^{2}/18.63^{b3}/15.69^{b4}
*rps*12	42.69^{a1}/45.08^{a2}/37.15^{a3}/37.63^{a4}/40.18^{a5}/34.84^{b1}/48.61^{b2}/44.61^{b3}/37.14^{b4}	15.68^{a}/19.12^{a2}/21.55^{a3}/16.59^{a4}/17.51^{a5}/42.01^{b2}/19.12^{b3}/16.91^{b4}/20.74^{b1}	13.6^{a}/17.07^{a2}/19.51^{a3}/14.52^{a4}/15.45^{a5}/18.7^{b}/40^{b2}/17.07^{b3}/14.84^{b4}
*sec*Y	33.93^{a1}/41.52^{a2}/39.84^{a3}/34.25^{a4}/46.09^{a5}/37.23^{b1}/39.23^{b2}/38.39^{b3}/35.96^{b4}	15.68^{a}/19.24^{a2}/16.53^{a3}/13.06^{a4}/19.86^{a5}/25.30^{b1}/13.95^{b2}/19.6^{b3}/17.98^{b4}	13.59^{a}/17.19^{a2}/14.46^{a3}/10.93^{a4}/17.81^{a5}/23.27^{b}/11.84^{b2}/17.55^{b3}/15.92^{b4}
*rpl*20	33.72^{a1}/32.31^{a2}/30.48^{a3}/32.66^{a4}/29.44^{a5}/40.01^{b1}/33.76^{b2}/32.9^{b3}/26.35^{b4}	16.36^{a}/14.92^{a2}/21.09^{a3}/17.45^{a4}/15.08^{a5}/24.11^{b1}/22.82^{b2}/15.08^{b3}/18.11^{b4}	14.29^{a}/12.82^{a2}/19.05^{a3}/15.38^{a4}/12.99^{a5}/22.08^{b}/20.78^{b2}/12.99^{b3}/16.05^{b4}
*sdh*D	61.00^{a1}/32.51^{a2}/29.6^{a3}/26.13^{a4}/35.89^{a5}/53.18^{b1}/35.29^{b2}/44.5^{b3}/38.18^{b4}	42.00^{a}/14.6^{a2}/17.07^{a3}/10.96^{a4}/12.16^{a5}/24.52^{b1}/18.31^{b2}/14.6^{b3}/14.7^{b4}	40^{a}/12.5^{a2}/15^{a3}/8.75^{a4}/10^{a5}/22.5^{b}/16.25^{b2}/12.5^{b3}/12.5^{b4}
*atp*9	27.1^{a1}/31.08^{a2}/0.338^{a3}/28.81^{a4}/30.09^{a5}/30.75^{b1}/23.55^{b2}/22.77^{b3}/27.4^{b4}	44.86^{a}/17.64^{a2}/16.37^{a3}/22.82^{a4}/17.64^{a5}/20.22^{b1}/12.54^{b2}/15.08^{b3}/17.64^{b4}	42.86^{a}/15.58^{a2}/14.29^{a3}/20.78^{a4}/15.58^{a5}/18.18^{b}/10.39^{b2}/12.99^{b3}/15.58^{b4}

Note: a1 represents *Gracilaria chilensis*; a2 represents *Gracilaria vermiculophylla*; a3 represents *Gracilaria changii*; a4 represents *Gracilaria edulis*; a5 represents *Gracilaria salicornia*; b1 represents *Gracilariopsis lemaneiformis*; b2 represents *Gracilariopsis andersonii*; b3 represents *Gracilariopsis oryzoides*; b4 represents *Gracilariopsis chorda*

The distance between the scattered points and the standard curve, that is, the difference between the actual value and the expected value of ENC, can reflect the main factors affecting the codon preference. When the actual value of ENC is closer to the

expected value, base mutation becomes an important factor affecting the preference for codon usage. Conversely, it indicates that the preference of codons is influenced by the pressure of environmental selection. The ENC-plot plot obtained using origin is shown in Figure 3-1. It was observed that the distribution relationship between ENC and GC3 in *Gracilaria* and *Gracilariopsis* is roughly the same. Most of the scattered points are concentrated in the area close to the standard curve, among which a few scattered points are below the standard curve, and the vast majority are above the standard curve. Therefore, the preferences of the codons of *Gracilaria* and *Gracilariopsis* are affected by both mutations and environmental selection pressures



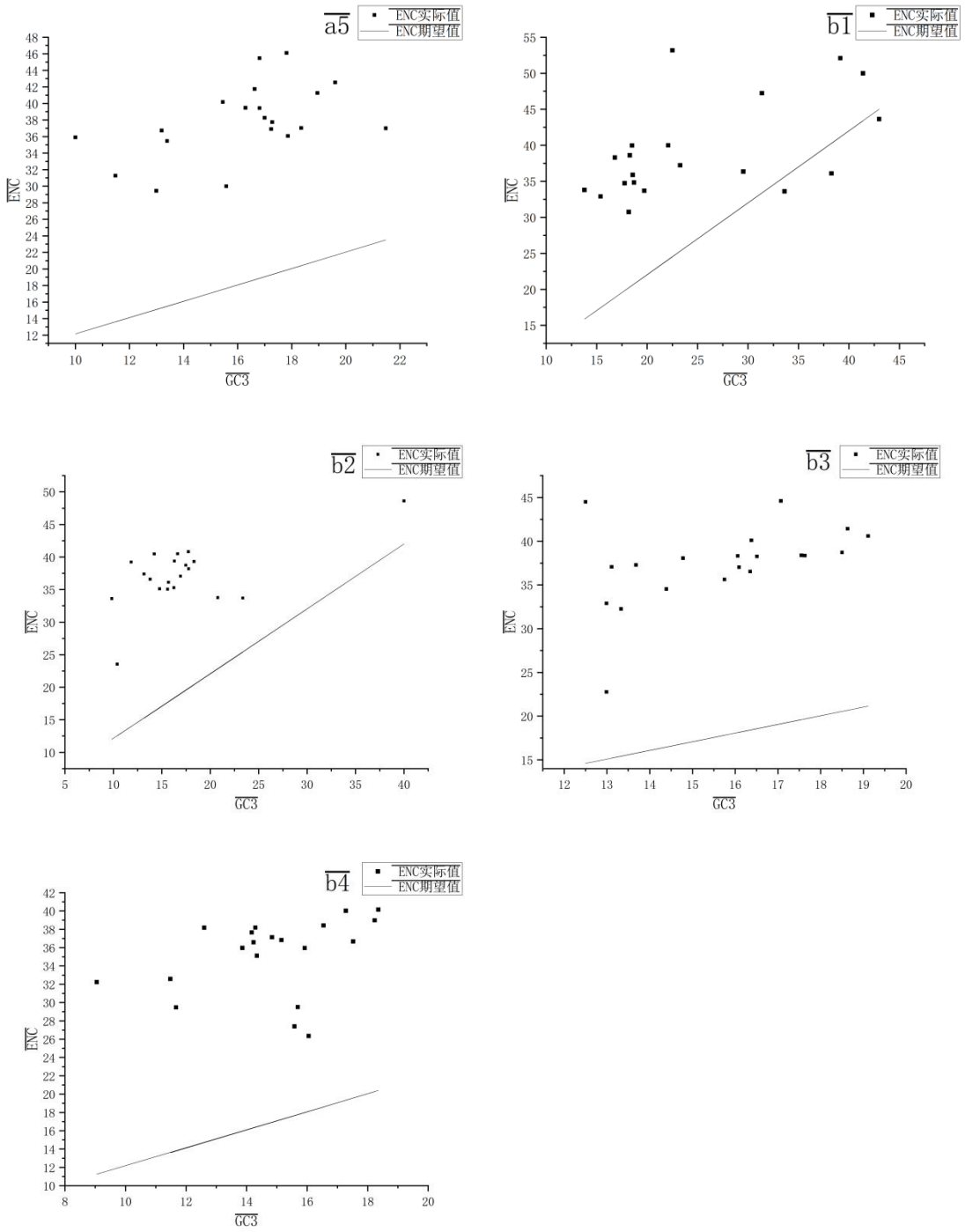


Figure 3-1 ENC-plot analysis of mitochondrial genomes of *Gracilaria* and *Gracilariopsis*

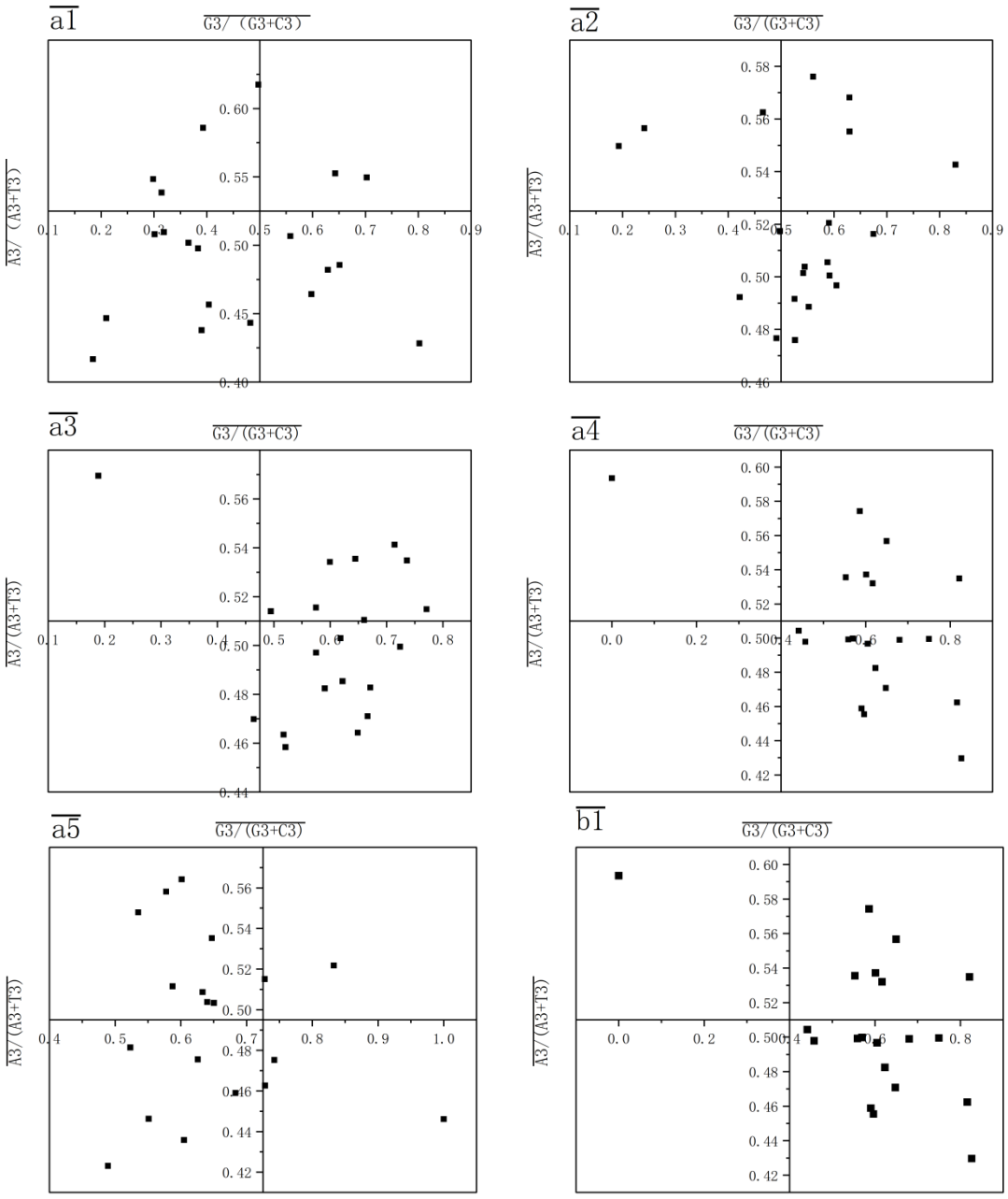
Note: a1 represents *Gracilaria chilensis*; a2 represents *Gracilaria vermiculophylla*; a3 represents *Gracilaria changii*; a4 represents *Gracilaria edulis*; a5 represents *Gracilaria salicornia*; b1 represents *Gracilariopsis lemaneiformis*; b2 represents *Gracilariopsis andersonii*; b3 represents *Gracilariopsis oryzoides*; b4 represents *Gracilariopsis chorda*

3.2.3 PR2-PLOT ANALYSIS

In PR2-plot, the scatter plot is plotted with A3/ (A3+T3) of the amino acids encoded by the four codons at the 3rd codon as the vertical coordinate and G3/ (G3+C3) as the horizontal coordinate. The center point of the figure (A=T and C=G) represents unbiased usage, and the positional relationship between each point and the center represents the direction and degree of gene bias[Ошибка! Источник ссылки не найден.]. For example, if the point is close to the right side of the horizontal coordinate, it indicates that the third bit of the codon is more inclined to use G or C; If the point is closer to the upper part of the vertical coordinate, it indicates that A or U is more preferred.

By figure 3-2 can be observed that all coordinates points are not evenly distributed in four areas of the scatterplot, *Gracilariopsis* is the majority of gene distribution in the figure below and on the right side, and *Gracilaria* most genes distribution in the figure below, this shows that the mitochondrial genome of the genus *Gracilariopsis* codon in third base on the frequency of use for T > A and G > C; The frequency of use of the third base in the mitochondrial genomic codon of *Gracilaria* is T>A. If completely affected by the mutation, then A should be equal to T and C equal to G. Therefore, by analyzing the PR2-plot plots of both, the results indicate that the use preference of codons is jointly affected by factors such as mutations and natural selection[Ошибка! Источник ссылки не найден.]. Moreover, compared with the genus *Gracilariopsis*, the scattered distribution of individual genes in the mitochondrial genome of the genus *Gracilaria* is

closer to the centerline. It can be inferred that the codon preference of the mitochondrial genome of the genus *Gracilaria* is weaker.



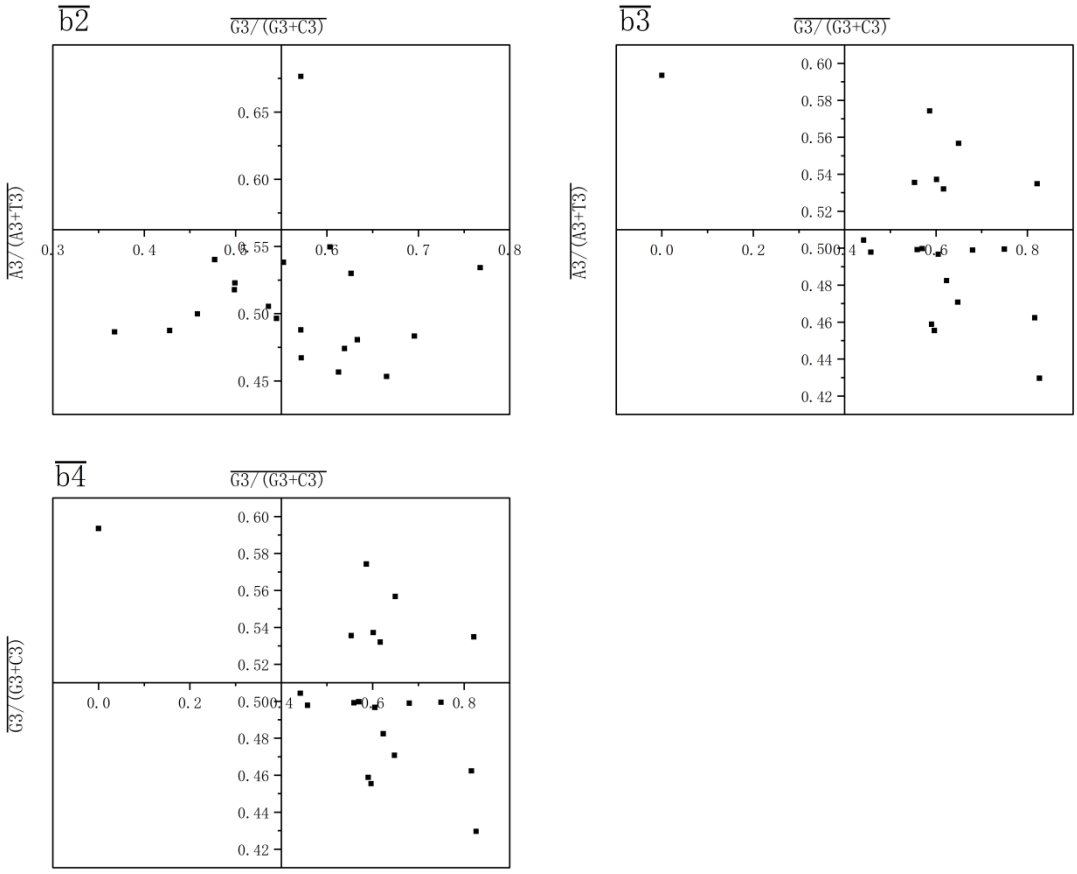


Figure 3-2 PR2-plot analysis of mitochondrial genomes of *Gracilaria* and *Gracilariopsis*

Note: a1 represents *Gracilaria chilensis*; a2 represents *Gracilaria vermiculophylla*; a3 represents *Gracilaria changii*; a4 represents *Gracilaria edulis*; a5 represents *Gracilaria salicornia*; b1 represents *Gracilariopsis lemaneiformis*; b2 represents *Gracilariopsis andersonii*; b3 represents *Gracilariopsis oryzoides*; b4 represents *Gracilariopsis chorda*

Summary of the chapter Ⅲ

1. The codon usage patterns were studied by using methods such as codon composition analysis, ENC-plot analysis and PR2-plot analysis.

2. The results show that the two types of algae have certain preferences on codon composition, part of the codon usage frequency is higher; ENC-plot analysis showed

that codon preference was affected by various factors such as gene expression levels and base composition; The PR2-plot analysis reveals the characteristics of codon usage from the perspective of base composition.

3. Through these analysis, understand the two types of algae using law of mitochondrial genome codes, for further study of gene expression regulation, and evolutionary adaptation mechanism provides important clues, and enrich the understanding of algae mitochondrial genome genetic information.

Data processing before establishment

Select the *Rhodymenia pseudopalmata* species (Genbank Registry number: NC_023252) as a group, to extract this study of 11 species (including group) after total of protein-coding genes using Gblocks online tools (http://www.phylogeny.fr/one_task.cgi?task_type=gblocks) screening conservative district. Files exported in the ".fasta "format will subsequently undergo format conversion according to the method of the established tree.

4.1.2 CONSTRUCTION OF THE BI PHYLOGENETIC TREE

The BI (Bayesian Inference) phylogenetic tree is a species evolution relationship tree constructed through the Bayesian method, which is used to display the evolutionary history and kinship among species**Ошибка! Источник ссылки не найден.**.

Using BI system evolutionary tree, specific steps are as follows: will be done before the fasta format files for data processing to get Mega open, then output to the nexus format file. Open it with Notepad and modify the opening command:

#NEXUS

[Title]

begin data;

dimensions ntax=10(The number of species to be established) nchar=4089;(Amino acid length)

format missing=? gap=- matchchar=. datatype=protein;

Matrix

Modify the end instruction:

end;

begin mrbayes;

outgroup NC_023252(outgroup);

prset aamodelpr=mixed;

mcmc ngen = 1000000 printfreq = 1000 samplefreq = 100 nchains = 4 temp = 02 savebrlens = yes;

end;

Then save the edited file. Double-click mrbayes.exe and enter the command "execute+ file name". After the operation is completed, remove 25% of the aged samples. The final phylogene tree obtained is beautified using figtree.

4.1.3 CONSTRUCTION OF THE ML PHYLOGENETIC TREE

The ML phylogenetic tree, also known as the maximum likelihood phylogenetic tree, assesses the probability of observing actual sequence data under a given evolutionary model, and seeks the tree structure and branch length that can maximize this probability to infer the evolutionary relationship between species or genes.

Specific steps are as follows: data processing done before the fasta format file with bioedit opened, and the output of phy format. Then put the.phy file in the folder "Standard-RaxML-master". In this folder, use the Command Prompt (cmd) to enter the command "raxmlHPC -m PROTGAMMAAUTO-p 12345-x 12345 -# 1000 -o The Genbank login number of the outside group -s.phy The full name of the file -f a -n "outfile0419 (outfile+ Date)" .

The ML phylogenetic tree obtained by running the program is beautified using figtree.

4.2 RESULTS AND ANALYSIS

Twelve common genes were screened out, including atp6, cox1, cox2, cox3, cytb, nad1, nad2, nad3, nad4, nad4L, nad5 and nad6. The amino acid site becomes 4089.

The length of the branch reflects the species, the genetic distance between the longer the branch, the greater the genetic differences. The algal phylogenetic tree obtained in this study was analyzed:

BI method of reconstruction of phylogenetic tree as shown in figure 4-1, the result shows that this study selected 10 species were divided into *Gracilariopsis* and *Gracilaria* belong to two branches, the branch in *Gracilaria* genera, *G. salicornia* and *G. changii* were first divided into a (posterior probability for 1), and *G. chilensis* gather again for a (posterior probability to 1), It then formed a branch with the *G. edulis* (with a posterior probability of 0.9999), and finally formed a branch with the *G. vermiculophylla* (with a posterior probability of 0.9998). Among the branches of the genus *Gracilariopsis*, *Gp. chorda* and *Gp. lemaneiformis* form one branch (posterior probability is 1), while *Gp. andersonii* and *Gp. oryzoides* form one branch (posterior probability is 1), and they are relatively closely related.

ML reconstruction of phylogenetic tree as shown in figure 4-2, phylogenetic tree topology and BI method of almost the same, different is the *G. vermiculophylla* and *Gp. lemaneiformis* is gathered for a (since the fair value of 59%), but because of the fair value is low, so can't *G. vermiculophylla* as *Gracilariopsis*. This situation might be due to the inaccurate clustering relationship of *G. vermiculophylla* caused by the different models used by the software.

But the two methods of reconstruction of phylogenetic tree by all present a family of *Gracilariopsis* and *Gracilaria* genera are divided into two independent branch, so this study for *Gracilariopsis* is independent in *Gracilaria* provides some evidence.

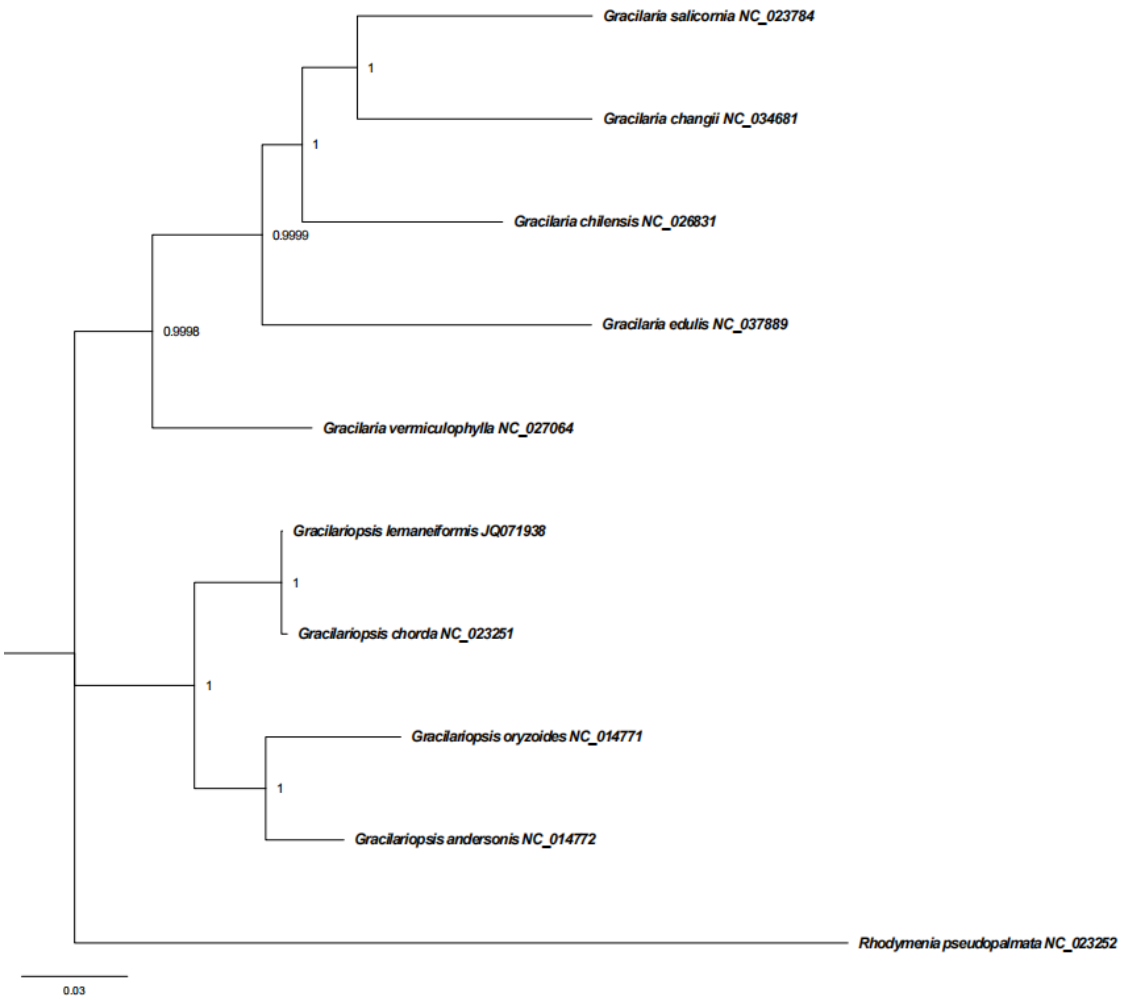


Figure 4-1 BI phylogenetic tree

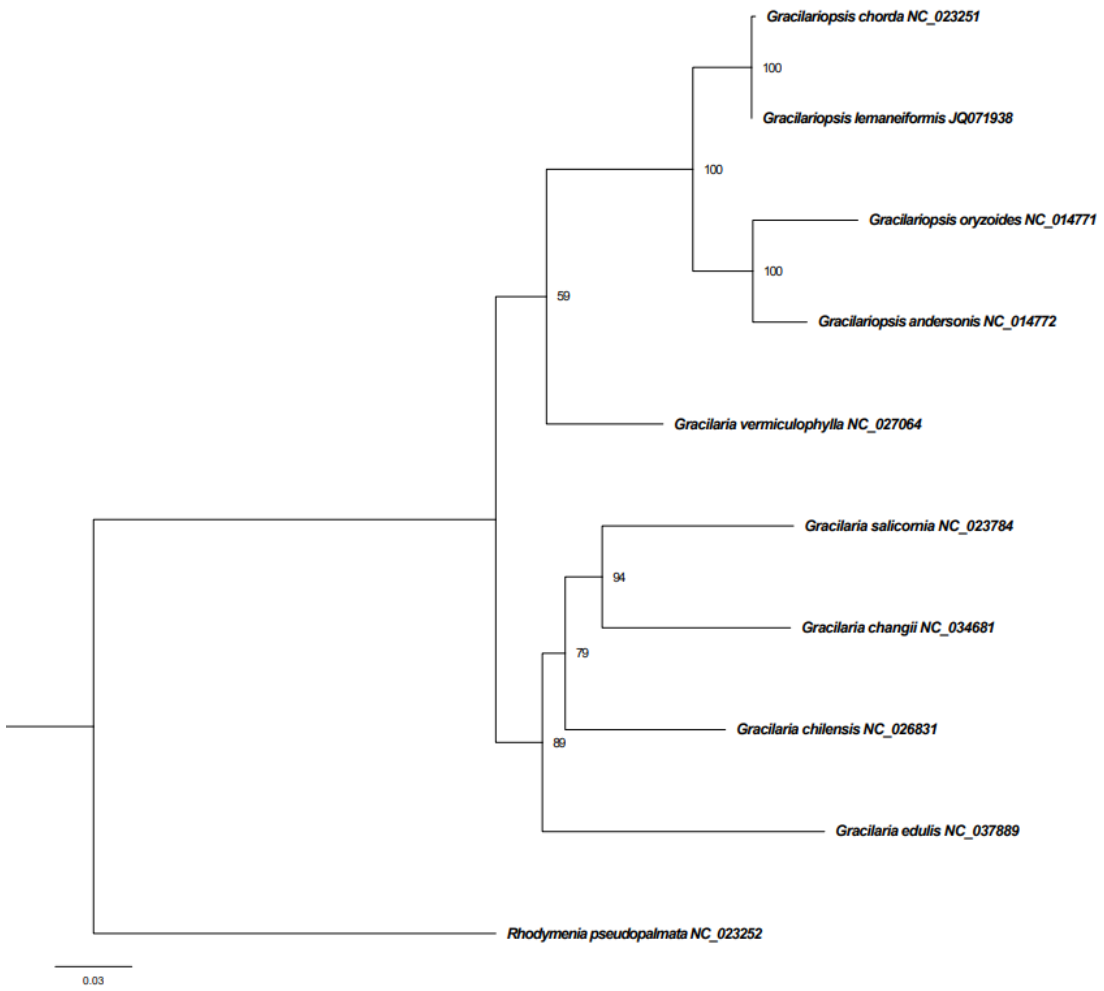


Figure 4-2 ML phylogenetic tree

Summary of the chapter Ⅳ

1. Aiming to explore the phylogenetic relationship between *Gracilaria* and *Gracilariopsis* algae, a phylogenetic tree was constructed by obtaining relevant sequences and making comparisons using ML and BI methods.

2. Results show that the constructed phylogenetic tree clearly presented *Gracilaria* and *Gracilariopsis* belong to the genetic relationship of algae, made clear its position in the red algal evolutionary lineages. This result provides molecular biological evidence for the systematic classification of *Gracilaria* and *Gracilariopsis* algae, which is helpful for understanding their evolutionary history and species differentiation mechanisms. At the

same time, it also offers an important reference basis for subsequent research on the evolutionary relationship and biodiversity of red algae.

Conclusion

Use of biological information analysis software, this study on the molecular level of *Gracilaria* and *Gracilariopsis* is a total of 10 species of mitochondrial genome has carried on the comparative analysis, the following conclusions:

(1) genome features: The mitochondrial genomes of *Gracilaria* and *Gracilariopsis* have similar characteristics, and the number and types of genes they contain are also similar. There are subtle differences in tRNA genes, rRNA genes and orf genes. Compared with the genus *Gracilaria*, the orf in the genus *Gracilariopsis* is more diverse, which is reflected in the differences in number and variety. Deletions, duplications or rearrangements of tRNA often occur in the trnN-trnA rich region of the mitochondrial genomes of the two genera. Moreover, a rearrangement phenomenon was found in the trnH gene of the mitochondrial genome of the genus *Gracilariopsis*, which can be used as evidence to distinguish the two genera.

(2) Both mitochondrial genome codon to unrael the differences between related parameters. The ENC values of *Gracilaria* and *Gracilariopsis* both exceeded 35, indicating that the mitochondrial genomes of both showed a relatively weak preference in the use of codons. In addition, the parameters such as CAI, CBI, and Fop of the mitochondrial genomes of *Gracilaria* and *Gracilariopsis* are slightly different, which indicates that the encoded proteins may have hydrophobicity.

(3) This study reconstruction of phylogenetic tree showed that *Gracilariopsis* and *Gracilaria* genera are divided into two, for asparagus is independent in greasewood belong to provide some evidence. *Gracilaria* and *Gracilariopsis* in the mitochondrial genome genetic composition, arrangement and evolutionary characteristics vary, for algae system development, and provided the foundation for molecular genetic diversity research.

In the future, the sample size can be expanded to add more species and closely related genus samples within the two genera, construct a more complete phylogenetic tree, and clarify the evolutionary relationships of mitochondrial genomes within and

between genera. Combining the nuclear genome to analyze the inter-genus evolutionary relationship, and combining the nuclear genome and chloroplast genome data to carry out multi-genome joint analysis, comprehensively analyze the evolutionary dynamics of algae and the co-evolution laws of the genome; Deeply explore mitochondrial functional genes and investigate their roles in environmental adaptation. The development of efficient molecular identification tools using mitochondrial markers facilitates the selection and breeding of superior varieties of *Gracilariopsis* and the ecological restoration of *Gracilaria* algae.

This field of research will provide more for algae biology and biotechnology development rich theoretical support, to promote the depth of the research and application based on practice, push related research on algae genetic breeding and conservation.

REFERENCE

1. Nan Fangru. Systematic Classification and Evolutionary History of Different Lineages of Freshwater Red Algae [D]. Shanxi University, 2017.
2. Zhao Huiying. Phylogenetic Analysis of Red Algae and Capsule Algae Endosymbiotic Systems Based on Chloroplast Genomes [D]. Shanxi University, 2024.
3. Zhang Hongxia. Preliminary Analysis of Antibacterial Substances in HT-6 Induced by Pathogenic Phytophthora and Their Disease Prevention Effects [D]. Hebei University, 2020. DOI: 10.27103/d.cnki.ghebu.2020.001013.
4. Qian Hao. Complete Mitochondrial Genome Amplification and Systematic Evolution Analysis of *Gracilaria* vermiculophylla [D]. Ocean University of China, 2013.
5. Li Tingting, Chen Bin, Chen Shengping, et al. Molecular Systematic Analysis of ITS Sequences of 5 Species of Sea Weeds in the Families *Gracilaria*ceae and Axonema [J]. Journal of Sun Yat-sen University (Natural Science Edition), 2012, 51(04): 97-105.
6. Guo Weihua. Construction of Mapping Populations of *Gracilariopsis* lemaneiformis and Screening of Parental SSR and RSAP Molecular Markers [D]. Ocean University of China, 2014.
7. Liu Chenlin, Li Yan, Wang Xiao, et al. DNA Barcoding Records of Several Gelatinous Red Algae in Dongshan Bay, Fujian Province [J]. Advances in Marine Science, 2024, 42(03): 568-577.
8. Wang Yanping. Preliminary Study on the Experimental Ecology of *Gracilaria* [D]. Guangdong Ocean University, 2017.
9. Yang Fang. Study on the Growth and Nitrogen Nutritional Characteristics of Three Branching Varieties of Fine Basal *Gracilaria* Populations [D]. Guangxi University, 2014.

10. Fu Feng. Breeding of Ultraviolet-induced Varieties and Selection of Drought-resistant Rapid-Growing Excellent Strains of *Gracilaria* [D]. Ocean University of China, 2013.
11. Qiang Xi, Wang Xuelai, Niu Jianfeng, et al. Differences in Algal Phycobilin and Agar Content in Different Germplasm Resources of *Gracilaria* and Comprehensive Extraction Process Research [J]. Marine Science, 2022, 46(11): 107-115.
12. Cui Zhenhao, Tang Wenting, Zhang Xiaochuan, et al. Prokaryotic Expression, Enzymological Characteristics and Response Analysis of Phospholipase D from *Gracilariopsis* [J/O]. Marine Science, 1-16 [2025-02-14].
13. Li Nenghui, Huang Qing, Li Hang, et al. Taxonomic Study of Four *Gracilaria* Species in Zhanjiang Sea Area Based on Morphological and Molecular Data [J]. Tropical Marine Journal, 2024, 43(02): 34-47.
14. Zhao Xiaobo, Pang Shaojun, Liu Feng. Systematic Analysis of Three Molecular Sequences of Several *Gracilaria* Sea Weeds [J]. Marine Science, 2013, 37(06): 8-15.
15. Zhao Xiaobo. Research and Application of Gene Barcoding (DNA Barcode) in Marine Algae Systematics [D]. Graduate School of Ocean University of China (Institute of Oceanography), 2013.
16. Li Nenghui, Huang Qing, Li Hang, et al. Taxonomic Study of Four *Gracilaria* Species in Zhanjiang Sea Area Based on Morphological and Molecular Data [J]. Tropical Marine Journal, 2024, 43(02): 34-47.
17. Ding Lanping, Yang Nan, Liu Jinmei, et al. Current Research Status and Existing Problems of Classification of Marine Economic Red Algae in *Gracilaria*ceae [J]. Journal of Applied Oceanography, 2020, 39(01): 144-151.
18. Ye Feiyu, Wei Yakang, Wang Guifeng. Structure and Assembly Path of Mitochondrial Respiratory Chain Complex I [J]. Chinese Journal of Biochemistry and Molecular Biology, 2022(038-001). DOI: 10.13865/j.cnki.cjbmb.2021.04.1696.

19. Li Tianyong. Study on the Complete Mitochondrial Genome of Undaria pinnatifida and Comparative Analysis of Algal Mitochondrial Genomes [D]. Ocean University of China, 2014.
20. Yu Haibin. The Effect of Mitochondrial Glycerol-3-Phosphate Acetyltransferase (GPAM) Gene Silencing on Milk Fat Metabolism in Cattle [D]. Jilin University, 2020. DOI: 10.27162/d.cnki.gjlin.2020.007254.
21. Guo Zhonglong, Yuan Minglong. Advances in Mitochondrial Genomics of Hemiptera Insects [J]. Science China: Life Sciences, 2016, 46(02): 151-166.
22. Ran W, Zhao Y, Chen J, et al. New Complete Mitogenomes of Four Species in Tribe Erythroneurini (Hemiptera: Cicadellidae: Typhlocybinae) from China and Their Phylogenetic Analysis [J]. Journal of Asia-Pacific Entomology, 2025, 28(1): 102378-102378.
23. Yin Jing. Research on the Protective Mechanism of Tongping Yangxin Formula against H9c2 Cell Injury Induced by High Glucose [D]. Hebei United University, 2014.
24. Xia Yu. Cell Organelle Genome and Red Algae Phylogenetic Study of Hypnea cervicornis [D]. Qilu University of Technology, 2023.
25. Liu Tao. Illustrated Guide to Common Large Seaweeds in the South China Sea [M]. Beijing: Ocean Press, 2017.
26. Mao Shuo. A Brief Discussion on the Origin of Mitochondria and Chloroplasts [J]. Middle School Biology Teaching, 2019, (04): 46-48.
27. Tao Yong. Discussing Core Competencies of High School Biology [J]. Middle School Biology Teaching, 2018, 000(006): 68-70.
28. Dai Chunyan. Practical Research on the Construction of Structure and Function View in High School Biology Teaching [D]. Nanjing Normal University, 2021. DOI: 10.27245/d.cnki.gnjsu.2021.000834.
29. A S M ,G J W ,J A R , et al.The Origin of Mitochondrial Cristae from Alphaproteobacteria.[J].Molecular biology and evolution,2017,34(4):943-956.

30. Yuan Xiaochu. Discussion on Jianglii Mitochondrial Genome and Management of Aquatic Genetic Resources [D]. Shandong: Ocean University of China, 2020.
31. He Baoli, Jiao Jianlin, Chen Liling, et al. Study on the Growth Hormone Gene Polymorphism of Diannan Small-Eared Pig [J]. Journal of Kunming Medical University, 2012, 33(11): 23-25.
32. Li Yucui, Zhang Chunbao, Zhao Hongkun, et al. Extraction Method of High-Purity Mitochondrial Genome DNA of Soybean Suitable for Genome Sequencing [J]. Genomics and Applied Biology, 2015, 34(08): 1734-1738. DOI: 10.13417/j.gab.034.001734.
33. Lam H T T ,Dinh M Q ,Nguyen D H T , et al.Genetic diversity and phylogenetic tree of Butis genus in the Vietnamese Mekong Delta based on mitochondrial cytochrome b gene sequences[J].The Egyptian Journal of Aquatic Research,2024,50（4）:498-505.
34. Li Ying. Transcriptome and Mitochondrial Genome Analysis of Wheat Type B Cytoplasmic Incompatibility Lines [D]. Northwest A&F University, 2022. DOI: 10.27409/d.cnki.gxbnu.2022.000092.
35. Li En. Discussion on the Phylogenetic Relationship of Snake Suborder Based on Mitochondrial Genome [D]. Anhui Normal University, 2014.
36. Chang Meihui, Chu Xiaoyu, Wu Ningfeng. Research Progress on the Effect of Translation Rate on Recombinant Protein Expression [J]. China Agricultural Science and Technology News, 2017, 19(04): 45-50. DOI: 10.13304/j.nykjdb.2016.483.
37. Yang Ping. Analysis of Codon Usage Characteristics of NHX Genes in Arabidopsis thaliana [J]. Journal of Nanjing Xiaozhuang University, 2011, 27(06): 80-84.
38. Li Tingting , Ma Zheng , Ding Tiemei , et al. Codon usage bias and phylogenetic analysis of chloroplast genome in 36 *Gracilaria*ceae species[J]. Functional & Integrative Genomics, 2024, 24 （2）: 45-45.

39. Bao Guoyuan, Li Wenxin, Yang Xingguang, et al. Analysis of Codon Usage Preferences in Mitochondrial and Chloroplast Genomes of Sea Beet [J]. Jiangsu Journal of Agricultural Sciences, 2023, 39(09): 1804-1817.
40. Yang Xiangyan, Cai Yuanbao, Tan Qinliang, et al. Analysis of Codon Preference in Chloroplast Genomes of Pineapple [J]. Tropical Crop Journal, 2022, 43(3): 8. DOI: 10.3969/j.issn.1000-2561.2022.03.001.
41. Lin Xinmao, Huang Jiaquan. Analysis of Codon Preference in Chloroplast and Nuclear Genomes of Cactus Family [J]. Molecular Plant Breeding, 2024, 22(22): 7400-7412. DOI: 10.13271/j.mpb.022.007400.
42. Mao Liyan, Huang Qiwei, Long Lingyun, et al. Analysis of Codon Preference in Chloroplast Genomes of 7 Species of Nymphaea [J]. Journal of Northwest Forestry College, 2022, 37 (02): 98-107.
43. Xie Liuyi, Ren Xuan, Ying Wenbo, et al. Comparative Analysis of Codon Preference in Chloroplast Genomes of Cultivated Barley and Wild Barley [J]. Journal of Yunnan Nationalities University (Natural Science Edition), 1-13.
44. Zhao Guohe. Preliminary Study on Codon Site Characteristics and Comparison of Evolutionary Dynamics of Histone Genes [D]. Nanjing Normal University, 2007.
45. Qiu Chuntao, Liang Fangmei, Lu Ying, et al. Analysis of Codon Preference in Mitochondrial Genomes of Fine Grained Sea Weed [J]. Molecular Plant Breeding, 2021, 1-17.
46. Zhang Yang, Sun Shuguang, Li Qing, et al. Analysis of Codon Usage Preferences in Chloroplast Genomes of Lettuce [J]. Henan Agricultural Sciences, 2022, 51(10): 114-124. DOI: 10.15933/j.cnki.1004.
47. Li Jiangfei, Li Yaqi, Tang Junrong, et al. Codon Preference Pattern of Chloroplast Genomes of Mountain Pine [J]. Journal of Biology, 2023, 40(01): 52-59.
48. Dai Jiangpeng, Cai Yiming, Liu Qiaozhen, et al. Analysis of Codon Preference in Chloroplast Genomes of 7 Species of Glycyrrhiza [J]. Chinese Herbal Medicines, 2023, 54(9): 2907-2916. DOI: 10.7501/j.issn.0253-2670.2023.09.023.

49. Gao Can, Fan Zhi Feng, Ma Changle. Analysis of Codon Preference in Chloroplast Genomes of Huangyao Dahua [J]. Journal of Southwest Forestry University (Natural Science), 2023, 43(05): 66-76.
50. Long Lingyun, Huang Qiwei, Li Huimin, et al. Analysis of Codon Usage Preferences in Chloroplast Genomes of Golden Orange [Available Online]. Molecular Plant Breeding.
51. Zhao Qinghao. Phylogenetic and Biogeographic Research on Antaeidae Subfamily [D]. Shanghai Normal University, 2021. DOI: 10.27312/d.cnki.gshsu.2021.000948.
52. Cui Yutong. Research on 8 Species of Sargassum Mitochondrial Genomes and Construction of Algae Organelle Genome Database [D]. Yantai University, 2021. DOI: 10.27437/d.cnki.gytdu.2021.000260.