



The complete plastid genomes of *Betaphycus gelatinus*, *Eucheuma denticulatum*, and *Kappaphycus striatus* (Solieriaceae: Rhodophyta) and their phylogenetic analysis

Jing Zhang¹ · Na Liu^{2,3,4} · Maria Dyah Nur Meinita⁵ · Xumin Wang⁶ · Xianming Tang⁷ · Guoliang Wang⁸ · Yuemei Jin² · Tao Liu²

Received: 16 August 2019 / Revised and accepted: 27 March 2020
© Springer Nature B.V. 2020

Abstract

Betaphycus, *Eucheuma* and *Kappaphycus* (Gigartinales, Florideophyceae) are the most commercially important genera of the family Solieriaceae that produce carrageenan. Here, three complete plastid genomes of *Betaphycus gelatinus*, *Eucheuma denticulatum* and *Kappaphycus striatus* were fully sequenced using next-generation sequencing technology. Genome organizations and gene contents of the three plastid genomes were highly alike. They all had circular mapping organizations and the sizes were 178,394 bp (*B. gelatinus*), 177,003 bp (*E. denticulatum*), and 176,763 bp (*K. striatus*). They encoded almost the same set of plastid genes (238–240), including 202 to 204 protein-encoding genes, 30 transfer RNA genes (tRNAs), 3 ribosomal RNA genes (rRNAs), 2 misc_RNAs (*ffs*, *rnpB*), and 1 transfer-messenger RNA gene (tmRNA). One group II intron interrupting the *trnM* gene was identified in each of these three plastid genomes. Other three plastid genomes from species of the order Gigartinales including *Kappaphycus alvarezii*, *Chondrus crispus* and *Mastocarpus papillatus* have been reported. The plastid genome organization at the level of the order Gigartinales was highly conserved. Co-linear analysis among the six plastid genomes of the Gigartinales showed the considerable sequence synteny with the exception of one remarkable gene rearrangement. The approximately 12.5-kb gene fragment from gene *psaM* to *yef21* in plastid genomes of the four species of the Solieriaceae was completely reversed compared to that of *M. papillatus* and *C. crispus*. It might be used as the potential phylogenetic markers uniting the species of the Solieriaceae. In addition, phylogenetic analysis based on 138 shared protein-encoding genes from 53 Florideophyceae plastid genomes indicated all species were clearly divided into five clades corresponding to their subclasses. The results suggested there was a non-monophyletic relationship of the order Gigartinales. Four species of the family Solieriaceae formed one clade and *E. denticulatum* was basal relative to the other three species. The novel plastid genomes expand the available plastid pool for red algae which would facilitate the phylogenetic study in algae.

Keywords Rhodophyta · Co-linear analysis · Comparative analysis · Phylogenetic analysis · Plastid genome · Solieriaceae

Jing Zhang and Na Liu contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10811-020-02120-5>) contains supplementary material, which is available to authorized users.

✉ Tao Liu
liutao@ouc.edu.cn

¹ School of Bioengineering, Qilu University of Technology (Shandong Academy of Sciences), 3501 Daxue Road, Jinan 250353, Shandong Province, China

² College of Marine Life Sciences, Ocean University of China, 5 Yushan Road, Qingdao 266003, Shandong Province, China

³ College of Life Sciences, Beijing Normal University, 19 Xijiekouwai Road, Beijing 100875, China

⁴ National Institute of Biological Sciences, 7 Kexueyuan Road, Beijing 102206, China

⁵ Fisheries and Marine Science, Jenderal Soedirman University, Purwokerto 53122, Indonesia

⁶ College of Life Sciences, Yantai University, 30 Qingshan Road, Yantai 264005, Shandong Province, China

⁷ Hainan Academy of Ocean and Fisheries Sciences, 54 Haifu Road, Haikou 570203, Hainan Province, China

⁸ Beijing Academy of Agriculture and Forestry Sciences, 11 Shuguanghuayuan Road, Beijing 100097, China

Introduction

Species of the family Solieriaceae belonging to the order Gigartinales are mostly distributed in the warm-temperate and tropical waters throughout the world (<http://www.algaebase.org/>, Guiry and Guiry 2020). They provide important raw materials to extract carrageenan (Pereira et al. 2015), which has been widely used in food and chemical industries. In the Solieriaceae, *Betaphycus*, *Eucheuma* and *Kappaphycus* mainly producing β -carrageenan, ι -carrageenan, and κ -carrageenan, respectively (Pereira et al. 2015), are the most important economic genera (Li et al. 2018). They have drawn much attention in recent years and many studies on culture techniques (Adnan and Porse 1987; de Góes and Reis 2011), genetic diversity (Liu et al. 2009; Zhang et al. 2011; Lim et al. 2014), and phylogenetic analysis (Fredericq et al. 1999; Dumilag et al. 2014; Li et al. 2018) have been carried out.

Species from *Betaphycus*, *Eucheuma* and *Kappaphycus* are short of diagnostic morphological features due to their high phenotypic plasticity in response to different environments and culture times (Doty and Norris 1985). Therefore, all these three genera are facing similar taxonomical challenges. Recently, DNA sequences from nucleus (ITS and 28S rRNA), plastid (*rbcL* and *rbcS*), and mitochondrion (*cox1*, *cox2–3* spacer) have been utilized for the phylogenetic systematics of *Betaphycus*, *Eucheuma* and *Kappaphycus* (Fredericq et al. 1999; Conklin et al. 2009; Liu et al. 2012; Tan et al. 2013; Dumilag et al. 2014; Sun et al. 2014; Lim et al. 2017). However, the previous works have been based on limited sequences which might not fully explore the evolutionary relationship of the family Solieriaceae. Compared to a fraction of genes, complete genomes contained adequate phylogenetic information and could be used to resolve complex phylogenetic relationships (Zhou et al. 2012; Janouškovec et al. 2013; Ma et al. 2014). Therefore, comprehensive surveys at the level of the complete genome are needed.

To our knowledge, organellar DNAs are characterized by uniparental inheritance and compact genome structure compared to the nuclear genomes, thus they become the efficient molecular tools for evolutionary studies. Plastids carrying genetic information are semi-autonomous photosynthetic organelles. Plastids of red algae are descendants of a primary endosymbiotic relationship in which a cyanobacterium was engulfed by a heterotrophic eukaryote (Reyes-Prieto et al. 2007). Therefore, plastid genomes could provide valuable sequence information to trace the origin and evolution of photosynthesis in eukaryotes. Additionally, the smaller genome size and higher copy number per cell than nuclear genomes have made the plastid genomes easier for sequencing. Based on these characters, plastid genomes have been recognized as an effective approach to clarifying phylogenetic relationships (Gao et al. 2010). Red algal plastid genomes are particularly

interesting due to the large gene content, compact organization, and comparatively slow evolution. This suggests that they might represent the ancestral state of primary plastids (Janouškovec et al. 2013). Several recent surveys have used plastid genomes to carry out the reconstructions and analysis of phylogenetic relationships for red algae (Wang et al. 2013a; Liu et al. 2018, 2019; Zhang et al. 2018).

As the next-generation sequencing technologies become less expensive and more efficient (Heather and Chain 2016), more and more sequencing projects of complete plastid genomes have been set up. Currently, nearly 300 plastid genomes are available among the algae and they are listed in Table S1. However, only three of them were from species of the order Gigartinales including one each from three families Solieriaceae (*Kappaphycus alvarezii*), Gigartinaceae (*Chondrus crispus*) and Phylloporaceae (*Mastocarpus papillatus*) (Collén et al. 2013; Sissinia et al. 2016; Liu et al. 2019). Here, we first characterized the complete plastid genomes of *Betaphycus gelatinus*, *Eucheuma denticulatum* and *Kappaphycus striatus* and expanded the available plastid pool of red algae. We obtained information about gene content, genome structure, and organization of these plastid genomes, and compared them with other available plastid genomes of Gigartinales species. Further, we used this plastid genome data for reconstruction of phylogenetic relationships and taxonomic study about Florideophyceae species with emphasis on the relationships of the family Solieriaceae within the order Gigartinales. The plastid genome information provided a promising resource to conduct the molecular investigations and study the phylogenetics of red algae.

Materials and methods

Sample collection and DNA extraction

Fresh thalli of *Betaphycus gelatinus* (specimen number: 2016050103, Fig. S1a) and *Eucheuma denticulatum* (specimen number: 2016070040, Fig. S1b) were collected from Lingshui County, Hainan Province, China (18° 26' 8" N, 110° 4' 1" E) and thalli of *Kappaphycus striatus* (specimen number: 2016100062, Fig. S1c) were collected from Takalar, Indonesia (5° 23' 46" S, 119° 27' 6" E). Thalli were cultivated at 24–26 °C in sterilized filtered seawater under fluorescent light (80–110 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$; 12 h light/dark cycles). Thalli were washed in autoclaved seawater and dried with paper towels before use. These specimens were all stored in the Culture Collection of Seaweed at Ocean University of China.

Genomic DNA was extracted from approximately 2 g of fresh tissue using the improved cetyltrimethylammonium bromide (CTAB) method (Sun et al. 2011). β -Mercaptoethanol was used to inhibit the oxidation of phenols in the samples and

two concentrations of CTAB (2% and 10%) were added to remove polysaccharides. The quality and quantity of DNA were detected using NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific, USA).

High-throughput sequencing and assembly

Three short-insert libraries were constructed using approximately 5 µg of purified DNA for each species following the instructions of manufacturer (Illumina Inc., USA). Approximately 9 Gb of paired-end sequence data was randomly extracted from the total sequencing output, as input into SOAPdenovo (Luo et al. 2012) with default assembly parameters to assemble the plastid genomes. Two plastid genomes from *Chondrus crispus* (GenBank accession number: NC_020795) and *Kappaphycus alvarezii* (GenBank accession number: KU892652) were used as the reference sequences to determine the proportion of plastid-related contigs. Subsequently, all plastid-related contigs were aligned and ordered into a circular structure using CodonCode Aligner (CodonCode Corporation, USA) (Liu et al. 2019).

Annotation and comparative genome analysis

The protein-encoding genes, rRNA genes, introns, and tmRNA gene of the three species were annotated based on those of *C. crispus* and *K. alvarezii* using Geneious R10 (Biomatters Ltd., New Zealand; available from <http://www.geneious.com/>). The tRNA genes were predicted using tRNAscan-SE version 1.21 (<http://lowelab.ucsc.edu/tRNAscan-SE/>) (Schattner et al. 2005). The physical maps of these three plastid genomes were obtained using Organellar Genome DRAW (OGDRAW) version 1.3.1 (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>) (Greiner et al. 2019). Sequence alignment and base composition were conducted using MEGA 6.0 (Tamura et al. 2013).

The plastid genomes of six reported species from the order Gigartinales were aligned using progressive Mauve genome aligner at the default settings with Geneious software (Kearse et al. 2012). Comparison of the plastid genomes among six species from the order Gigartinales and *Riquetophycus* sp. from the order Peyssonneliales was made using Mauve in order to better clarify the evolutionary relationship of four species from the family Solieriaceae within the order Gigartinales.

Phylogenetic analysis

Phylogenetic analysis of Florideophyceae was conducted using 138 shared plastid protein-encoding genes from 50 plastid genomes available at the GenBank database and three plastid genomes obtained in this work. *Cyanidioschyzon merolae* (GenBank accession number: NC_004799) served

as the outgroup. Each protein sequence was aligned individually using MEGA 6.0 and then the entire concatenated alignment was generated manually using BioEdit (Hall 1999). The concatenated protein alignments with conserved regions were generated and poorly conserved regions were removed using the Gblocks server (http://phylogeny.lirmm.fr/phylo.cgi/one_task.cgi?task_type=gblocks) (Castresana 2000), then the alignment was reduced from the original 33,839 positions to 26,849. To reconstruct the phylogenetic tree, the best-fit model for maximum-likelihood (ML) was selected using ProtTest 3.4.2 (Darriba et al. 2011). ML analysis was conducted using RAxML (Stamatakis 2006) with 1000 replications under the CpREV + G + I + F model. The Bayesian inference (BI) was performed based on MrBayes version 3.1.2 (Huelsenbeck and Ronquist 2001) using CpREV model. The phylogenetic analysis was performed using two independent runs with four Markov Chains, which ran for 1,000,000 generations until the average standard deviation of split frequencies was below 0.01. Output trees were sampled every 100 generations. In addition, the first 25% of samples was removed as burn-in. The remaining trees were used to build a 50% majority rule consensus tree accompanied with posterior probability values. FigTree version 1.3.1 (<http://tree.bio.ed.ac.uk/>) was used to display and edit the phylogenetic tree (Rambaut 2009).

To further elaborate the phylogenetic relationship of *Betaphycus*, *Eucheuma* and *Kappaphycus*, seven plastid genes (*atpB*, *infB*, *psbA*, *psbB*, *rbcL*, *rps7*, *rps12*) combined with six mitochondrial genes (*atp6*, *cob*, *cox1*, *nad1*, *nad4*, *nad5*) were employed to reconstruct phylogenetic trees by ML and BI methods. *Riquetophycus* sp. was used as the outgroup.

Results

Genome features

The plastid genomes of *B. gelatinus*, *E. denticulatum* and *K. striatus* were assembled as a single circular molecule with the length of 178,394 bp, 177,003 bp, and 176,763 bp, respectively. The average GC contents of the plastid genomes were 28.93% (*B. gelatinus*), 29.61% (*E. denticulatum*), and 29.35% (*K. striatus*), which are similar to those previously reported Gigartinales species. In general, Florideophyceae species have large plastid genomes ranged from 91 to 194 kb in size and lower GC contents of about 29.00%.

All of the three complete plastid genomes were submitted to GenBank under the following accession numbers: MN240356 (*B. gelatinus*), MN240357 (*E. denticulatum*) and MN240358 (*K. striatus*). General features of the six reported complete plastid genomes from the order Gigartinales including four species from Solieriaceae (*B. gelatinus*, *E. denticulatum*, *K. striatus* and *K. alvarezii*), one species

from Gigartinales (*C. crispus*) and one species from Phyllophoraceae (*M. papillatus*) are summarized in Table 1.

The three plastid genomes were predicted to encode a total of 238–240 genes including 202–204 protein-encoding genes, 30 tRNA genes, 3 rRNA genes, 2 misc_RNAs (*ffs*, *rnpB*), and 1 tmRNA gene (*ssrA*), which are in the same range as those of reported Gigartinales species (Table 1). The only difference in gene content among the three genomes was the absence of two protein-encoding genes (*pbsA*, *ycf57*) in *K. striatus*, and which are present in *B. gelatinus* and *E. denticulatum*. Genes were encoded on both the heavy and light strands and each strand in two opposite directions had approximately the same encoding proportion (Fig. 1). All tRNA genes ranged from 71 to 89 bp in size, and one group II intron interrupting the *tmMe* gene was identified in all the three plastid genomes. For these three plastid genomes, all of the rRNA genes retained a single copy, which was in accordance with the majority of Florideophyceae species. Compared with previously

published plastid genomes of the Gigartinales species, there was a high conservation in gene content and genome organization. Only the plastid genome of *M. papillatus* showed a smaller gene content (234 genes), tmRNA, misc_RNA, or intron was not identified.

The coding sequence regions in the plastid genomes of *B. gelatinus*, *E. denticulatum* and *K. striatus* ranged from 153,570 bp to 154,927 bp, accounting for 87.96%, 87.41%, and 86.88% of the corresponding plastid genome. The total spacer sizes in *B. gelatinus*, *E. denticulatum*, and *K. striatus* plastid genomes were 19,497 bp, 20,284 bp, and 21,226 bp which were smaller than *C. crispus* (24,270 bp) and *M. papillatus* (32,023 bp). Together with *K. alvarezii*, the four plastid genomes from the family Solieriaceae had more compact organization. Moreover, the general organization of *K. striatus* plastid genome was approximately the same as that of congeneric *K. alvarezii* by overall comparison. Insertions or deletions

Table 1 General features of the complete plastid genomes available in the order Gigartinales

	<i>B. gelatinus</i> *	<i>E. denticulatum</i> *	<i>K. striatus</i> *	<i>K. alvarezii</i> #	<i>C. crispus</i> #	<i>M. papillatus</i> #
			Solieriaceae		Gigartinales	Phyllophoraceae
Accession number	MN240356	MN240357	MN240358	KU892652	NC_020795	NC_031167
Genome size (bp)	178,394	177,003	176,763	178,205	180,086	184,382
Overall G + C content (%)	28.93	29.61	29.35	29.56	28.73	29.05
Predicted coding sequence (%)	87.96	87.41	86.88	86.19	85.44	82.63
Pairs of gene overlaps	11	11	10	10	7	8
Total spacer size (bp)	19,497	20,284	21,226	22,644	24,270	32,023
Predicted total gene content	240	240	238	238	240	234
rRNA/tRNA/CDS content	3/30/204	3/30/204	3/30/202	3/30/202	3/30/204	3/29/202
Intron number	1	1	1	1	1	0
tmRNA	1	1	1	1	1	0
misc_RNA	2	2	2	2	2	0
Start codons: ATG	193	191	189	189	190	200
Start codons: GTG	8	10	10	10	9	1
	(<i>chlI</i> , <i>infC</i> , <i>petN</i> , <i>psbC</i> , <i>rbcS</i> , <i>rpl3</i> , <i>rps8</i> , <i>ycf63</i>)	(<i>atpD</i> , <i>chlI</i> , <i>infC</i> , <i>petN</i> , <i>psbC</i> , <i>rbcS</i> , <i>rpl24</i> , <i>rps8</i> , <i>ycf91</i> , <i>ycf63</i>)	(<i>chlI</i> , <i>infC</i> , <i>petN</i> , <i>psbC</i> , <i>rbcS</i> , <i>rpl3</i> , <i>rpl24</i> , <i>rps8</i> , <i>ycf63</i> , <i>ycf64</i>)	(<i>chlI</i> , <i>infC</i> , <i>petN</i> , <i>psbC</i> , <i>rbcS</i> , <i>rpl3</i> , <i>rpl24</i> , <i>rps8</i> , <i>ycf63</i> , <i>ycf64</i>)	(<i>chlI</i> , <i>infC</i> , <i>psbC</i> , <i>rbcS</i> , <i>rpl3</i> , <i>rps8</i> , <i>trpG</i> , <i>ycf37</i> , <i>ycf63</i>)	(<i>infC</i>)
Start codons: TTG	3	3	3	3	5	0
	(<i>ycf27</i> , <i>ycf57</i> , <i>ycf20</i>)	(<i>ycf27</i> , <i>orf712</i> , <i>trxA</i>)	(<i>ycf27</i> , <i>trxA</i> , <i>ycf20</i>)	(<i>ycf27</i> , <i>trxA</i> , <i>ycf20</i>)	(<i>Orf64</i> , <i>ycf27</i> , <i>ycf20</i> , <i>trxA</i> , <i>ycf65</i>)	
Start codons: ATA	0	0	0	0	0	1(<i>ycf86</i>)
Stop codons: TAA	152	140	142	142	158	150
Stop codons: TAG	36	44	39	39	33	33
Stop codons: TGA	16	20	21	21	13	19
Reference	This work	This work	This work	Liu et al. 2019	Janouskovec et al. 2013	Sissinia et al. 2016

*Obtained in this work

To ensure the comparisons with accuracy and consistency, the plastid genome sequences of *K. alvarezii*, *C. crispus* and *M. papillatus* have been re-annotated and reanalyzed with the same method as those in the plastid genomes of *B. gelatinus*, *E. denticulatum* and *K. striatus*

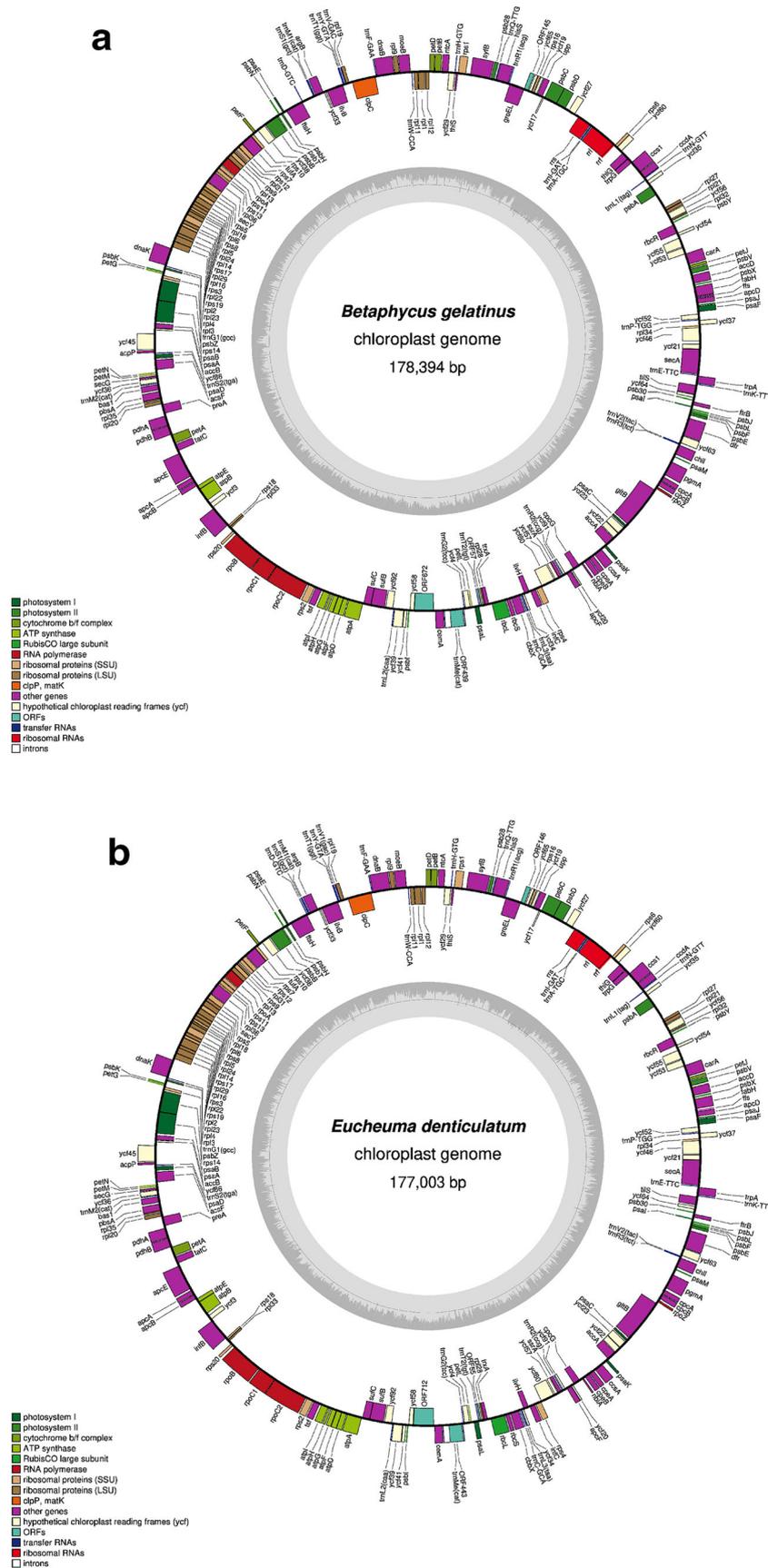


Fig. 1 Gene maps of plastid genomes of *B. gelatinus* (a), *E. denticulatum* (b), and *K. striatus* (c). Genes on the outside of the maps are transcribed in a clockwise direction, whereas those on the inside of the maps are transcribed counterclockwise

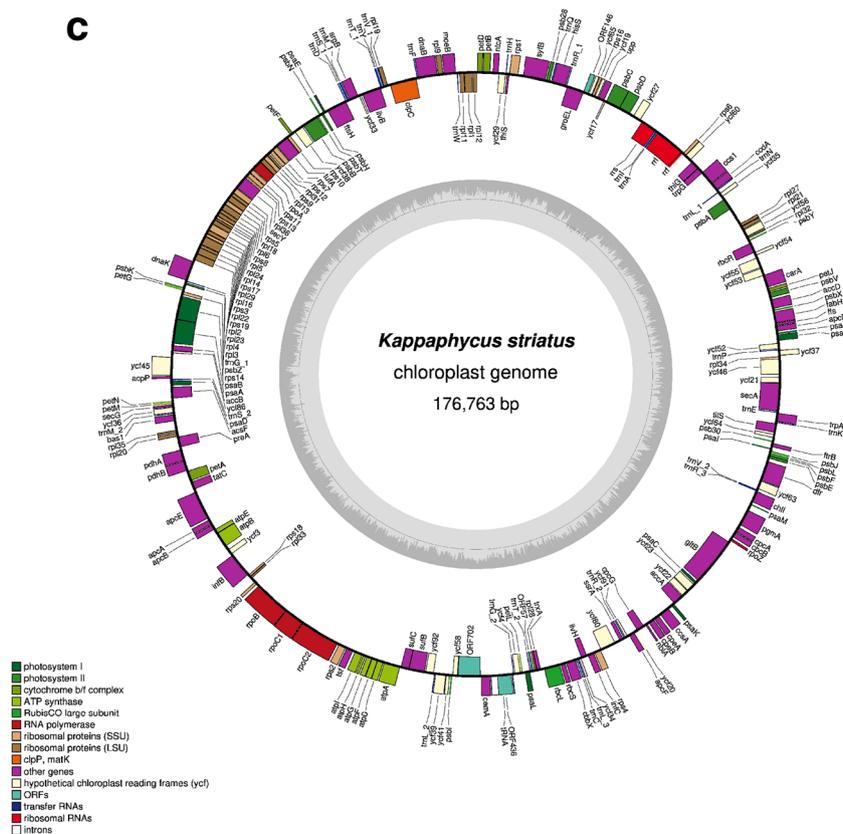


Fig. 1 (continued)

were rare, except for the existence of 1426 nucleotides, one intergenic region which was located between *rrs* and *ycf27* in the plastid genome of *K. alvarezii*. The total intergenic region size of *K. striatus* was 21,226 bp whereas that of *K. alvarezii* was 22,644 bp. The lengths of these two plastid genomes varied mainly due to the difference of the intergenic region.

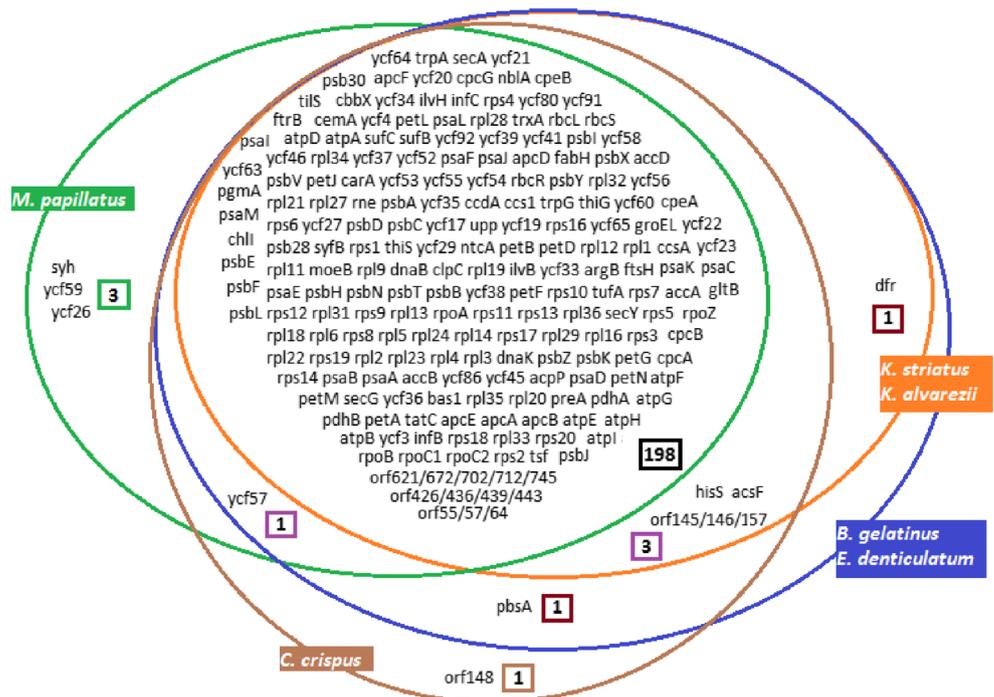
Six reported plastid genomes of the order Gigartinales were compact. Ten pairs of gene overlaps were found in plastid genome of *K. striatus* with overlap length of 1–26 bp (*carA-ycf53*, *ccs1-trpG*, *ycf60-rps6*, *psbD-psbC*, *orf146-groEL*, *trnH-ycf29*, *rpl24-rpl14*, *rpl23-rpl4*, *atpF-atpD*, and *psaL-trnT*), which were the same as those of *K. alvarezii*. These ten overlaps were conserved in the plastid genomes of the four Solieriaceae species. Besides the ten conservative overlaps, plastid genome of *B. gelatinus* had one additional overlap (*pbsA-rpl35*) and *E. denticulatum* with another one (*rpl21-rpl27*). There were seven overlaps in plastid genome of *C. crispus* and eight in *M. papillatus*. Four overlaps of *trnH-ycf29*, *rpl24-rpl14*, *rpl23-rpl4*, and *atpF-atpD* were conserved in all reported plastid genomes of the order Gigartinales. Notably, the overlapping regions of *rpl23-rpl4* were also observed in some plastid genomes of brown algae and diatom indicating the high conservation.

Protein-encoding genes

In terms of predicted protein-encoding gene content, *B. gelatinus*, *E. denticulatum*, and *K. striatus* encoded 202–204 plastid genes which were very similar with the available plastid genomes of the Gigartinales. Several ancient genes such as *glnB*, *chlB*, *chlL*, and *chlN* were absent in plastid genomes of four species from the Solieriaceae. In total, 198 protein-encoding genes were shared by the six plastid genomes of the Gigartinales that represented the majority of the plastid gene content (97.06% of all plastid genes of *B. gelatinus*, *E. denticulatum*, *C. crispus*, and 98.02% of *K. striatus*, *K. alvarezii*, and *M. papillatus*). Venn diagram comparing gene contents revealed the shared and/or unique genes in plastid genomes of the Gigartinales (Fig. 2). The plastid genomes of *B. gelatinus* and *E. denticulatum* contained six additional genes (*orf145/146*, *hisS*, *acsF*, *pbsA*, *ycf57*, and *dfr*), whereas two species of *Kappaphycus* did not have genes *pbsA* and *ycf57*. Gene *orf148* was specific to plastid genome of *C. crispus* which did not have gene *dfr*. The plastid genome of *M. papillatus* did not have genes *orf145/146/157*, *hiss*, *acsF*, *pbsA*, and *dfr* while had three specific genes (*syh*, *ycf59*, and *ycf26*).

Nearly all the protein-encoding genes used ATG as the start codon in plastid genomes of *B. gelatinus* (193, 94.61%), *E. denticulatum* (191, 93.63%), and *K. striatus* (189,

Fig. 2 Venn diagram of protein-encoding gene content of plastid genomes from six reported species of the order Gigartinales



93.56%). In addition to ATG, TTG and GTG were also used as start codons. TTG was used as the start codon for genes *ycf27*, *ycf57*, and *ycf20* in plastid genome of *B. gelatinus*, for genes *ycf27*, *orf712*, and *trxA* in *E. denticulatum* and for genes *ycf20*, *ycf27*, and *trxA* in *K. striatus*. About 8–10 plastid genes

including *atpD*, *chlI*, *infC*, *petN*, *psbC*, *rbcS*, *rpl3*, *rpl24*, and *rps8*, etc., used GTG as the start codon (Table 1).

All three typical stop codons (TAA, TAG, and TGA) were identified with an obvious preference to TAA amounting for 74.51% in plastid genome of *B. gelatinus*, 68.63% in

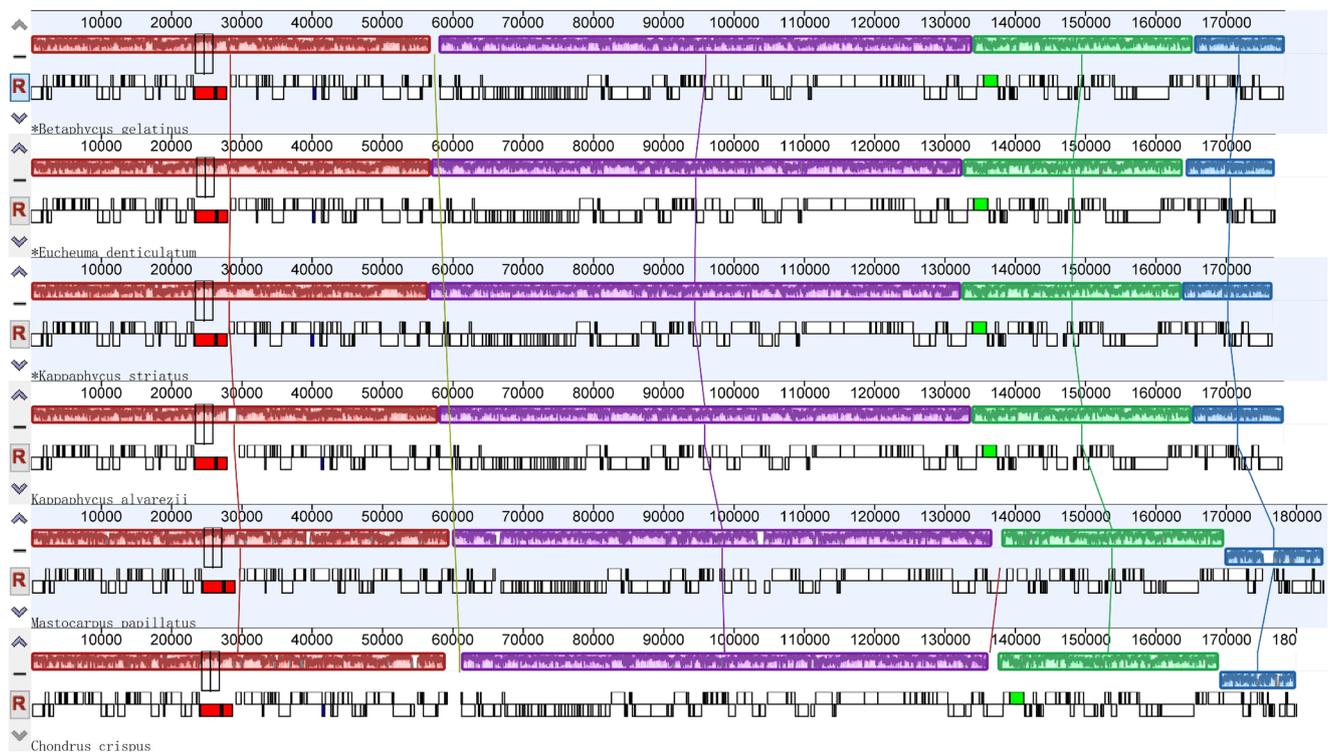


Fig. 3 Whole-genome multiple alignments of six plastid genomes from the order Gigartinales using Mauve software

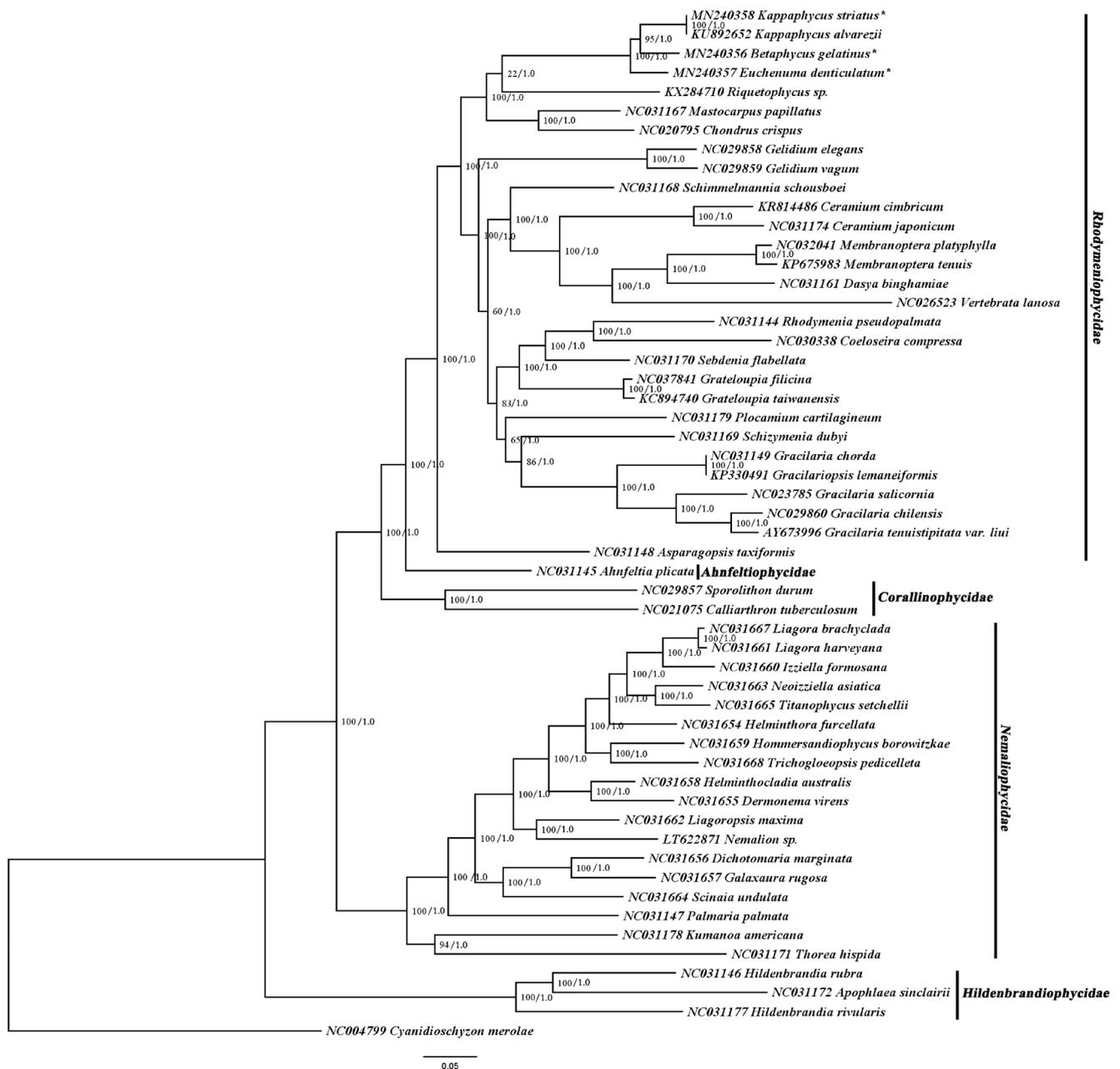


Fig. 4 Phylogenetic tree (ML and BI) of 53 species from Florideophyceae based on 138 shared plastid protein-encoding genes. The support values for each node are shown from maximum-likelihood

bootstrap and Bayesian posterior probabilities. Asterisks indicate newly sequenced *B. gelatinus*, *E. denticulatum*, and *K. striatus* in this work

E. denticulatum and 70.29% in *K. striatus*. They were similar to those of *K. alvarezii*, *C. crispus* and *M. papillatus*.

Comparative analysis

Detailed characteristics of gene order based on co-linear analysis for six complete plastid genomes of the order Gigartinales are shown in Fig. 3. The results showed that six plastid genomes were highly conserved and there was considerable sequence synteny among all the reported species except one

remarkable gene rearrangement. The approximately 12.5-kb gene fragment from gene *psaM* to *ycf21* in plastid genomes of four Solieriaceae species (*B. gelatinus*, *E. denticulatum*, *K. striatus* and *K. alvarezii*) was completely reversed to that of other two Gigartinales species (*M. papillatus* and *C. crispus*). Co-linear results of the plastid genomes among six species from the order Gigartinales and *Riquetophycus sp.* from the order Peyssonneliales showed that gene order from gene *psaM* to *ycf21* of *Riquetophycus sp.* was completely opposite to those of four species of the Solieriaceae, while

consistent with those of *M. papillatus* and *C. crispus*. Additionally, another reversed fragment with the length of about 18.5 kb from gene *ycf46* to *trnN* was identified in plastid genomes of the order Gigartinales compared to *Riquetophycus* sp. (Fig. S2).

Phylogenetic analysis

Phylogenetic analysis was performed based on 138 shared protein-encoding genes from 53 available plastid genomes of Florideophyceae using both ML and BI methods in order to elucidate the global phylogeny. The ML and BI phylogenetic results both showed that all species were clearly divided into five groups corresponding to their subclasses: Rhodymeniophycidae, Ahnfeltiophycidae, Corallinophycidae, Nemaliophycidae and Hildenbrandiophycidae (Fig. 4). Six species of the order Gigartinales were separated into two clades. Four species of the Solieriaceae and *Riquetophycus* sp. from the order Peyssonneliales formed one clade, while *C. crispus* (Gigartinaceae) and *M. papillatus* (Phyllophoraceae) formed another clade. The phylogenetic trees inferred by ML and BI methods were concordant except for the clade about four Solieriaceae species and *Riquetophycus* sp. with low bootstrap support (22) but high posterior probability value (1.0). The phylogenetic relationship of the family Solieriaceae within the order Gigartinales was reconstructed. Four species of the Solieriaceae, *B. gelatinus*, *E. denticulatum*, *K. striatus* and *K. alvarezii*, formed one sub-clade. Within this sub-clade, *K. striatus* and *K. alvarezii* had a closer relationship with *B. gelatinus* than *E. denticulatum*. The overall topologies from these four species of the family Solieriaceae based on seven plastid genes and six mitochondrial genes were consistent with the above results (Fig. S3).

Discussion

In this study, the complete plastid genomes of *B. gelatinus*, *E. denticulatum* and *K. striatus* from the family Solieriaceae were characterized which expanded the available plastid pool of red algae.

Most of the genes found in plastid genomes from the red lineage have been identified in this work. Among the plastid genomes of red algae, species of the Solieriaceae maintained less ancient gene content. Genes *glnB*, *chlB*, *chlL*, and *chlN* were absent in the family Solieriaceae while were present in the family Bangiaceae (Wang et al. 2013a). The absence of these ancient genes revealed that species of the Solieriaceae were more advanced multicellular red algae. Additionally, we found plastids of red algae had the larger coding capacity compared with reported plastid genomes of brown algae (Wang et al. 2013b; Zhang et al. 2015; Liu and Pang 2016). According to a report about plastid genomes of algae, the large

protein-encoding gene content might play a crucial role in stabilizing the photosynthetic mechanism to resist high stress associated with adaption to intertidal environments (Ji et al. 2010). Additionally, plastid genomes from most species of Florideophyceae including *B. gelatinus*, *E. denticulatum*, and *K. striatus* retained a single copy of the rDNA operons. In contrast, most of the early-diverged classes in the red algae contain two copies of operons. Previous reports demonstrated that duplicated rDNA operons or inverted repeats containing rDNAs were widely distributed in primary and secondary plastids as well as in cyanobacterial genomes (Wang et al. 2013b; Lee et al. 2016; Liu and Pang 2016). This suggests there was the loss of one copy in plastid genomes of the most Florideophyceae species in the evolution.

The gene content in plastid genomes among six species of the Gigartinales was similar. The six plastid genomes have a common core set of 198 protein-encoding genes which represents the majority of the plastid gene content. Here, missing or unique genes are mainly related to genes *orf* and *ycf* such as *orf145/146/157*, *orf148*, and *ycf57*, *ycf59*, and *ycf26*. The functions of these genes are to encode the conserved hypothetical proteins; therefore, the effect of their deletion in the plastid genomes was difficult to access. However, we noticed that gene *pbsA* was absent in the plastid genomes of *K. striatus*, *K. alvarezii* and *M. papillatus*, while it was present in most marine red algal species (Cho et al. 2018). Gene *pbsA* was known for encoding heme oxygenase contributing to the iron controlling system. It has two additional copies (HMOX1 and HMOX2) that have been reported in the nuclear genomes of some red algae. It is well known that genes could be transferred to the nucleus through endosymbiotic gene transfer (Graf et al. 2017). Although gene loss is a common event throughout plastid genome evolution, previous research showed that functional requirement as demonstrated in *pbsA* might have prevented gene loss during the long evolutionary history of red algal plastid genomes (Cho et al. 2018). In *Chlamydomonas reinhardtii* from Chlorophyta, two nuclear-encoded heme oxygenases have been named as HMOX1 and HMOX2 and there is no gene *pbsA* in its plastid genome (Duanmu et al. 2013). Therefore, we deduce that red algae lacking gene *pbsA* in the plastid genome might have its homologs in the nucleus. However, we were unable to locate the homology sequences to gene *pbsA* in the nuclear genomes of *K. striatus*, *K. alvarezii* and *M. papillatus* because there are no corresponding nuclear genomes available.

Among all the available plastid genomes from the Gigartinales species, the majority of protein-encoding plastid genes used ATG as the start codon, and a few genes used GTG and TTG as the start codons. In particular, gene *ycf86* used ATA as the start codon in the plastid genome of *M. papillatus*. Apart from the standard start codon ATG, GTG was another commonly used start codon, which was mainly present in bacteria. Analysis of organellar genomes revealed that several

plastid genes and mitochondrial genes used GTG as the start codon in both red algae and brown algae (Corguille et al. 2009; Wang et al. 2013a, 2013b; Zhang et al. 2013; Li et al. 2018; Liu et al. 2019). In four species of the Solieriaceae, the start codon GTG was mainly utilized in the specific plastid genes, showing a conserved evolutionary process of these genes. TTG, which was known as unusual start codon in eubacteria and archaea, was located in frame upstream from the originally proposed ATG start codon (Golderer et al. 1995). Heretofore, TTG as the start codon has been found in some plastid genes in the red lineage, while it has not been identified in brown algal plastid genomes which are descended from the red algae plastid. This provides the evidence that red algae are more primitive than brown algae.

Co-linear analysis indicated six reported plastid genomes from the order Gigartinales were conserved in genome structure and organization, except the specific gene rearrangement of an 12.5-kb gene fragment from *psaM* to *ycf21*. This rearranged fragment was identified for the first time in the plastid genome of *K. alvarezii* from the family Solieriaceae (Liu et al. 2019). The three new plastid genomes of the family Solieriaceae had the same gene inversion compared with other two species of the order Gigartinales, which implied that this phenomenon might be unique to the Solieriaceae. Several inversion events have been recognized as effective phylogenetic markers, such as a 30-kb inversion found in all living vascular plants (Raubeson and Jansen 1992) and two short inversions putatively shared by all ferns (Roper et al. 2007; Gao et al. 2009). We might propose 12.5-kb inversion to be potential phylogenetic markers uniting the Solieriaceae species. In plastid genomes of Florideophyceae, another four gene inversions among three orders (Gracilariales, Gigartinales and Peyssonneliales) have been reported (Liu et al. 2019), indicating the rearrangements could occur at the order level in red algae. Furthermore, our findings showed that gene inversions of plastid genomes could also occur at the family level. In general, we assume gene rearrangements such as reordering of genetic elements mainly were caused by repeated inversion. However, no inverted sequences were detected at the junction area between the rearranged fragments and their contiguous regions and no shared feature was determined. There is an alternative point that illegitimate recombinations are thought to be the major cause of gene rearrangements (Boudreau and Turmel 1995). The mechanism of gene rearrangement deserves to be researched in depth.

The phylogenetic relationship revealed by this study is similar with previous reports (Yang et al. 2015; Lee et al. 2016; Liu et al. 2019). However, the phylogenetic analysis within the family Solieriaceae is inconsistent with previous observation from mitochondrial genome. Within the Solieriaceae clade, *E. denticulatum* was basal relative to other three species in this study, whereas *B. gelatinus* was basal species based on concatenated mitochondrial genes (Li et al. 2018). We reconstructed phylogenetic trees based on both plastid and

mitochondrial genes in order to further verify the phylogenetic relationship of *Betaphycus*, *Eucheuma* and *Kappaphycus*. The overall topologies (ML and BI) were consistent with the trees constructed using the plastid genomes. It was supposed that plastid genomes could provide more evolutionary information compared with mitochondrial genomes (Xu et al. 2018).

Here, one interesting finding of phylogenetic analysis was the relationship between four species of the family Solieriaceae and other two Gigartinales species (*C. crispus* and *M. papillatus*). Based on mitochondrial genomes, *C. crispus* and *M. papillatus* were clustered with four Solieriaceae species firstly in the ML tree (bootstrap support = 48), but they were firstly clustered with *Riquetophycus* sp. in the BI tree (posterior probabilities = 0.57) (Li et al. 2018). In this study, four Solieriaceae species and *Riquetophycus* sp. clustered together firstly with low bootstrap support (22) in the ML tree but with high posterior probability values (1.0) in the BI tree. Therefore, it was difficult to state these phylogenetic relationships precisely due to the low bootstrap and posterior probability values. The uncertain phylogenetic results may be due to limited taxon sampling and insufficient genomic data available. Combined with the results of co-linear analysis, we deduced the family Solieriaceae in the order Gigartinales had relatively significant differentiation. All of them might imply there was a non-monophyletic relationship of the order Gigartinales. In any case, we could not fully resolve the phylogenetic relationships among the family Solieriaceae based on the existing molecular data. Therefore, further investigations, not only on molecular phylogeny of more taxa but also on the cytology, should be conducted to refine the above research.

Funding information This work was supported by the China-ASEAN Maritime Cooperation Fund; Public Science and Technology Research Funds Projects of Ocean (Grant No. 201405020); the Natural Science Foundation of Shandong Province (Grant No. ZR2019MC049).

References

- Adnan H, Porse H (1987) Culture of *Eucheuma cottonii* and *Eucheuma spinosum* in Indonesia. *Hydrobiologia* 151:355–358
- Boudreau E, Turmel M (1995) Gene rearrangements in *Chlamydomonas* chloroplast DNAs are accounted for by inversions and by the expansion/contraction of the inverted repeat. *Plant Mol Biol* 27: 351–364
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552
- Cho CH, Choi JW, Lam DW, Kim KM, Yoon HS, Devarenne TP (2018) Plastid genome analysis of three nemaliophycidae red algal species suggests environmental adaptation for iron limited habitats. *PLoS One* 13:e0196995

- Collén J, Porcel B, Carré W, Ball SG, Chaparro C, Tonon T, Barbeyron T, Michel G, Noel B, Valentin K, Elias M, Artiguenave F, Arun A, Aury JM, Barbosa-Neto JF, Bothwell JH, Bouget FY, Brillet L, Cabello-Hurtado F, Capella-Gutiérrez S, Charrier B, Cladière L, Cock JM, Coelho SM, Colleoni C, Czjzek M, da Silva C, Delage L, Denoed F, Deschamps P, Dittami SM, Gabaldón T, Gachon CM, Groisillier A, Hervé C, Jabbari K, Katinka M, Kloareg B, Kowalczyk N, Labadie K, Leblanc C, Lopez PJ, McLachlan D, Meslet-Cladière L, Moustafa A, Nehr Z, Nyvall Collén P, Panaud O, Partensky F, Poulain J, Rensing SA, Rousvoal S, Samson G, Symeonidi A, Weissenbach J, Zambounis A, Wincker P, Boyen C (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc Natl Acad Sci U S A* 110:5247–5252
- Conklin KY, Kurihara A, Sherwood AR (2009) A molecular method for identification of the morphologically plastic invasive algal genera *Euclidean* and *Kappaphycus* (Rhodophyta, Gigartinales) in Hawaii. *J Appl Phycol* 21:691–699
- Corguille GL, Pearson G, Valente M, Viegas C, Gschloessl B, Corre E (2009) Plastid genomes of two brown algae, *Ectocarpus siliculosus* and *Fucus vesiculosus*: further insights on the evolution of red-algal derived plastids. *BMC Evol Biol* 9:253
- Darriba D, Taboada GL, Doallo R, Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27:1164–1165
- Doty MS, Norris JN (1985) *Euclidean* species (Solieriaceae, Rhodophyta) that are major sources of carrageenan. In: Abbott IA (ed) *Taxonomy of economic seaweeds*. California Sea Grant College Program, La Jolla, pp 47–61
- Duanmu D, Casero D, Dent RM, Gallaher S, Yang W, Rockwell NC, Martin SS, Pellegrini M, Niyogi KK, Merchant SS, Grossman AR, Lagarias JC (2013) Retrograde bilin signaling enables *Chlamydomonas* greening and phototrophic survival. *Proc Natl Acad Sci U S A* 110:3621–3626
- Dumilag RV, Liao LM, Lluisma AO (2014) Phylogeny of *Betaphycus* (Gigartinales, Rhodophyta) as inferred from *COI* sequences and morphological observations on *B. philippinensis*. *J Appl Phycol* 26:587–595
- Fredericq S, Freshwater DW, Hommersand MH (1999) Observations on the phylogenetic systematics and biogeography of the Solieriaceae (Gigartinales, Rhodophyta) inferred from *rbcL* sequences and morphological evidence. *Hydrobiologia* 398:25–38
- Gao L, Yi X, Yang YX, Su YJ, Wang T (2009) Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. *BMC Evol Biol* 9:130
- Gao L, Su YJ, Wang T (2010) Plastid genome sequencing, comparative genomics, and phylogenomics: current status and prospects. *J Syst Evol* 48:77–93
- de Góes HG, Reis RP (2011) An initial comparison of tubular netting versus tie-tie methods of cultivation for *Kappaphycus alvarezii* (Rhodophyta, Solieriaceae) on the south coast of Rio de Janeiro state, Brazil. *J Appl Phycol* 23:607–613
- Golderer G, Dlaska M, Grobner P, Piendl W (1995) TTG serves as an initiation codon for the ribosomal protein MvaS7 from the Archaeon *Methanococcus vannielii*. *J Bacteriol* 177:5994–5996
- Graf L, Kim YJ, Cho GY, Miller KA, Yoon HS (2017) Plastid and mitochondrial genomes of *Coccoloba langsdorffii* (Fucales, Phaeophyceae) and the utility of molecular markers. *PLoS One* 12:e0187104
- Greiner S, Lehwark P, Bock R (2019) Organellar genome DRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res* 47:59–64
- Guiry MD, Guiry GM (2020) AlgaeBase. World-wide electronic publication, National University of Ireland, Galway <https://www.algaebase.org>; searched on 06 March 2020
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Heather JM, Chain B (2016) The sequence of sequencers: the history of sequencing DNA. *Genomics* 107:1–8
- Huelsenbeck JP, Ronquist F (2001) MrBayes: bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755
- Janouškovec J, Liu SL, Martone PT, Carré W, Leblanc C, Keeling PJ (2013) Evolution of red algal plastid genomes: ancient architectures, introns, horizontal gene transfer, and taxonomic utility of plastid markers. *PLoS One* 8:e59001
- Ji L, Xie SL, Feng J (2010) Progress in chloroplast genome of algae. *Acta Botan Boreali-Occiden Sin* 1:214–220 (in Chinese with English abstract)
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649
- Lee JM, Cho CH, Park SI, Ji WC, Song HS, West JA, Bhattacharya D, Yoon HS (2016) Parallel evolution of highly conserved plastid genome architecture in red seaweeds and seed plants. *BMC Biol* 14:75
- Li Y, Liu N, Wang X, Tang X, Zhang L, Meinita MDN, Wang G, Yin H, Jin Y, Wang H, Liu C, Chi S, Liu T, Zhang J (2018) Comparative genomics and systematics of *Betaphycus*, *Euclidean*, and *Kappaphycus* (Solieriaceae: Rhodophyta) based on mitochondrial genome. *J Appl Phycol* 30:3435–3443
- Lim PE, Tan J, Phang SM, Nikmatullah A, Hong DD, Sunarpi H, Hurtado AQ (2014) Genetic diversity of *Kappaphycus* Doty and *Euclidean* J. Agardh (Solieriaceae, Rhodophyta) in Southeast Asia. *J Appl Phycol* 26:1253–1272
- Lim PE, Yang LE, Tan J, Meggs CA, Brodie J (2017) Advancing the taxonomy of economically important red seaweeds (Rhodophyta). *Eur J Phycol* 52:438–451
- Liu F, Pang SJ (2016) Chloroplast genome of *Sargassum horneri* (Sargassaceae, Phaeophyceae): comparative chloroplast genomics of brown algae. *J Appl Phycol* 28:1419–1426
- Liu CL, Huang XH, Liu JG (2009) The primary application of ISSR analysis on strains identification of *Kappaphycus alvarezii*. *Mar Sci* 33:50–53
- Liu T, Sun Y, Jin YM, Liu C, Zhang L, Zhang S (2012) Study on molecular systematics of four seaweed species in family Solieriaceae (Rhodophyta). *Periodical Ocean Univ China* 42:39–48 (in Chinese with English abstract)
- Liu T, Yang SG, Wu XY, Chen WZ, Zhang J, Tang XM (2018) Complete plastid genome of *Gracilaria bailiniae* (Rhodophyta) and phylogenetic analysis. *Mitochondrial DNA B* 4:10–11
- Liu N, Zhang L, Tang XM, Wang XM, Meinita NMD, Wang GL, Chen W, Liu T (2019) Complete plastid genome of *Kappaphycus alvarezii*: insights of large-scale rearrangements among Florideophyceae plastid genomes. *J Appl Phycol* 31:3997–4005
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1:18
- Ma PF, Zhang YX, Zeng CX, Guo ZH, Li DZ (2014) Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (Poaceae). *Syst Biol* 63:933–950
- Pereira L, Meireles F, Abreu HT, Ribeiro-Claro PJA (2015) A comparative analysis of carrageenans produced by underutilized versus industrially utilized macroalgae (Gigartinales, Rhodophyta). In: Kim S-K, Chojnacka K (eds) *Marine Algae Extracts*. Weinheim, Wiley-VCH, pp 277–294
- Rambaut A (2009) FigTree v1.3.1. 2006–2009. Accessed on November 29, 2012, Program package available at <http://tree.bio.ed.ac>

- Raubeson LA, Jansen RK (1992) Chloroplast DNA evidence on the ancient evolutionary split in vascular land plants. *Science* 255:1697–1699
- Reyes-Prieto A, Weber AP, Bhattacharya D (2007) The origin and establishment of the plastid in algae and plants. *Annu Rev Genet* 41:147–168
- Roper JM, Kellon Hansen S, Wolf PG, Karol KG, Mandoli DF, Everett KDE, Kuehl J, Boore JL (2007) The complete plastid genome sequence of *Angiopteris evecta* (G. Forst.) Hoffm. (Marattiaceae). *Am Fern J* 97:95–106
- Schattner P, Brooks AN, Lowe TM (2005) The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 33:686–689
- Sissinia MN, Navarrete-Fernandez TM, EMC M, Freese JM, Gentilhomme AS, Huber SR, Mumford TF, Hughey JR (2016) Mitochondrial and plastid genome analysis of the heteromorphic red alga *Mastocarpus papillatus* (C. Agardh) Kützing (Phyllophoraceae, Rhodophyta) reveals two characteristic florideophyte organellar genomes. *Mitochondrial DNA B* 1:676–677
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690
- Sun XY, Luo D, Zhao C, Li W, Liu T (2011) DNA extraction and PCR analysis of five kinds of large seaweed under different preservation conditions. *Mol Plant Breed* 9:1680–1691 (in Chinese with English abstract)
- Sun Y, Liu C, Chi S, Tang XM, Chen FX, Liu T (2014) Molecular systematic of four Solieriaceae algae based on ITS and *rbcL* sequence. *Preiodical Ocean Univ China* 44:54–60 (in Chinese with English abstract)
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Tan J, Lim PE, Phang SM (2013) Phylogenetic relationship of *Kappaphycus* Doty and *Euclidean* J. Agardh (Solieriaceae, Rhodophyta) in Malaysia. *J Appl Phycol* 25:13–29
- Wang L, Mao YX, Kong FN, Li GY, Ma F, Zhang BL, Sun P, Bi G, Zhang F, Xue H, Cao M (2013a) Complete sequence and analysis of plastid genomes of two economically important red algae: *Pyropia haitanensis* and *Pyropia yezoensis*. *PLoS One* 8:e65902
- Wang XL, Shao ZR, Fu WD, Yao JT, Hu QP, Duan DL (2013b) Chloroplast genome of one brown seaweed, *Saccharina japonica* (Laminariales, Phaeophyta): its structural features and phylogenetic analyses with other photosynthetic plastids. *Mar Genomics* 10:1–9
- Xu KP, Tang XH, Bi GQ, Cao M, Wang L, Mao XY (2018) The first complete organellar genomes of an Antarctic red alga, *Pyropia endiviifolia*: insights into its genome architecture and phylogenetic position within genus *Pyropia* (Bangiales, Rhodophyta). *J Oceanol Limnol* 36:1315–1328
- Yang EC, Kim KM, Kim SY, Lee J, Boo GH, Lee JH, Nelson WA, Yi G, Schmidt WE, Fredericq S, Boo SM, Bhattacharya D, Yoon HS et al (2015) Highly conserved mitochondrial genomes among multicellular red algae of the Florideophyceae. *Genome Biol Evol* 7:2394–2406
- Zhang T, Shi X, He JH, Chen QF, Feng ZH, He PM (2011) Preliminary analysis on identification *Euclidean* and *Kappaphycus* (Rhodophyta) by ISSR. *Acta Oceanol Sin* 33:173–178 (in Chinese with English abstract)
- Zhang J, Wang XM, Liu C, Jin YM, Liu T (2013) The complete mitochondrial genomes of two brown algae (Laminariales, Phaeophyceae) and phylogenetic analysis within *Laminaria*. *J Appl Phycol* 25:1247–1253
- Zhang L, Wang XM, Liu T, Wang GL, Chi S, Liu C, Wang H (2015) Complete plastid genome sequence of the brown alga *Undaria pinnatifida*. *PLoS One* 10:e0139366
- Zhang J, Tang XM, Chen WZ, Liu T, Li Y (2018) The complete plastid genome of *Grateloupia filicina* (Rhodophyta) and phylogenetic analysis. *Mitochondrial DNA B* 3:1172–1173
- Zhou X, Xu S, Xu J, Chen B, Zhou K, Yang G (2012) Phylogenomic analysis resolves the interordinal relationships and rapid diversification of the laurasiatherian mammals. *Syst Biol* 61:150–164

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