

Research Article

Comparative Analysis of MicroRNAs between Sporophyte and Gametophyte of *Porphyra yezoensis*

Linwen He,^{1,2} Aiyu Huang,^{1,2} Songdong Shen,³ Jianfeng Niu,¹ and Guangce Wang¹

¹Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences (IOCAS), Nanhai Road 7, Qingdao 266071, China

²School of Earth Science, Graduate University of Chinese Academy of Sciences, Yuquan Road 19, Beijing 100049, China

³College of Life Sciences, Soochow University, Renai Road 199, Suzhou 215123, China

Correspondence should be addressed to Guangce Wang, gcwang@qdio.ac.cn

Received 29 March 2012; Revised 6 August 2012; Accepted 22 August 2012

Academic Editor: Paul Denny

Copyright © 2012 Linwen He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Porphyra yezoensis Ueda is an intertidal marine red algae that has received increasing attention as a model organism owing to its important role in biological research and the agronomic industry. The two generations of *Porphyra yezoensis*, the sporophyte and the gametophyte, have the same genome but show great differences in many aspects, including structural features, habitat, and gene expression. To identify miRNAs and their probable roles in *P. yezoensis* development, we constructed and sequenced libraries of small RNA from *P. yezoensis* sporophytes and gametophytes. The sequencing data were analyzed, and 14 miRNAs were identified, with only one common to these two samples. Our results show that *P. yezoensis* has a complex small RNA processing system containing novel miRNAs that have no identifiable homolog in other organisms. These miRNAs might have important regulatory roles in development of the different generations of *P. yezoensis*.

1. Introduction

Porphyra yezoensis Ueda is an intertidal marine red algae that has received increasing attention as a model organism owing to its important role in biological research and the agronomic industry [1]. *P. yezoensis* is one of the most valuable marine crops in the world and is cultivated widely in Asia, especially in Japan, China, and Republic of Korea [2]. As a red algae, whose relationship to other groups is by no means certain, *P. yezoensis* has many characteristics of lower eukaryotes, including the location of the small subunit of Rubisco in the chloroplast genome [3], the same accessory pigments as cyanobacteria [4], and the high degree of similarity between many of its genes and their homologs in bacteria. Owing to these features, red algae are believed to be original or degraded eukaryotic organisms [5].

P. yezoensis has a unique dimorphic life-cycle consisting of two generations, a microscopic diploid filamentous sporophyte and a macroscopic haploid foliate gametophyte, with completely different morphology [6]. The sporophyte and the gametophyte have the same genome but show great

differences in many aspects, including structural features, habitat, and gene expression. The sporophyte is densely tufted with uniseriate filaments, whereas the gametophyte is monolayered. The sporophyte enters and germinates in shells, whereas the gametophyte lives on static substrates of the intertidal zone, experiencing stress caused by strong light, high temperature, and desiccation during low tide [7–9]. Analysis of expressed sequence tag (EST) groups generated from sporophytes and gametophytes of *P. yezoensis* found that only 22.5% of groups commonly occurred in both generations, indicating great differences in gene expression resulting in morphological differences between the two generations [10]. These characteristics prompted us to hypothesize that the gene expression regulators (e.g., microRNAs (miRNAs)) of the two generations might show different specificity.

miRNAs are important noncoding small RNAs (sRNAs) that can influence the output of quantity genes in eukaryotes by targeting mRNAs for translational repression or cleavage [11–13]. miRNAs have been widely studied only recently, but they are attracting a great deal of attention and are being

studied in many organisms. miRNAs are believed to exist in animals, plants, and viruses with a high degree of conservation in each kingdom [14, 15]. The expression of miRNAs has a spatiotemporal pattern [11, 12, 16–18], and each of them influences the transcription and translation of specific genes [15]. miRNAs are of simple structure and have important roles in gene regulation in various processes [15], including developmental patterning, cell proliferation, tumor generation [19], stress resistance [19], auxin response [20, 21], fat metabolism, and miRNA biogenesis [22, 23]. miRNAs have been studied extensively in higher plants and in animals, and Liang et al. [24] identified miRNAs from the *P. yezoensis* sporophyte but the likely roles of miRNA in the development of different generations of *P. yezoensis* remain unknown.

In the present study, we constructed sRNA libraries from the sporophyte and the gametophyte of *P. yezoensis* then used high-throughput Solexa technology to deeply sequence the sRNAs. The sequencing data were analyzed, and miRNAs were identified from both samples studied. This study has provided insights into the expression and function of small silencing RNAs in *P. yezoensis*.

2. Materials and Methods

2.1. Culture of Sporophytes and Gametophytes of *P. yezoensis*. The *P. yezoensis* sporophytes and gametophytes of *P. yezoensis* were available in our laboratory. Both were cultured by constant aeration in Provasoli enriched seawater (PES) culture medium made with steam-sterilized local seawater supplemented with inorganic nutrients and vitamins (filter sterilized), with renewal of the culture medium every week. As the sporophytes and gametophytes grow well under different conditions, they were cultured at different temperatures and under different light intensities. The sporophytes were grown at 15°C under an illumination intensity of 50 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ with a 12 h dark/12 h light photoperiod. The gametophytes were induced from conchocelis; basically, a shell with *P. yezoensis* conchocelis was cultured at 26°C with PES medium in a 500 mL glass beaker under an illumination intensity of 20 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ with a 12 h dark/12 h light photoperiod. After several weeks, the shell was placed at 22°C and some fibers were placed on it for attachment of the conchospores. The fibers with conchospores were cultured at 10°C under an illumination intensity of 50 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ with a 12 h dark/12 h light photoperiod. Blades were selected randomly for further culture in 1-liter glass beakers.

2.2. RNA Extraction, Library Construction, and Sequencing. Total RNA was extracted with Trizol reagent according to the manufacturer's protocol (Invitrogen, USA). sRNAs with a size range of 18–28 nt were collected and sequenced. Basically, sRNAs were separated by electrophoresis in denaturing polyacrylamide gels, and fragments of 18–28 nt were recovered by gel purification. A pair of adaptors was ligated sequentially to their 5'- and 3'-ends, and the ligated sRNAs were used as template for cDNA synthesis, using the SuperScript II Reverse Transcription Kit (Invitrogen, USA).

The cDNAs were then amplified by PCR, and the products were sequenced directly with a Solexa 1G Genome Analyzer (Figure 1).

2.3. Initial Processing of Reads. A perl script was used to complete the initial processing of reads. Low-quality tags were removed, 3' adaptor sequences were trimmed, adaptor contamination tags were removed, and tags smaller than 18 nt were filtered out. The Short Oligonucleotide Analysis Package (SOAP) [25] was used to map the remaining sRNA sequences (clean reads) to the *P. yezoensis* EST sequences; all hits were reported and mismatch was not allowed. All the clean reads were aligned against noncoding RNA from Rfam (www.sanger.ac.uk/resources/databases/rfam.html) and GenBank (<http://www.ncbi.nlm.nih.gov/>) to identify noncoding RNAs (rRNA, tRNA, snRNA, and snoRNA) fragments, using blastn [26] with an *e*-value of 0.01 as cutoff. All the clean reads were compared to all known plant miRNAs available from miRBase (miRBase Sequence Database version 15; <http://www.mirbase.org/>) to identify homologs of known miRNAs (≤ 2 mismatches with other known miRNAs). Tags from clean reads were aligned with each other to identify potential small interfering (siRNA) candidates; the two perfectly complementary sRNAs with 2 nt overhangs at the 3'-end were considered to be siRNA. All clean reads were classified according to their identity with the sRNA categories mentioned above. In the case that an sRNA was mapped to more than one category, the following priority rule was adopted: rRNA and so forth (in which GenBank > Rfam) > known miRNA > siRNA [27].

2.4. miRNA Identification. miRNAs were identified by structure filtering. As the genome of *P. yezoensis* was not sequenced, EST sequences were scanned to identify potential hairpin regions for homologs of known miRNAs and the remaining nonannotated sRNAs (Figure 1). sRNAs with more than one read and ≤ 20 perfect matches to EST sequences were folded with 300 nt of upstream and downstream flanking sequences and examined for secondary structures to identify potential miRNAs on the basis of the criteria described in the following. Precursors with minimum free energy (MFE) $\leq -18 \text{ kcal mol}^{-1}$ checking by Mfold [28, 29], $\geq 16 \text{ bp}$ and ≤ 4 bulges or asymmetries between miRNA and miRNA*, with space between miRNA and miRNA* $\leq 300 \text{ nt}$, mature sequence length 18–25 nt, and a flank sequence length of 20 nt, were considered to be potential *P. yezoensis* pre-miRNAs.

2.5. miRNA Target Prediction. As the genome of *P. yezoensis* has not been sequenced, EST sequences were used to detect potential miRNA targets with parameters and conditions according to the criteria suggested earlier [30, 31]. Basically, ≤ 4 mismatches between the sRNA and the target; ≤ 2.5 mismatches in positions 1–12; no mismatch in positions 10 or 11; no adjacent mismatch in positions 2–12 (counting from the 5'-end of the miRNAs and G-U bases as 0.5 mismatch). Additionally, an MFE of $\geq 74\%$ of the MFE of

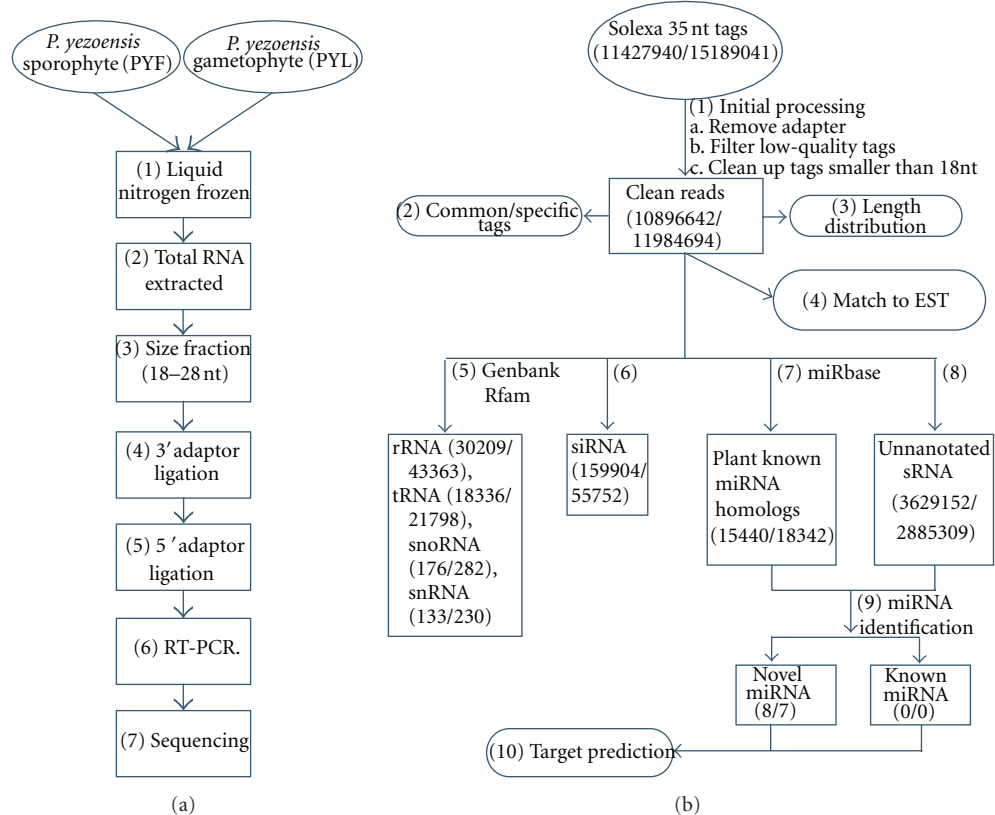


FIGURE 1: A flow chart of the procedure for sample preparation and sequencing and for the processing of reads. (a) A flow chart of the procedure for sample preparation and sequencing. (1) Sporophytes and gametophytes of *P. yezoensis* were frozen rapidly in liquid nitrogen and stored at -80°C before RNA extraction. (2) Total RNA was extracted using the Trizol reagent method. (3) RNA 18–28 nt fragments were gel-purified. (4) A 3' adaptor was ligated to the 3' end of sRNAs. (5) A 5' adaptor was ligated to the 5' end of sRNAs. (6) sRNAs were amplified by RT PCR. (7) Sequencing. (b) A flow chart of the procedure for processing reads; the numbers in parentheses represent the total reads from PYF and PYL, respectively. (1) Initial processing: remove adapter, filter out low-quality tags, and clean up tags smaller than 18nt. (2) Common/specific tags identified between samples. (3) Length distribution analysis of clean reads. (4) Clean reads matched to *P. yezoensis* EST sequences using SOAP [25]. (5) Clean reads compared to noncoding RNAs from GenBank and Rfam. (6) siRNA identified. (7) Plant miRNA homologs identified. (8) Annotated sRNAs. (9) miRNA identified by hairpin structure filtering. (10) Target prediction.

the miRNA bound to its perfect complement was required for the miRNA/target duplex.

2.6. Experimental Verification of the Expression of *P. yezoensis* miRNAs. RT PCR was used to detect the expression of *P. yezoensis* miRNAs. cDNAs were synthesized using the NCode VILO miRNA cDNA Synthesis Kit (Invitrogen) according to the manufacturer's protocol. The cDNA was used for PCR amplification of *P. yezoensis* miRNAs. The sense primers were designed according to each miRNA and the antisense primer was the universal primer supplied in the cDNA synthesis kit.

3. Results

3.1. Library Construction, Sequencing, and Initial Processing of Reads. In order to identify miRNAs and their likely roles in *P. yezoensis* development, we constructed and sequenced sRNA libraries from the *P. yezoensis* sporophyte (PYF) and gametophyte (PYL). The two runs yielded a set of

26,616,981 total signatures. After filtering out low-quality data and elimination of adaptor contamination (Figure 1), we obtained sRNAs with a distribution of lengths ranging from 10 to 43 nt, with sequences of 21–22 nt as the major component (Figure 2). miRNAs were commonly larger than 17 nt, so we removed sequences smaller than 18 nt and obtained 10,896,642 and 11,984,694 total sequences, representing 3,853,350 and 3,025,076 unique, although sometimes partially overlapping, clean reads from PYF and PYL, respectively (Table 1). Of these unique sequences, ~75% (2,879,648) and ~74% (2,235,970) were sequenced only once, indicating a diverse sRNA set in *P. yezoensis*.

Although many sRNAs expressed in *P. yezoensis* remained unidentified, they were annotated and classified into non-coding RNAs (Rfam, GenBank), homologs of plant miRNAs (miRBase), or siRNAs according to their identity with sequences from these databases or their characteristics of being siRNA. In the case that some sRNAs were mapped to more than one category, the following priority rule was adopted: rRNA and so forth (in which GenBank > Rfam) > known miRNA > siRNA [27]. Only a few of the sRNAs

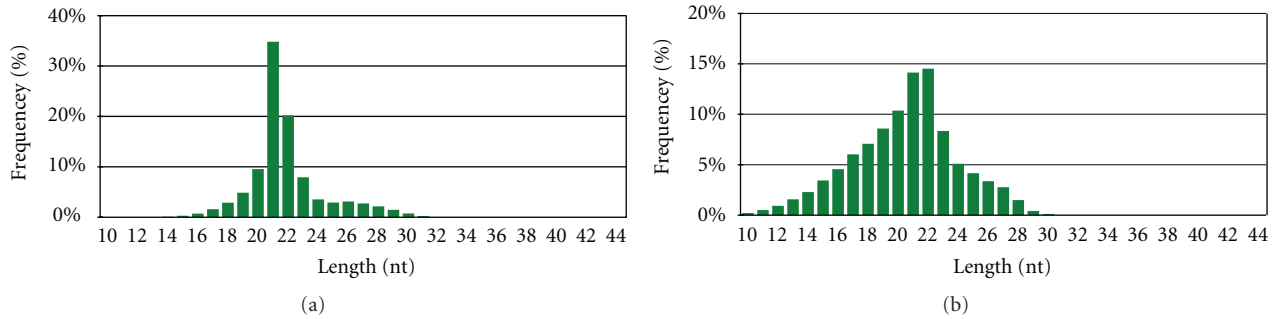


FIGURE 2: Length distributions of unique sRNA sequences in *P. yezoensis*. The length occurrence of each unique sequence read was counted: (a) PYF; (b) PYL.

TABLE 1: Categorization of *P. yezoensis* small RNAs.

Category	PYF				PYL			
	Unique sRNA	Percent (%)	Total sRNA	Percent (%)	Unique sRNA	Percent (%)	Total sRNA	Percent (%)
Total	3853350	100%	10896642	100%	3025076	100%	1.2E + 07	100%
miRNA	15440	0.40%	104081	0.96%	18342	0.61%	322689	2.69%
rRNA	30209	0.78%	197902	1.82%	43363	1.43%	705106	5.88%
siRNA	159904	4.15%	903198	8.29%	55752	1.84%	259369	2.16%
snRNA	133	0.00%	663	0.01%	230	0.01%	581	0.00%
snoRNA	176	0.00%	345	0.00%	282	0.01%	1797	0.01%
tRNA	18336	0.48%	215562	1.98%	21798	0.72%	669520	5.59%
Unannotated	3629152	94.18%	9474891	86.95%	2885309	95.38%	1E + 07	83.65%

were annotated; $\sim 85\%$ total reads and $\sim 95\%$ unique reads remained nonannotated. Of all the annotated sRNAs, siRNAs were the most abundant sequences retrieved from the *P. yezoensis* unique sRNA pools, with the highest read frequency of all sRNA categories in both samples: 4.15% for PYF and 1.84% for PYL (Table 1). Yet, in the total sRNA pools, siRNA represented a significant part (8.29%) in PYF, while rRNA (5.88%) and tRNA (5.59%) were the significant components in PYL. Homologs of known plant miRNAs accounted for $\sim 0.4\%$ and $\sim 0.61\%$ of the unique sequences in PYF and PYL, respectively; whereas in total sequence pools, the numbers were $\sim 0.96\%$ and $\sim 2.69\%$ in PYF and PYL, respectively, indicating that homologs of miRNA might be expressed differently in PYF and PYL. sRNAs mapped to snRNA and snoRNA were rare, and the remaining sRNAs were not annotated. Analysis of common and specific sequences showed that only $\sim 6\%$ of the unique sequences were shared by the two samples (Table 2), suggesting the presence of a diverse set of endogenous sRNAs in *P. yezoensis*.

3.2. miRNAs in *P. yezoensis*. The identification of a diverse set of sRNAs in *P. yezoensis* prompted us to examine whether they were indeed functional. We used homologs of known plant miRNAs and the remaining nonannotated sRNAs to identify candidates of known and novel miRNA families in *P. yezoensis*, respectively (Figure 1). Finally, we identified 14 candidate miRNAs in *P. yezoensis*, including two mature miRNAs identified earlier [24].

Each miRNA had a single precursor. The length of pre-miRNA ranged from 47 to 246 nt, with a mean of 135 nt

TABLE 2: Common and specific small RNAs between PYF and PYL.

Class	Unique sRNA	Percent (%)	Total sRNA	Percent (%)
Total sRNAs	6501964	100.00%	22881336	100.00%
PYF and PYL	376462	5.79%	11112797	48.57%
PYF specific	3476888	53.47%	6738816	29.45%
PYL specific	2648614	40.74%	5029723	21.98%

(Table 3). The MFE range was from -135 to -20 kcal mol $^{-1}$, with a mean of -65 kcal mol $^{-1}$, similar to the computational prediction values of Arabidopsis miRNA precursors (-57 kcal mol $^{-1}$) and much lower than that of tRNA (-27.3 kcal mol $^{-1}$) and rRNA (-33 kcal mol $^{-1}$) [32]. The 14 miRNAs were designated Pye-miR 1–14. Sequence terminal variation analysis showed that the percentage of length heterogeneity was $\sim 14\%$ for the 5' end of *P. yezoensis* miRNAs and 32% for the 3' end with either a 3 nt deletion or a 3 nt extension (Figure 3).

3.3. Expression Patterns of *P. yezoensis* miRNA Candidates during Different Generations. To investigate the likely roles of miRNAs in *P. yezoensis* development, we sequenced sRNAs from *P. yezoensis* sporophytes and gametophytes. A total of 14 miRNAs were identified, of which only 1 was sequenced in both samples, 7 were sequenced exclusively from PYF, and 6 were sequenced exclusively from PYL. This indicated that they might have an important role in *P. yezoensis* development. To determine the likely regulated genes, we predicted targets for these miRNAs on the basis of the rules

TABLE 3: Characteristics of *P. yezoensis* pre-miRNA sequences.

Id	Mfe ^a	Length ^b	seq ^c
Pye-miR1	-104	227	ACGCGCACAGCCTCGCCAACCGCGACGTCCGCCATCGCCA CGACCGCCCCCGTCGACAGCTCAACGGTGGCGGCGGCGGG GAAGCACGCCGGTGTGTGTCCGCGCCGTGCCGTTGGCC CCCTCCGCGACATCGCCGACTTGATCGCCGCGCCCGTGT ACACGCACAGCGGGTGTCCACCTGTGAGATGGGGT <u>TGGA</u> <u>CTCGATGGTGTAGGTGAGAAGGTGGCG</u>
Pye-miR2	-34.9	76	GAAACTCAGGACAGCGACGACGACGACCGCAAGCCG CGCCCGTCGCCGTCGTACTAGCTGGGCTGATGACAA CGTGTGCCGCCGTACATCCGAAGTCATCGAGCTGCAGTTC AGGACATCTTCGTGCTGGCGGACGGCGCCTTCTCCATCCAC GTCAACCGTTACAAGAACACGGAGGGACGTGACGACCCTC GCCGCTGGTGTACACCATTCCACGTGACCCCATGCTGGT <u>G</u> <u>CACGACCCGGTCTGGCATTGCTGCGCACC</u>
Pye-miR3	-72	193	GCGCAGCGCGCCGAGGGCGATCGCCGCGCGGCCACCGC CGACCCCCCAGCTCCCCCTCCGCCCCGGCAGCATGGGC GTGGGCGTCAGCGCGGCCGGTGGAGGCTGGAGCGGTCAAC GCCCGGGCACGCC GCCACCGCCG <u>TGGTGTAGGTGCTGGCACGGACGCCGCAAC</u> GACGATATCCCCCTGCAGCGTCTGGGGTGTACAACCCCG CCGAGGTGCCCGTGGAGACGGCGGTACCACCGACGACGC GGCGTCGGCGGCCACGTCCAGCGCGTCCGCCCCCGCACC GACCGCAGCGCGCCGCGCGGCCATTGATGTACACCAGGT GCCCCGGCGTGGCGGTGAGCGCGTGGCCGCTGCGGGTGGT GGCG
Pye-miR4	-28.8	47	GCGCAGCGCGCCGAGGGCGATCGCCGCGCGGCCACCGC CGACCCCCCAGCTCCCCCTCCGCCCCGGCAGCATGGGC GTGGGCGTCAGCGCGGCCGGTGGAGGCTGGAGCGGTCAAC GCCCGGGCACGCC
Pye-miR5	-74.4	138	GCGCAGCGCGCCGAGGGCGATCGCCGCGCGGCCACCGC CGACCCCCCAGCTCCCCCTCCGCCCCGGCAGCATGGGC GTGGGCGTCAGCGCGGCCGGTGGAGGCTGGAGCGGTCAAC GCCCGGGCACGCC
Pye-miR6	-71.2	134	GCGCAGCGCGCCGAGGGCGATCGCCGCGCGGCCACCGC CGACCCCCCAGCTCCCCCTCCGCCCCGGCAGCATGGGC GTGGGCGTCAGCGCGGCCGGTGGAGGCTGGAGCGGTCAAC GCCCGGGCACGCC
Pye-miR7	-135	246	GCCACCGCCG <u>TGGTGTAGGTGCTGGCACGGACGCCGCAAC</u> GACGATATCCCCCTGCAGCGTCTGGGGTGTACAACCCCG CCGAGGTGCCCGTGGAGACGGCGGTACCACCGACGACGC GGCGTCGGCGGCCACGTCCAGCGCGTCCGCCCCCGCACC GACCGCAGCGCGCCGCGCGGCCATTGATGTACACCAGGT GCCCCGGCGTGGCGGTGAGCGCGTGGCCGCTGCGGGTGGT GGCG
Pye-miR8	-117	226	GCGCGCCCCG <u>TCGGACGGGACGAGGGCAGCAAGCCGGCGC</u> TTATGGCCGCGCGCGCCTCCTGCGTGTGCTGCGTCCGGG AGAGCGGCTCGAGGTCGCCGGCGAGCCTCGTCTGCATGGC CGCCAGCTCGTGCTTTGGGCATGCCACTTGGGCACGATG CCGGAGCGCACCTCACCTTCAGCGTGTGAGCCGCGTGCC CTTTCTCGTTCGAGGCGCCGACGCT
Pye-miR9	-20.3	71	TGGTGGGTGTGTCTCTGTGTGTTCTGGGTGCTACGCGCC <u>TAAGGTACGTAATAAACCCTACACCCCTTCCT</u>
Pye-miR10	-62.2	153	GTGGCTGGTACACAACAAGTACACGCGCTCTGAGATGGG CCGGAAGGCAGTCCGTGCCGGCGTCAAGGCCATGTACGC GTACCTCGGCGTCACTGACCGCGAGCGCGATGACGACGT CGGCAGTGGGACTGTCTTGGCATCTCTATTGGCCA ATCCTTGGCCGCTCGGTGAGGGCTCGGACCTGGTCCAAG GCCTTAGCGAGCTTACTCTGGAGGTGAGAGATGACCTCCG CCTGGGTGGGGTCTGCCCGCCGAGAGTCTGAGCCCAT CCCTTGCCTGGACGCAAGGAA
Pye-miR11	-79.4	141	ATCCTTGGCCGCTCGGTGAGGGCTCGGACCTGGTCCAAG GCCTTAGCGAGCTTACTCTGGAGGTGAGAGATGACCTCCG CCTGGGTGGGGTCTGCCCGCCGAGAGTCTGAGCCCAT CCCTTGCCTGGACGCAAGGAA

TABLE 3: Continued.

Id	Mfe ^a	Length ^b	seq ^c
Pye-miR12	-38.8	79	AAAAGCCGCC TGGGGATGGGCATTGAAGGCGTCCGCGGCT TCCTGAAAGCGCTTTGTGCGCGTACCCCGCGCGCCACA GCTGTCGCGTCACAGCTCCAGCGCGCGTGC GCGGGCCGCG
Pye-miR13	-51.2	83	ACGGCCGCGGCC TGTCATGGTCGCTTGGGCACTGATGCGG TGG
Pye-miR14	-24.2	72	CATTGCCAACCGCT TTCGTGACTCTCGGCATGGGAGAGAGCC GCGATTATGGATTGCGGAGAGAGGACAGTGG

^a Minimum free energy (cal/mol) of pre-miRNAs, predicted by mfold.

^b Length of pre-miRNAs.

^c Sequence of pre-miRNAs, the mature miRNA was indicated in italics.

for plant miRNA target prediction suggested by Allen et al. [30]. Most miRNAs found none or < 5 targets, except that 36 EST contigs were suggested as targets for Pye-miR4. However, the *P. yezoensis* genome is not sequenced and most ESTs are not annotated, so it is difficult to determine whether these miRNA targets have any functional bias.

3.4. siRNA in *P. yezoensis*. To identify siRNA in *P. yezoensis*, we compared tags from clean reads against each other. A pair of perfectly complementary sRNAs with a 2 nt overhang at the 3'-end was considered to be siRNA. Potential siRNAs were found to be expressed in *P. yezoensis* with 159,904 (4.15%) and 55,752 (1.84%) unique sequences; and 903,198 (8.29%) and 259,369 (2.16%) total sequences in PYF and PYL, respectively. Yet, owing to the lack of *P. yezoensis* genome information, we cannot determine their location. Thus, we cannot determine if they are phased relative to each other or if they have a role in silencing repetitive sequences in *P. yezoensis*, as for other organisms.

3.5. Homologs of Known Plant miRNAs in *P. yezoensis*. We compared all the clean reads with all known plant miRNAs available from miRBase (miRBase Sequence Database version 15; <http://www.mirbase.org/>) to identify homologs of known miRNAs. If a *P. yezoensis* sRNA exhibited homology with ≤ 2 mismatches (or 90% identity) with other known miRNAs, it was considered as a homolog of known miRNAs and these can be classified into 449 known miRNA families, allowing one or two mismatches between sequences (additional data file 1, see Supplementary Material available online at doi:10.1155/2012/912843). We compared them to miRNAs from 34 other species, including *Arabidopsis thaliana*, *Oryza sativa*, *Sorghum bicolor*, *Pinus taeda*, *Physcomitrella patens*, and *Chlamydomonas reinhardtii* (additional data file 1). Among these homologs, 56 were expressed in *A. thaliana*, 37 in *O. sativa*, 23 in *P. patens*, and only 4 in *C. reinhardtii*.

3.6. Experimental Validation of *P. yezoensis* miRNAs. We used RT PCR to detect the expression of *P. yezoensis* miRNAs and 4 were validated by PCR amplification. PCR products of the expected sizes (60–80 bp) were amplified (Figure 4), recovered, and sequenced, increasing confidence in their

expression. Some larger PCR products might result from precursor RNAs.

4. Discussion

sRNAs of 21 nt were the most abundant in PYF, consistent with an earlier report [24], whereas in PYL the length of enriched sRNAs is 22 nt. This is different from the case in *A. thaliana*, where 24 nt sRNAs represent a major part [33, 34]. Earlier reports indicated that *P. patens* and *C. reinhardtii* also lacked the enrichment of 24 nt sRNAs [35–37]. The length of sRNAs is determined by the species of enzymes that participate in their processing. For example, DCL2 produced sRNAs of 24 nt, whereas DCL1 produced sRNAs of 21 nt [37]. The lack of enrichment of 24 nt sRNAs in *P. yezoensis*, *C. reinhardtii*, and *P. patens* indicated that the RNA processing complexes in these lower photosynthetic organisms might differ from those of *A. thaliana*. The size of the most abundant sRNAs from PYF was different from that of PYL, indicating that RNA processing enzymes and level of expression might be different within species and even in different generations of the same organism.

To identify potential known miRNAs in *P. yezoensis*, we compared all sRNAs to all known plant miRNAs in miRBase and found quantity homologs; however, these homologs did not meet the criteria we used for miRNA precursor filtering. The most straightforward interpretation for this is the lack of genome information for *P. yezoensis*, although scenarios that *P. yezoensis* contains novel miRNAs that have no identifiable homologs in other organisms cannot be ruled out. The unicellular green algae *C. reinhardtii* lacks homologous miRNAs with other organisms and even with other green algae [37]. Thus, we proposed that *P. yezoensis* has novel miRNAs that have no sequence homology with others, as for *C. reinhardtii*. We used the remaining nonannotated sRNAs to identify potential novel miRNAs in *P. yezoensis*. In all, 14 miRNAs were identified, indicating that *P. yezoensis* does have novel miRNAs that lack sequence homology with other known miRNAs.

Two mature miRNAs, Pye-miR2 and Pye-miR13, were identified earlier and designated m0001 and m0005 [24], respectively. Notwithstanding, the length of the precursor of Pye-miR2 is different from that of m0001, designated by Liang et al. [24]. Interestingly, Pye-miR13, which was

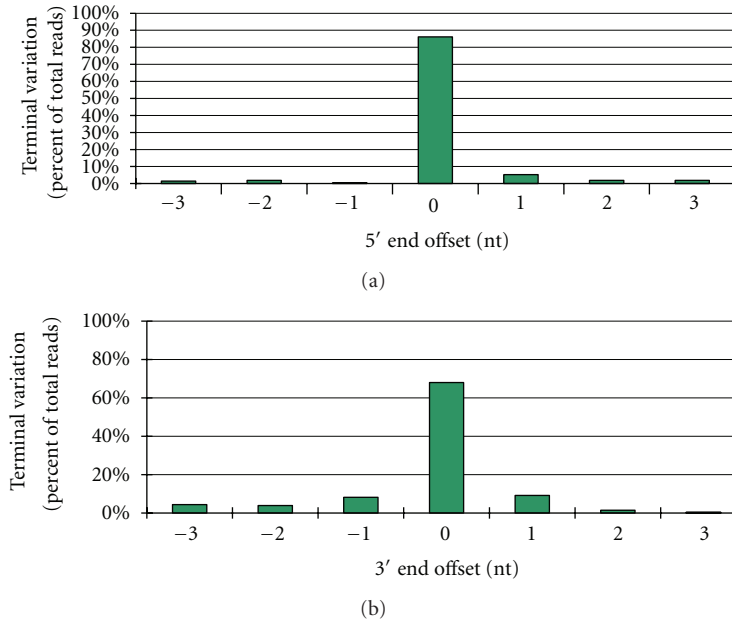


FIGURE 3: *P. yezoensis* miRNA sequence terminal variation analysis. Each unique sequence with 3' terminal nucleotide deletion or 5' terminal nucleotide extension corresponding to the mature miRNA selected is assigned a negative offset number, whereas the unique sequence with 3' terminal nucleotide extension or 5' terminal nucleotide deletion is assigned a positive offset number. In all cases, the percentage of heterogeneity for each unique *P. yezoensis* miRNA was obtained by dividing the read number of each variant by the total read number.

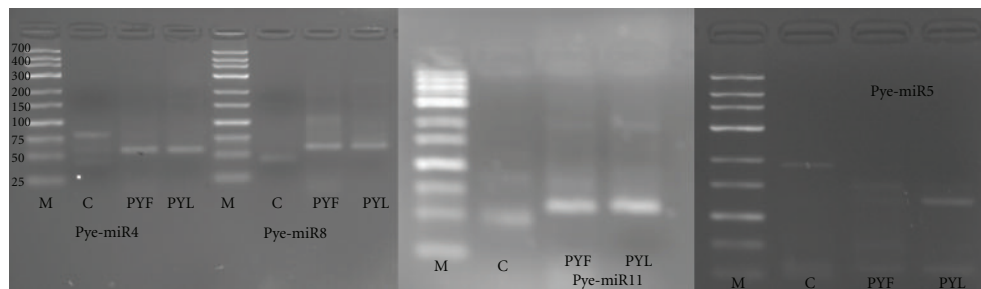


FIGURE 4: Experimental validation of some *P. yezoensis* miRNAs. M: marker; C: negative control.

sequenced from sporophyte of *P. yezoensis* and designated m0005 by Liang et al. [24], was found in gametophyte but not sporophyte of *P. yezoensis* in this study. This indicated that *P. yezoensis* might contain more miRNAs, which cannot be identified effectively by high-throughput sequencing.

Some of the homologs of known miRNAs were highly expressed. For example, homologs of miR3445 were sequenced 21,384 times in PYF and homologs of miR1442, miR1154, and miR1211 were sequenced 144,756, 41,203, and 22,573 times in PYL, respectively. miR3445 was reported only in *Arabidopsis lyrata* and was not reported even in *A. thaliana*. In this study, we sequenced homologs of miR3445 nearly 20,000 times in PYF but only 9 times in PYL, indicating its important role in PYF. Interestingly, homologs of miR1442, which were sequenced only from salt-stressed but not drought-stressed or untreated libraries of *O. sativa* [38], were most abundant in PYL with only 1 read in PYF. As mentioned in Section 1, sporophytes enter and germinate in

shells, whereas gametophytes live on static substrates of the intertidal zone and experience stress from various sources during low tide [7–9]. We proposed that homologs of miR1442 might have an important role in salt stress in PYL. Talmor-Neiman et al. [39] reported that miR1211 is expressed at a very low level in the *P. patens* gametophyte and they proposed that it might be expressed in other stages of development, such as in the sporophyte. However, we sequenced homologs of miR1211 nearly 20,000 times in PYL (gametophyte) but only 13 times in PYF (sporophyte), indicating that its role in *P. yezoensis* might be different from that in *P. patens*.

It was suggested that miRNA might participate in sexual differentiation in *Porphyra* [40]. In this study, we identified 14 miRNAs from *P. yezoensis* and only 1 was sequenced from both PYF and PYL; the others were sequenced exclusively from either PYF or PYL, indicating that different miRNAs are expressed and regulated gene expression between different

generations of *P. yezoensis*. This might be a reason for the differences of EST between the two generations [10]. Besides, the two generations have the same genome but they differ in many aspects, including habitat. Sporophytes live in shells and experience little change of environment. By contrast, gametophytes live on static substrates of the intertidal zone and experience stress caused by strong light, high temperature, and desiccation during low tide [7–9], which might result in expression of stress response-related miRNAs, as mentioned above for homologs of miR1442. Consistent results have been reported for other species. Homologs of miR2119, which might be drought responsive [41], presented in gametophyte *Porphyra haitanensis* with 1,479,099 copies. In the study of *Phaeodactylum tricornutum*, we identified more miRNAs from nitrogen-limited and silicon-limited samples [42]. This difference of expression of miRNAs in two generations of *P. yezoensis* might result in their different tolerance of various stresses. In fact, the gametophyte is very tolerant of stress caused by strong light, high temperature, and desiccation, whereas the sporophyte is not.

We intended to identify the exact functions of *P. yezoensis* miRNAs. Different targets have been suggested but, owing to the lack of genome sequences, we cannot determine whether these miRNA targets have any functional bias. Additionally, we used criteria for plant miRNA target prediction to identify *P. yezoensis* miRNA targets; so some bona fide targets might have been missed because the interaction of miRNAs with mRNA in this lower photosynthetic organism might differ from that in higher plants.

5. Conclusion

The results of this study show that *P. yezoensis* has a complex sRNA processing system containing novel miRNAs that have no identifiable homolog in other organisms and that might have important regulator roles in *P. yezoensis* development.

Abbreviations

PYF: The sporophyte of *Porphyra yezoensis*
 PYL: The gametophyte of *Porphyra yezoensis*.

Authors' Contribution

L. He and A. Huang equally contributed to this work.

Acknowledgments

The work was supported by the National Natural Science Foundation of China (41176134, 30830015, and 30970302), 973 Project (2011CB411908), nonprofit research project for the State Oceanic Administration (201105023-7) and the Qingdao High-Tech Research and Development Plan (11-2-3-8(5)-jch).

References

- [1] N. Saga and Y. Kitade, "Porphyra: a model plant in marine sciences," *Fisheries Science-Tokyo*, vol. 68, pp. 1075–1078, 2002.
- [2] C. A. Lembi and J. R. Waaland, *Algae and Human Affairs*, Cambridge University Press, 1988.
- [3] M. S. Shivji, "Organization of the chloroplast genome in the red alga *Porphyra yezoensis*," *Current Genetics*, vol. 19, no. 1, pp. 49–54, 1991.
- [4] K. M. Cole and R. G. Sheath, *Biology of the Red Algae*, Cambridge University Press, 1990.
- [5] C. Hoek, D. G. Mann, and H. M. Jahns, *Algae: An Introduction to Phycology*, Cambridge University Press, Cambridge, UK, 1995.
- [6] H. Iwasaki, "The life-cycle of *Porphyra tenera* in vitro," *The Biological Bulletin*, vol. 121, no. 1, p. 173, 1961.
- [7] A. Yamazaki, K. Nakanishi, and N. Saga, "Axenic tissue culture and morphogenesis in *Porphyra yezoensis* (Bangiales, Rhodophyta)," *Journal of Phycology*, vol. 34, no. 6, pp. 1082–1087, 1998.
- [8] Y. Kitade and S. Fukuda, "Preliminary study on early development of monospore of *Porphyra yezoensis* (Bangiales, Rhodophyta)," *Fish Genetics and Breeding Science*, vol. 28, pp. 27–34, 1999.
- [9] S. Behura and S. Sahoo, "Porphyra: the economic seaweed as a new experimental system," *Current Science*, vol. 83, no. 11, p. 1313, 2002.
- [10] E. Asamizu, M. Nakajima, Y. Kitade, N. Saga, Y. Nakamura, and S. Tabata, "Comparison of RNA expression profiles between the two generations of *Porphyra yezoensis* (rhodophyta), based on expressed sequence tag frequency analysis," *Journal of Phycology*, vol. 39, no. 5, pp. 923–930, 2003.
- [11] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, and T. Tuschl, "Identification of novel genes coding for small expressed RNAs," *Science*, vol. 294, no. 5543, pp. 853–858, 2001.
- [12] N. C. Lau, L. P. Lim, E. G. Weinstein, and D. P. Bartel, "An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*," *Science*, vol. 294, no. 5543, pp. 858–862, 2001.
- [13] R. C. Lee and V. Ambros, "An extensive class of small RNAs in *Caenorhabditis elegans*," *Science*, vol. 294, no. 5543, pp. 862–864, 2001.
- [14] B. J. Reinhart, E. G. Weinstein, M. W. Rhoades, B. Bartel, and D. P. Bartel, "MicroRNAs in plants," *Genes & Development*, vol. 16, no. 13, pp. 1616–1626, 2002.
- [15] D. P. Bartel, "MicroRNAs genomics, biogenesis, mechanism, and function," *Cell*, vol. 116, no. 2, pp. 281–297, 2004.
- [16] A. E. Pasquinelli, B. J. Reinhart, F. Slack et al., "Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA," *Nature*, vol. 408, no. 6808, pp. 86–89, 2000.
- [17] A. Bashirullah, A. E. Pasquinelli, A. A. Kiger, N. Perrimon, G. Ruvkun, and C. S. Thummel, "Coordinate regulation of small temporal RNAs at the onset of *Drosophila* metamorphosis," *Developmental Biology*, vol. 259, no. 1, pp. 1–8, 2003.
- [18] L. P. Lim, N. C. Lau, E. G. Weinstein et al., "The microRNAs of *Caenorhabditis elegans*," *Genes & Development*, vol. 17, no. 8, pp. 991–1008, 2003.
- [19] B. P. Lewis, I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, and C. B. Burge, "Prediction of mammalian microRNA targets," *Cell*, vol. 115, no. 7, pp. 787–798, 2003.
- [20] M. W. Rhoades, B. J. Reinhart, L. P. Lim, C. B. Burge, B. Bartel, and D. P. Bartel, "Prediction of plant microRNA targets," *Cell*, vol. 110, no. 4, pp. 513–520, 2002.

- [21] K. D. Kasschau, Z. Xie, E. Allen et al., "P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA function," *Developmental Cell*, vol. 4, no. 2, pp. 205–217, 2003.
- [22] J. Brennecke, D. R. Hipfner, A. Stark, R. B. Russell, and S. M. Cohen, "Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*," *Cell*, vol. 113, no. 1, pp. 25–36, 2003.
- [23] P. Xu, S. Y. Vernooy, M. Guo, and B. A. Hay, "The *Drosophila* microRNA *mir-14* suppresses cell death and is required for normal fat metabolism," *Current Biology*, vol. 13, no. 9, pp. 790–795, 2003.
- [24] C. Liang, X. Zhang, J. Zou, D. Xu, F. Su, and N. Ye, "Identification of miRNA from *Porphyra yezoensis* by high-throughput sequencing and bioinformatics analysis," *PLoS One*, vol. 5, no. 5, Article ID e10698, 2010.
- [25] R. Li, Y. Li, K. Kristiansen, and J. Wang, "SOAP: short oligonucleotide alignment program," *Bioinformatics*, vol. 24, no. 5, pp. 713–714, 2008.
- [26] S. F. Altschul, T. L. Madden, A. A. Schäffer et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, vol. 25, no. 17, pp. 3389–3402, 1997.
- [27] J. M. Calabrese, A. C. Seila, G. W. Yeo, and P. A. Sharp, "RNA sequence analysis defines Dicer's role in mouse embryonic stem cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 46, pp. 18097–18102, 2007.
- [28] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner, "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure," *Journal of Molecular Biology*, vol. 288, no. 5, pp. 911–940, 1999.
- [29] D. H. Mathews, M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker, and D. H. Turner, "Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 19, pp. 7287–7292, 2004.
- [30] E. Allen, Z. Xie, A. M. Gustafson, and J. C. Carrington, "MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants," *Cell*, vol. 121, no. 2, pp. 207–221, 2005.
- [31] R. Schwab, J. F. Palatnik, M. Riester, C. Schommer, M. Schmid, and D. Weigel, "Specific effects of microRNAs on the plant transcriptome," *Developmental Cell*, vol. 8, no. 4, pp. 517–527, 2005.
- [32] E. Bonnet, J. Wuyts, P. Rouzé, and Y. Van de Peer, "Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences," *Bioinformatics*, vol. 20, no. 17, pp. 2911–2917, 2004.
- [33] Z. Xie, E. Allen, N. Fahlgren, A. Calamar, S. A. Givan, and J. C. Carrington, "Expression of Arabidopsis MIRNA genes," *Plant Physiology*, vol. 138, no. 4, pp. 2145–2154, 2005.
- [34] R. Rajagopalan, H. Vaucheret, J. Trejo, and D. P. Bartel, "A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*," *Genes & Development*, vol. 20, no. 24, pp. 3407–3425, 2006.
- [35] T. Arazi, M. Talmor-Neiman, R. Stav, M. Riese, P. Huijser, and D. C. Baulcombe, "Cloning and characterization of microRNAs from moss," *The Plant Journal*, vol. 43, no. 6, pp. 837–848, 2005.
- [36] M. J. Axtell, J. A. Snyder, and D. P. Bartel, "Common functions for diverse small RNAs of land plants," *The Plant Cell*, vol. 19, no. 6, pp. 1750–1769, 2007.
- [37] T. Zhao, G. Li, S. Mi et al., "A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*," *Genes & Development*, vol. 21, no. 10, pp. 1190–1203, 2007.
- [38] R. Sunkar, X. F. Zhou, Y. Zheng, W. Zhang, and J. K. Zhu, "Identification of novel and candidate miRNAs in rice by high throughput sequencing," *BMC Plant Biology*, vol. 8, article 25, 2008.
- [39] M. Talmor-Neiman, R. Stav, W. Frank, B. Voss, and T. Arazi, "Novel micro-RNAs and intermediates of micro-RNA biogenesis from moss," *The Plant Journal*, vol. 47, no. 1, pp. 25–37, 2006.
- [40] N. A. Blouin, J. A. Brodie, A. C. Grossman, P. Xu, and S. H. Brawley, "Porphyra: a marine crop shaped by stress," *Trends in Plant Science*, vol. 16, no. 1, pp. 29–37, 2011.
- [41] C. Arenas-Huertero, B. Pérez, F. Rabanal et al., "Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress," *Plant Molecular Biology*, vol. 70, no. 4, pp. 385–401, 2009.
- [42] A. Huang, L. He, and G. Wang, "Identification and characterization of microRNAs from *Phaeodactylum tricornutum* by high-throughput sequencing and bioinformatics analysis," *BMC Genomics*, vol. 12, article 337, 2011.