Turk. J. Fish.& Aquat. Sci. 23(7), TRJFAS21732
https://doi.org/10.4194/TRJFAS21732

RESEARCH PAPER

Turkish Journal of FISHERIES and AQUATIC SCIENCES

Differential Gene Expression Analysis of *Neopyropia yezoensis* (Bangiales, Rhodophyta) Mutant with High-growth and Development of SSR Marker Using RNA-seq

Seo-jeong Park[1], Jong-il Choi[1,*]

[1]Chonnam National University, Department of Biotechnology and Bioengineering, Gwangju, Republic of Korea, 61186

How to Cite

Park, S., Chol, J. (2023). Differential Gene Expression Analysis of *Neopyropia yezoensis* (Bangiales, Rhodophyta) Mutant with High-growth and Development of SSR Marker Using RNA-seq. *Turkish Journal of Fisheries and Aquatic Sciences*, *23(7)*, *TRJFAS21732.* https://doi.org/10.4194/TRJFAS21732

Article History
Received 11 April 2022
Accepted 17 January 2023
First Online 23 January 2023

Corresponding Author
Tel.: +82625301846

E-mail: choiji01@jnu.ac.kr

Keywords
Differentially expressed genes,
Neopyropia yezoensis
RNA-seq
SSR marker

Abstract

Neopyropia yezoensis is a type of red macroalgae that is becoming economically important owing to increasing demand. A high-growth-rate mutant, NyEMS, was recently developed using ethyl methanesulfonate to meet this demand. However, the high-growth-rate mechanisms of NyEMS have not yet been identified because of the poor understanding of *N. yezoensis* biomolecules. In this study, NyEMS was subjected to comparative transcriptome analysis using RNA-seq. As a result, total of 7,933 differentially expressed unigenes were identified. From the transcriptome analysis, it was demonstrated that increased proteasome expression and repressed 1-aminocyclopropane-l-carboxylate (ACC) production could be responsible for the enhanced growth rate of the mutant. Additionally, genes involved in vitamin B6 biosynthesis were up-regulated, while those related to phycobilisomes were down-regulated in the mutant, which could be responsible for its characteristic features. Moreover, 1,281 SNPs and 12,963 cSSRs were identified from assembled data. To determine possible biomarkers, primer pairs were designed from the identified cDNA-derived simple sequence repeats. This study expands our understanding of the genetic mechanisms of *N. yezoensis* for future research and will be helpful in the development of *N. yezoensis* biomarkers.

Introduction

Porphyrella, a genus of red macroalgae (Bangiales), is an economically important sea crop. *Porphyrella* grows in intertidal zones, which generates abiotic stresses including nutrient starvation, desiccation, and osmotic stress as well as variation in light intensity and temperature (Davison and Pearson 1996; Sun et al., 2015).

Some *Porphyrella* species such as *Neopyropia yezoensis*, *Neopyropia tenera*, and *Neoporphyra haitanensis* are cultivated as sea vegetables in East Asian countries, including Korea, Japan, and China (Niwa et al., 2005; Wang et al., 2013).

Porphyrella production continues to increase with the growing market demand, and 1,141,710 tons of *Porphyrella* were produced globally in 2014, compared to 52,940 tons in 1990 (FAO 2018; Jiang et al., 2018). Therefore, there is a need to develop new *Porphyrella* cultivars to improve production and meet this increasing demand.

One method to meet this increased demand involves the development of fast-growing *Porphyrella* strains. A high-growth-rate *N. yezoensis* mutant, NyEMS,

was isolated by mutagenesis using ethyl methanesulfonate (Lee and Choi 2018). The NyEMS mutant had a specific growth rate that was three times higher than that of wild-type *N. yezoensis* (NyWT). However, it is difficult to interpret the mechanisms behind this higher growth rate because the available genetic information for *N. yezoensis* is insufficient, despite the availability of published drafts and plastid genomes for the species (Nakamura et al., 2013; Wang et al., 2013).

High-throughput RNA sequencing (RNA-seq) via next-generation sequencing is a suitable approach to transcriptome analysis when an accurate reference genome is unavailable (Chu and Corey 2012; Garber et al., 2011, Grabherr et al., 2011). RNA-seq was used to interpret whole-transcriptome expression profiles of selected tissues or cells, and then to analyze the interesting metabolic pathways, regulations, and gene expression (Rama Reddy et al., 2015). Studies using this method have already altered the view of the extent and complexity of eukaryotic transcriptomes (Wang et al., 2010). Next-generation sequencing technology may facilitate the analysis of non-model organisms using the available bioinformatics analysis tools and assembly programs that do not require a reference genome (*de novo* assembly) (Feldmesser et al., 2014). *De novo* transcriptome assembly has been used for various non-model organisms, including algae (Bryant et al., 2017; Carvalho et al., 2018; Im et al., 2015).

Recently, different types of biomarkers were developed in plants, including seaweeds. Biomarkers have been used to identify species, to understand the metabolic changes in the cells, and to select specific genotypes during breeding. Some types of genetic variations such as single nucleotide polymorphisms (SNPs), amplified fragment length polymorphisms, randomly amplified polymorphic DNA, and microsatellites (or simple sequence repeats, SSRs) can be used as biomarkers (Ganal et al., 2009; Kim and Misra 2007; Sun et al., 2006). Of these, SSRs and SNPs are simple, stable, convenient, and abundant as well as highly reproducible and polymorphic (Park et al., 2009; Sun et al., 2006). SSRs and SNPs not only distinguish species or strains, but also contribute to molecular mechanisms and traits (Li et al., 2019; Tang et al., 2016). There were several reports on the development of SSR and SNP biomarkers in plants (Singh et al., 2013; Thakur and Randhawa, 2018). Previous studies have attempted to develop SSR and SNP markers for *N. yezoensis*; however, there have been few studies on *N. yezoensis* biomarkers (Huang and Yan 2019a, 2019b).

Therefore, the present study aimed to understand the mechanisms underlying the high growth rate of NyEMS via transcriptome analysis using RNA-seq. RNA-seq data were assembled *de novo*, and differentially expressed genes (DEGs) between NyWT and NyEMS were identified and analyzed. Furthermore, SNPs and SSRs were identified from the assembled transcriptome, and potential biomarkers were identified using SSRs.

Materials and Methods

Cultivation of *N. yezoensis*

The mutant strain NyEMS was isolated from wild-type NyWT in a previous study; it was found to have a higher growth rate than the wild-type (Lee and Choi 2018). NyWT and NyEMS were obtained from the Seaweed Research Center (National Fisheries Research and Development Institute, Haenam, Republic of Korea) and cultivated in 0.22 µM filtered Provasoli Enriched Seawater (PES) medium (Provasoli 1963) at 10°C under a 10h light:14h dark cycle with circulating filter-sterilized air. The PES medium was replaced with fresh medium weekly.

RNA Sequencing and *de novo* Assembly

NyWT and NyEMS were harvested after 4 weeks of cultivation, and the harvested algae were frozen in liquid nitrogen. There were 3 samples of NyWT and NyEMS for RNA sequencing. RNA extraction and sequencing were performed in triplicate using the Illumina Nextseq 500 platform at the Insilicogen Company (Yongin, Republic of Korea). To prepare RNA samples, gametophytes of *N. yezoensis* were first cleaned with sterilized water. After drying with hygroscopic filter paper, the samples were ground into powder with liquid nitrogen. RNA was extracted using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The quality and quantity of the purified RNA were determined by measuring the absorbance at 260 nm/280 nm (A_{260}/A_{280}) using a Nanodrop® ND-1000 spectrophotometer (LabTech, Holliston, MA, USA). RNA quality was further verified using a 2100 Bioanalyzer RNA Nanochip (Agilent, Santa Clara, CA), and all samples had RNA Integrity Number (RIN) values greater than 6.

The cDNA libraries were constructed following the manufacturer's instructions (Illumina). Briefly, poly(A) RNA was isolated from 5 µg of total RNA using Oligo (dT) magnetic beads. Following purification, the mRNA was fragmented into small pieces and the cleaved RNA fragments were used for first stand cDNA synthesis using reverse transcriptase and random primers. This was followed by second-strand cDNA synthesis. These cDNA fragments were then purified with a QiaQuick PCR extraction kit (Qiagen, Hilden, Germany). The cDNA fragments were then connected with sequencing adapters. After agarose gel electrophoresis, the suitable fragments were selected as templates for PCR amplification to create the final cDNA library (Xie et al., 2013). The library was sequenced using Illumina HiSeq 2500 at Macrogen (Seoul, Republic of Korea).

Raw data were trimmed using Trimmomatics v0.32 (Bolger et al., 2014) to remove adapter sequences and low-quality data, and FastQC v0.11.8 (Andrews 2010) was then used to confirm the quality of the reads. *N. yezoensis* sequences were primarily distinguished from trimmed reads using the draft *N. yezoensis* genome

(Nakamura et al., 2013) and BBmap/BBsplit (Bushnell 2014).

The top 10 putative contaminant sequences (*Marivita cryptomonadis, Lewinella cohaerens, Maribacter dokdonesis, Haliea salexigens, Sphingomonadales bacterium* EhC05*, Maribacter, Bacillus, Maribacter* sp. MAR_2009_60, *Porticoccus hydrocarbonoclasticus,* and *Marinobacter* sp. C18) were eliminated from the non-mapped reads of the *N. yezoensis* draft genome. Mapped reads from the *N. yezoensis* draft genome and non-mapped and decontaminated reads were assembled in Trinity v2.8.5 (Grabherr et al., 2011) following the *de novo* assembly workflow using paired-end reads. Each of the samples was compared pairwise to identify differences across samples (Wei et al., 2011). Contaminant sequences were re-identified and removed using BLASTn (Camacho et al., 2009) and Python2.7 in-house scripts (Eric, 2011). BLASTn was performed with match length ≥200 bp, identity percentage ≥95%, and an e-value cut-off of 1×10^{-5}.

The open reading frame was identified using TransDecoder (Haas et al., 2013), and contigs were clustered by CD-HIT-EST (Li and Godzik 2006). CD-HIT-EST-2D was used to generate shared sequences between NyWT and NyEMS, and this shared sequence data were used in later analyses (Yong et al., 2014). NyWT and NyEMS are separately assembled from each sample, and then CD-HIT-EST-2D is used to identify shared sequences. The threshold for CD-HIT-EST-2D generating shared sequences was 0.7-1.0.

Functional Annotation and Differentially Expressed Genes (DEGs)

Functional annotation was conducted using a modified Trinotate workflow (Bryant et al., 2017). The NCBI non-redundant (NR), Swiss-prot, Pfam, and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases were used for functional annotation. Amino acid sequences were identified using TransDecoder, and BLASTp was performed to identify descriptive annotations in Swiss-prot and NR using BLAST+ (Camacho et al., 2009) and DIAMOND aligner (Buchfink et al., 2015), respectively, with e-value cut-offs of 1×10^{-5}. Conserved amino acid domains were annotated using Pfam and hmmscan with e-value cut-offs of 1×10^{-5} (Finn et al., 2011). Gene Ontology (GO) was annotated and classified in Swiss-prot.

Unigene expression levels were estimated as transcripts per million (TPM), and RSEM (Li and Dewey 2011) was used to estimate gene abundance with shared sequences and trimmed reads. DEseq2 (Love et al., 2014) was used to identify DEGs, with a false discovery rate significance score of ≤ 0.05 and an absolute value of $\log_2$ ratio ≥ 1. BaCelLo (Balanced subcellular Localization predictor; http://gpcr2.biocomp.unibo.it/bacello/) and TargetP (http://www.cbs.dtu.dk/services/TargetP) were used to predict subcellular localization of DEGs.

Mining Candidate Biomarkers

Single nucleotide polymorphisms and cDNA-derived SSRs (cSSRs) were identified for biomarker development. BWA-MEM v0.7.17 (Li 2013) and FreeBayes v1.2.0 (Garrison and Marth 2012) were used to mine SNPs with trimmed reads and shared sequences between NyWT and NyEMS. Putative SNPs were filtered by the phred-scaled quality score (QUAL; with a cut-off higher than 0.3) and the total read depth (with a cut-off higher than 30).

Next, cSSRs were identified using the MicroSAtellite identification tool (MISA) (Beier et al., 2017), which is commonly used to mine SSRs in plants. MISA parameters were set to their defaults. Contigs containing SSRs were randomly selected, and primers were designed using BatchPrimer3 (You et al., 2008). Twelve pairs of primers were randomly selected and the results were confirmed with PCR. PCR was performed using the cDNA of NyWT and NyEMS as templates and *pfu-x* DNA polymerase (Solgent, Daejeon, Republic of Korea) and PCR mixtures were produced according to the manufacturer's instructions. PCR for cSSR amplification was conducted under the following conditions: denaturation at 98°C for 5 min, followed by 30 cycles of denaturation at 95°C for 30 sec, annealing at a suitable temperature for the respective primer pair for 90 sec, and amplification at 74°C for 30 sec; the cycle ended after a final extension at 74°C for 30 sec and a final elongation at 74°C for 10 min. The PCR products were confirmed by electrophoresis using 1.5% agarose gel, before being analyzed for sequencing.

Quantitative Real-time PCR (qRT-PCR)

N. yezoensis was harvested after 4 weeks of cultivation. The gametophytes were dried with paper towels before being ground in a mortar and pestle with liquid nitrogen. Following the manufacturer's instructions, total RNA was extracted from the resultant algal powder using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), and genomic DNA was excluded using the RNase-Free DNase Set (Qiagen). The extracted total RNA was confirmed using a Biodrop and 1.2% agarose gel. RIN was confirmed using 2100 Bioanalyzer (Bustin et al., 2009). Complementary DNA (cDNA) was synthesized using a PrimScript 1st strand cDNA Synthesis Kit (TaKaRa, Tokyo, Japan) with oligo dT primers following the manufacturer's instructions. The amount of RNA used as a template for retro transcription was 0.9 µg following to manufacturer's instructions. qRT-PCR was conducted using TB Green™ Premix Ex Taq™ (Ti RNaseH Plus) (TaKaRa) and the Illumina Eco Real-Time system. For qRT-PCR, 20-µl reaction mixtures containing the following were used: 10 µl 2× TB Premix Ex Taq (Ti RNaseH Plus), 0.4 µl forward primer, 0.4 µl reverse primer, 6.8 µl RNase-free water, 0.4 µl ROX Reference

Dye, and 25 ng/2 µl cDNA. Amplification efficiency of the primers was 87.3-103.5%. All reactions were conducted in triplicate. *PyGAPDH* was used as an internal control, and the *PyGAPDH* primer set used in the present study was previously published by Kong et al., (2015) (Genbank accession no. AB303420.1; F-primer, 5'-GGTGCGTCCAAGCATCTGA-3'; R-primer, 5'-TGGGTGTAGTCCTGGTCGTTC-3'). The primer sets were designed from randomly selected unigenes using the Primer3 website (Untergasser et al., 2012).

All experiments were performed in triplicate, and the relative expression levels were described using the $\log_2 2^{(-\Delta\Delta Ct)}$ method, as outlined by Livak and Schmittgen (2001).

Result

De novo Transcriptome Assembly and Functional Annotation of Shared Sequences Between Wild-type and Mutant

The total raw reads measured approximately 69 Gbp, with an average length per read of 151 bp. Data on the shared sequences between NyWT and NyEMS were generated, further information is provided in Table 1. In total, 226,293 unigenes were found among the sequences shared between NyWT and NyEMS (Table 1). For the shared sequences, the N50 was 615 bp and the N10 was 1,347 bp, which were sufficient for analysis.

In total, 188,176 unigenes (83.16 %) were annotated in at least one database from shared sequences. Moreover, most unigenes of shared sequences were annotated in NCBI-nr, with 178, 902 unigenes (79.06%), followed by Pfam, GO, Swiss-prot, and KEGG with 129,540 unigenes (57.24%), 123,044 unigenes (54.37%), 116,908 unigenes (51.66%), and 101,519 unigenes (44.86%). The construction of shared sequences helped reduce the amount of data, making it more suitable for analysis. The shared sequences contained 8,114 and 47,547 contigs that were annotated as proteins of *Porphyrella* and of eukaryotes without *Porphyrella*, respectively.

Analysis of Differentially Expressed Genes (DEGs) and Gene ontology (GO) Term Enrichment

Differentially expressed genes (DEGs) were detected and used for further analysis to examine the mechanisms underlying the high growth rate of the mutant. The DEGs were detected with a false discovery rate ≤ 0.05 and a $\log_2$ fold change ≥ 1 (Figure 1 and Figure 2). Shared sequences were identified; in total, 7,933 DEGs were identified by the analysis (6,552 up-regulated unigenes and 1,381 down-regulated unigenes) and some genes related with proteasomes, ethylene production, antioxidants, and phycobilisomes were found (Table 5-7).

Eighteen GO terms were enriched in DEGs from shared sequences (Figure 3). "Protein metabolic process" (GO:0019538), "macromolecule biosynthetic process" (GO:0009059), and "cellular macromolecule biosynthetic process" (GO:0034645) were the most enriched GO terms in the biological process category. In contrast, "intracellular" (GO:0005622), "intracellular ribonucleoprotein complex" (GO:0030529), and "ribosome" (GO:0005840) were the most enriched GO terms in the cellular component category. "Nucleic acid binding" (GO:0003676), "DNA binding" (GO:0003677), and "kinase activity" (GO:0016301) were the most enriched GO terms in the molecular function category.

Table 1. Statistical data of non-redundant contigs assembled by Trinity in shared sequences between wild-type (NyWT) and mutant *Neopyropia yezoensis* (NyEMS).

Results	Shared seq.
Number of Trinity genes	226,293
Number of transcripts	267,320
GC (%)	54.48
Contig N10	1,347
Contig N50	615
Median contig length	474
Average contig	580.92
Total assembled bases	155,290,389

Table 2. Summary statistics of mined cDNA-derived single sequence repeats (SSRs) from shared sequences between wild-type (NyWT) and mutant *Neopyropia yezoensis* (NyEMS), as determined by MISA. These results were provided by MISA for SSR detection.

Searching item	Results
Total number of sequences examined	270,518
Total size of examined sequences (bp)	156,782,316
Total number of identified SSRs	12,693
Number of SSR-containing sequences	10,645
Number of sequences containing more than one SSR	1,638
Number of SSRs present in compound formation	1,411

Identification of SNPs and cSSRs

In total, 1,281 SNPs were identified in the shared sequences, and 179 and 153 SNPs were detected in only NyWT and NyEMS, respectively. Moreover, 12,963 cSSRs were identified by MISA from shared sequences (Table 2). The cSSRs were classified by repeat motifs based on mono-, di-, tri-, penta-, and hexa-nucleotides. Tri-nucleotide repeat motifs were most commonly found in the shared sequence data, with CGA (727 counts), GAC (448 counts), and ACG (353 counts) being the three most common tri-nucleotide repeat motifs. Interestingly, 18 penta-nucleotides and 84 hexa-nucleotides motifs were also counted in the mined cSSRs, and 6 hexa-nucleotide motifs were repeated more than 10 times (Table 3).

Primer pairs were designed from the identified cSSRs using BatchPrimer3. Twelve primer pairs were

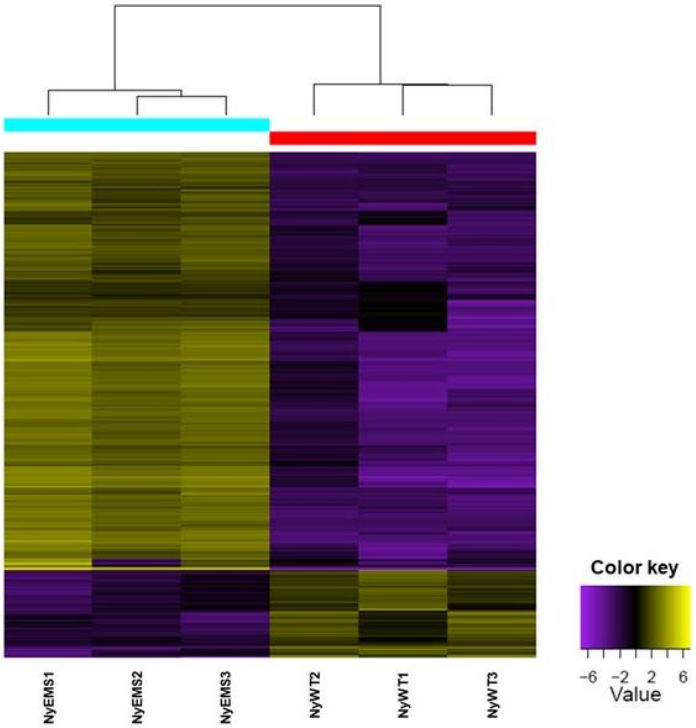

Figure 1. Hierarchical clustering heatmap of differentially expressed genes between wild-type (NyWT) and mutant *Neopyropia yezoensis* (NyEMS). The purple and yellow colors represent the down-regulated and up-regulated unigenes, respectively.

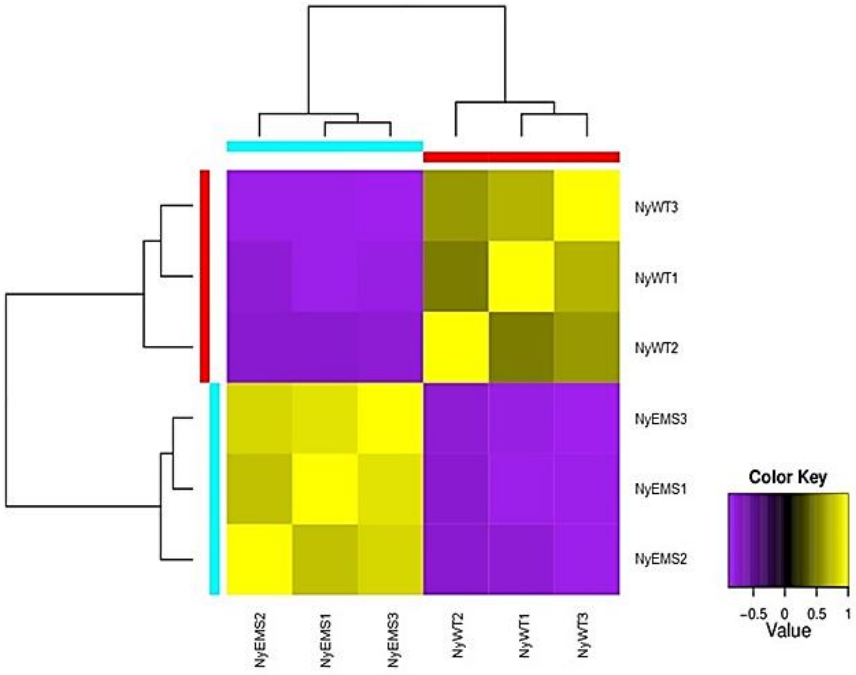

Figure 2. Hierarchical clustering heatmap based on log2-transformed count data for relationships among samples using Pearson's correlation coefficients. Purple and yellow colors represent the coefficients between samples.

randomly selected and used for PCR to identify a biomarker capable of distinguishing between NyWT and NyEMS. Eight primer pairs generated amplicons (Table 4); only one primer pair (NyCBM2) produced amplicons that showed different patterns (Figure 4). The other primer pairs generated the same amplicons between NyWT and NyEMS.

Validation of Transcriptome Analysis Using qRT-PCR

Eighteen unigenes were selected and used in qRT-PCR to validate the RNA sequencing results (Figure 5 (A)). The genes that were described in the discussion were also validated and the results were presented in Figure 5 (B). All genes showed similar expression patterns in the DEG analysis.

Discussion

NyWT and NyEMS were cultivated and prepared for mRNA sequencing under the same conditions. To eliminate the contaminants of various symbionts from *N. yezonesis* for downstream analysis, shard sequences were identified between NyWT and NyEMS (Gao et al., 2015; Ngan Tran and Choi 2020).

Assembled transcripts of shared sequences had lower GC content than the reported GC content of *N. yezoensis* genome which is approximately 64% (Nakamura et al., 2013; Wang et al., 2020). However, *N. yezoensis* genomes of mitochondria and chloroplast have low GC content of 32.7% and 33.1%, respectively (Nakamura et al., 2013; Xu et al., 2019). In this study, we did not remove organellar genomes from the assembled

Table 3. Summary statistics from MISA of the number of repeat motifs from mined cDNA-derived single sequence repeats in shared sequences between wild-type (NyWT) and mutant *Neopyropia yezoensis* (NyEMS).

Motifs	≤6	>6, ≤10	>10	Total
mono-nucleotide	0	1,603	2,441	4,044
di-nucleotide	749	534	28	1,311
tri-nucleotide	6,309	799	38	7,146
tetra-nucleotide	84	3	3	90
penta-nucleotide	13	5	0	18
hexa-nucleotide	58	20	6	84

Table 4. List of primer sequences used for single sequence repeat markers.

Name	Forward primer (5'- -3')	Revers primer (5'- -3')	T_a (°C)
NyCBM1	TTCTTTGTGTCCAACTTCTTC	AAGCGACCCATATAAGACTTC	52
NyCBM2	ACAGCAACAACCACAAATG	TCGTCAATAATGCCTGATG	52
NyCBM3	GTGGAAGGAGGAGATGATG	AGTCAAACTCGGCCACCA	55
NyCBM4	GATCTCCAGCATGAACATC	AGGCGTCGAAGTAGCTTTC	55
NyCBM5	GTGGTGGAGGTGATGTGTA	GGAGGAACGAAGGAGGTC	55
NyCBM6	CGTTCATTGGTATCAGACTTC	AAGATCGTCGTTGAACTCAG	55
NyCBM7	GAGGTTTGGGAACGGTGT	CACGACTGGTAGTAGTGCA	52
NyCBM8	CGACAACAACCTCAACAAC	CTGTTGCGACAATTGTGTT	52

Table 5. List of *Neopyropia yezoensis* unigenes annotated as relating to proteasomes.

Unigene	Fold change	Description	E-value
NyEMS_unigene_251545	1.72	Proteasome subunit beta type-7-B	1.4E-122
NyEMS_unigene_230732	1.57	26S proteasome regulatory subunit 7	2.1E-89
NyEMS_unigene_254931	2.83	26S proteasome regulatory subunit 8 homolog B	3.1E-73
NyEMS_unigene_179268	2.22	Putative 26S proteasome non-ATPase regulatory subunit 8 homolog B	2.7E-09
NyEMS_unigene_32572	2.16	20S core proteasome subunit beta 6	8.0E-57
NyEMS_unigene_249440	1.86	26S proteasome regulatory subunit 6B homolog	9.8E-161
NyEMS_unigene_10616	1.78	Probable 26S proteasome regulatory subunit rpn6	2.8E-65
NyEMS_unigene_263145	1.78	26S proteasome non-ATPase regulatory subunit 8 homolog A	8.6E-26
NyEMS_unigene_252040	1.77	26S proteasome regulatory subunit 8 homolog A	4.4E-119
NyEMS_unigene_20464	1.61	26S proteasome regulatory subunit 6B homolog	2E-198
NyEMS_unigene_236852	1.58	26S proteasome regulatory subunit	1.2E-89
NyEMS_unigene_80889	1.57	Proteasome subunit alpha type-2	1.4E-78
NyEMS_unigene_64315	3.99	26S proteasome regulatory subunit 7	2.3E-63
NyEMS_unigene_254932	4.46	26S proteasome regulatory subunit 8 homolog A	2.3E-22
NyEMS_unigene_78613	3.28	Proteasome subunit alpha type-7	3.6E-15
NyEMS_unigene_17647	2.82	Proteasome subunit beta type-5	1.9E-66
NyEMS_unigene_14513	2.53	Proteasome subunit, partial	4.4E-38
NyEMS_unigene_170864	1.52	Proteasome subunit alpha type 5	2.6E-38
NyEMS_unigene_193947	1.56	26s proteasome non-atpase regulatory	1.3E-44

transcripts to investigate all changes in gene expression in the whole algae, which would cause the low GC content of the assembled transcripts.

Some studies focusing on the genetic analysis of *N. yezoensis* have been published previously; and the newest genome of *N. yezoensis* was published in 2019 and 2020 (Wang et al., 2020; Xu et al., 2019). We aligned the *de novo* assembled transcripts with the newest *N. yezoensis* genome and organelle genome using BLAT (Kent 2002) and confirmed that approximately 70% of the *de novo* assembled transcripts were aligned to the genome.

It was difficult to distinguish between the wild-type and mutant strains using 18s rRNA sequencing. Treatment of EMS causes random mutagenesis, which changes nucleotides and large DNA structures including insertion and deletion of chromosomal segments. Microsatellite regions could be easily changed because of simple repeated sequences by mutation (Anmarkrud et al., 2008). For these reasons, SSRs were used as biomarkers to distinguish wild-type and mutant strains. cDNA-derived SSRs were investigated to develop a biomarker capable of discriminating between wild-type and mutant strains. CGA, GAC, and ACG were the three

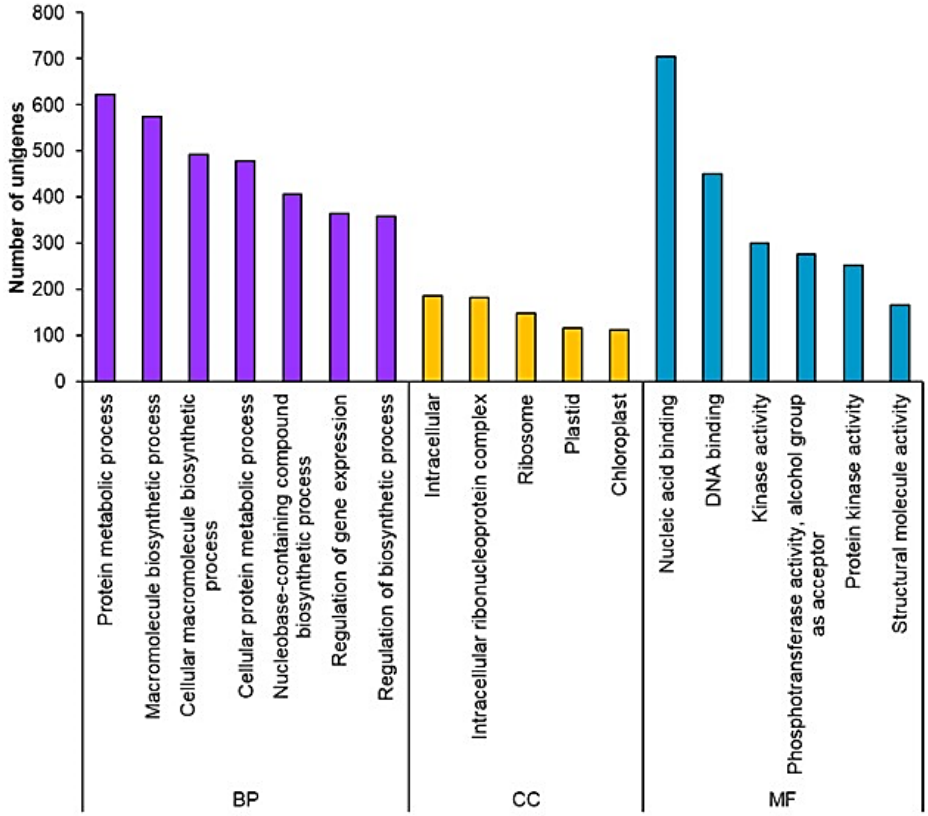

Figure 3. Distribution of gene ontology terms for differentially expressed genes in shared sequences between wild-type (NyWT) and mutant *Neopyropia yezoensis* (NyEMS). BP, biological process; CC, cellular component; MF, molecular function.

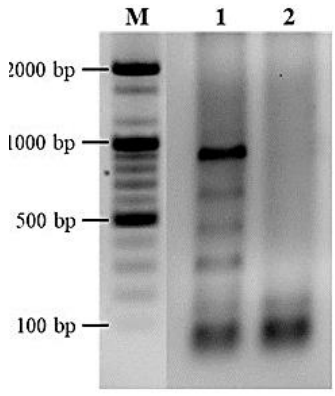

Figure 4. The cDNA-derived simple sequence repeat patterns of wild-type (NyWT) and mutant *Neopyropia yezoensis* (NyEMS) using the NyCBM2 primer pair. DNA electrophoresis was performed on a 1.5 % agarose gel. The M lane indicates the 100-bp DNA marker, and lanes 1 and 2 indicate identified amplicons from cDNA-derived SSR PCR products using NyCBM2. Lane 1, NyWT; and lane 2, NyEMS.

most common tri-nucleotide repeat motifs in the shared sequences. These three tri-nucleotide repeat motifs were not among the 10 most common tri-nucleotide repeat motifs reported in a previous study, which identified SSRs by sequencing *N. yezoensis* genomic DNA (Huang and Yan 2019b). It is plausible that this difference in the results is because we mined the SSRs from the transcriptome rather than from genomic DNA. Only the NyCBM2 primer pair produced amplicons showing different patterns between NyWT and NyEMS (Figure 4), indicating that the NyCBM2 primer pair could be used as a biomarker to distinguish between NyWT and NyEMS; however, additional experiments using more *Porphyrella* species are required to confirm this founding.

Previous proteomic analysis revealed that the expression of proteasome subunit α type 5 was up-regulated in NyEMS compared to NyWT (Lee and Choi 2018). Proteasome subunit α type 5 was also up-regulated as per the transcriptome analysis in the present study (Table 5). In addition to proteasome subunit α type 5, many unigenes annotated as proteasome subunits were up-regulated in NyEMS. The proteasome is a protein complex that possesses proteolytic activity targeted at misfolded or damaged proteins. Proteasomes control numerous signaling and metabolic pathways and can affect cell proliferation as well as expansion and stress responses (Kurepa et al., 2009; Tanaka 2009). Several small molecules known as ubiquitin are attached to the target proteins, and ubiquitination signals the shuttling of the target proteins to the proteasome.

The 26S proteasome (26SP) contains 19S regulatory particles (RPs) attached to both ends of the 20S core particle (20SP), which is also known as the cylindrical 20S proteasome and directly unfolds the target proteins (Tanaka 2009). 26SP molecules are distributed throughout the cytoplasm and particularly the nuclei of eukaryotes, suggesting that nuclear proteins require increased proteasome-dependent turnover. Proteasome subunit genes are expressed at higher levels in nuclear envelopes and mitotic spindles during plant cell division within meristems and young unexpanded organs and are expressed steadily at lower levels in mature tissues (Kurepa and Smalle 2008; Kurepa et al., 2009). The 26SP:20SP ratio is strictly regulated during development. A decrease in 26SP causes 20SP biogenesis, leading to senescence and cell death owing to the accumulation of misfolded proteins that are otherwise removed by 26SP. In *Arabidopsis* mutants, a decrease in 26SP results in accelerated senescence, indicating that 26SP is necessary to prevent premature aging (Kurepa et al., 2009). 26SP activity is increased by the up-regulation of 26SP genes in NyEMS, which may activate cell division and delay senescence. Further, delayed senescence could result in longer gametophytes in NyEMS owing to improved cell longevity, maintaining gametophytes for longer periods of time and increasing cell division.

1-aminocyclopropane-l-carboxylate (ACC) is the precursor to ethylene. Ethylene is a gaseous plant hormone that regulates root initiation, fruit ripening, flower blossoming, and plant responses to biotic and abiotic stresses (Liu et al., 2017; Nakatsuka et al., 1997). Earlier, ethylene was known to play an important role in promoting sexual reproduction in *N. yezoensis* (Uji et al., 2016); however, a recent study showed that ACC induces sexual reproduction in *N. yezoensis* independent of ethylene (Uji et al., 2020).

Table 6. List of *Neopyropia yezoensis* unigenes annotated as relating to ethylene reproduction.

Unigene		Fold change	Description	E-value
Ethylene productic	NyEMS_unigene_20617	−1.48	1-aminocyclopropane-1-carboxylate synthase 6	1.1E-78
Reproduction	NyEMS_unigene_248648	−3.81	ASPO1527	3.0E-201
	NyWT_unigene_2803	−3.00	ASPO2608	1.90E-34
	NyEMS_unigene_48773	−1.81	Aurora kinase A-A	1.1E-134
	NyEMS_unigene_95483	−2.81	Cyclin-U4-1	1.08E-66

Table 7 List of *Neopyropia yezoensis* unigenes annotated as relating to phycobilisomes and vitamin B6 biosynthesis.

	Unigene	Fold change	Description	E-value
Phycobilisomes	NyEMS_unigene_263879	−1.95	Allophycocyanin alpha subunit	9.07E-31
	NyEMS_unigene_263878	−2.08	Allophycocyanin beta subunit	6.3E-83
	NyEMS_unigene_160992	−1.31	Phycobilisome 7.8-kDa linker polypeptide, allophycocyanin-associated, core	6.1E-51
	NyEMS_unigene_261854	−1.13	Allophycocyanin gamma subunit	5.2E-85
	NyEMS_unigene_258877	−1.62	Allophycocyanin beta 18 subunit	2.2E-86
	NyEMS_unigene_264236	−1.92	Phycocyanin alpha subunit	2.9E-83
	NyEMS_unigene_264235	−1.69	Phycocyanin beta subunit	2.9E-86
	NyEMS_unigene_91456	−2.10	Phycobilisome 32.1-kDa linker polypeptide, phycocyanin-associated, rod 2	4.4E-170
	NyWT_unigene_549	−2.11	Phycobilisome 31.8-kDa linker polypeptide, phycoerythrin-associated, rod	2E-180
	NyEMS_unigene_22131	−3.37	Phycobilisome 27.9-kDa linker polypeptide, phycoerythrin-associated, rod	9.9E-118
Vitamin B6 biosynthesis	NyEMS_unigene_203220	2.11	Pyridoxamine 5′-phosphate oxidase	7.4E-32
	NyEMS_unigene_189639	−2.23	Pyridoxal 5′-phosphate synthase subunit	1.9E-50
	NyEMS_unigene_111496	1.89	Pyridoxal phosphate-dependent aminotransferase	3.4E-170

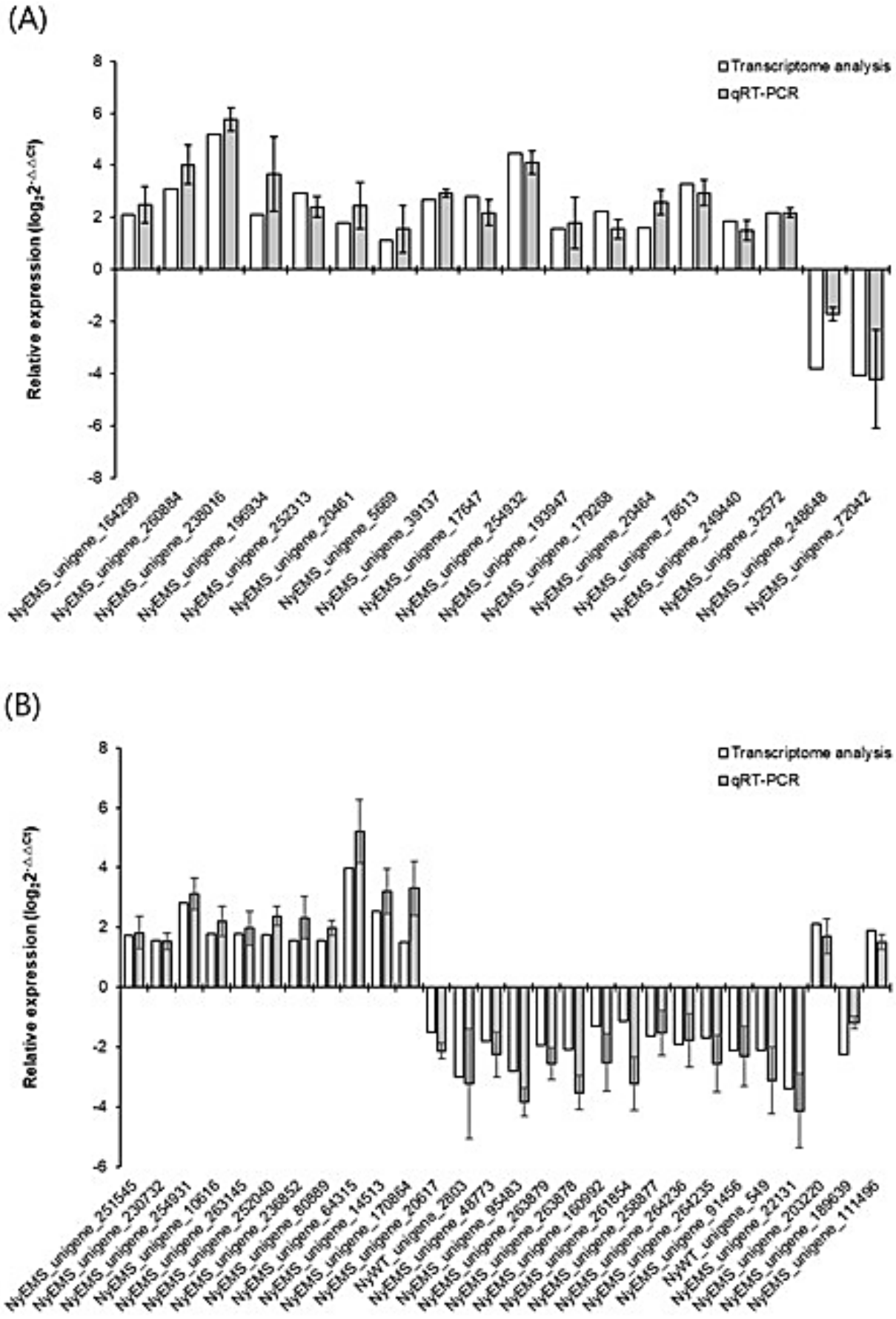

Figure 5. Results of qRT-PCR used to confirm the validity of the transcriptome analysis. (A) unigenes randomly selected, (B) unigenes described in Discussion section. White and gray boxes indicate the results of transcriptome analysis and qRT-PCR, respectively.

ACC synthase catalyzes S-adenosylmethionine, which is derived from methionine, into ACC (Nakatsuka et al., 1997). In NyEMS, ACC synthase expression was down-regulated (Table 6). Genes related to reproduction were also down-regulated in NyEMS (Table 6). Both ASPO2608 and ASPO1527 proteins have been reported to play roles in sexual and asexual reproduction in *N. yezoensis* (Kitade et al., 2008; Uji et al., 2016). As reported in a previous study, aurora kinase and cyclin, key regulators of cell division in sexual reproductive cells, were down-regulated leading to the formation of daughter cells with specific functions (Uji et al., 2020, 2016). *N. yezoensis* experiences a trade-off between vegetative growth and reproduction (Burg 1973; Uji et al., 2020, 2016). Therefore, reduced expression of ACC synthase may lead to reduced levels of ACC, and the decreased ACC levels may then suppress reproduction and enhance vegetative growth in NyEMS.

Phycobilisomes are light-harvesting antennae that are dominant in photosystem II in cyanobacteria and red algae (Watanabe et al., 2014). Phycobilisomes comprise pigmented proteins, known as phycobiliproteins. Phycoerythrin, phycocyanin, and allophycocyanin are the major classes of phycobiliproteins, with each phycobiliprotein exhibiting a different absorption wavelength (Grossman et al., 1993). Phycobilisomes absorb wavelengths of light that cannot be absorbed by chlorophyll and unidirectionally transfer energy to chlorophyll a in photosystem II (PSII). NyEMS contains less phycocyanin and other similar photosynthetic pigments than NyWT does (Lee and Choi 2018). Additionally, the expression of phycobilisome rod-core linker polypeptide was shown to be down-regulated in NyEMS. Genes related to phycobiliproteins were down-regulated in NyEMS in the present study (Table 7).

Transcriptome analysis showed that some genes related to vitamin B6 biosynthesis were up-regulated in NyEMS, similar to the proteomic results of Lee and Choi (2018) (Figure 6 and Table 7). Vitamin B6 is an effective antioxidant as well as an essential coenzyme for metabolic enzymes in organisms (Colinas et al., 2016; Tambasco-Studart et al., 2005; Titiz et al., 2006). An *Arabidopsis* mutant with accumulated vitamin B6 was shown to be more tolerant to oxidative stress, but mutants with reduced levels of vitamin B6 were found to be more sensitive to various stressors such as salinity, high light intensity, UV-B, and oxidative stress (Vanderschuren et al., 2013). Vitamin B6 has six different forms: pyridoxine, pyridoxamine, pyridoxal, pyridoxine 5'-phosphate, pyridoxamine 5'-phosphate, and pyridoxal 5'-phosphate (PLP) (Colinas et al., 2016; Tambasco-Studart et al., 2005). PLP is an active form of vitamin B6 that acts as a cofactor for many metabolic enzymes (Gerdes et al., 2012); it is synthesized by PLP synthase in addition to precursors from the glycolysis and pentose phosphate pathways. PLP synthase

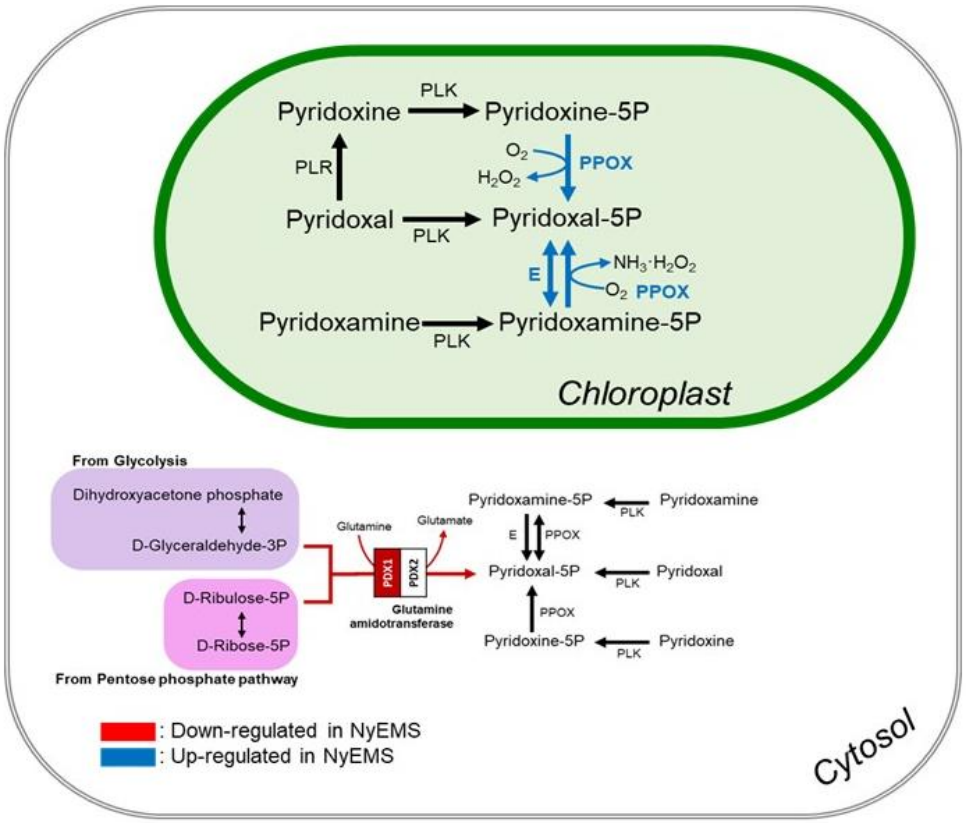

Figure 6. Description of differentially expressed genes related to vitamin B6 (pyridoxine) biosynthesis. The vitamin B6 biosynthetic pathway was described based on the description by Gerdes et al., (2012). *De novo* biosynthesis of vitamin B6 via the DXP-independent pathway has been described in the cytosol, and the salvage pathway has been described in chloroplasts.

comprises two proteins: pyridoxal phosphate synthase protein 1 (PDX1) and PDX2 (Tambasco-Studart et al., 2005). The expression of PDX1 was found to be decreased in the NyEMS transcriptome. The *pdx1*-deficient Arabidopsis mutant showed short shoot and root growth, a high chlorotic rate in leaves, and a high sensitivity to abiotic stress (Titiz et al., 2006).

However, PLP can also be synthesized from other vitamin B6 forms via the salvage pathway in the cytosol and chloroplasts (Gerdes et al., 2012). This salvage pathway was activated, and pyridoxine (pyridoxamine) 5'-phosphate oxidase (PPOX) was up-regulated in NyEMS in the present study. It has been reported that the down-regulation of PPOX decreases the levels of vitamin B6 and PLP and represses root growth and seed germination (González et al., 2007). Therefore, it is likely that the observed increase in PPOX expression compensates for the repression of PDX1 and increases vitamin B6 levels in NyEMS. The increase in vitamin B6 enhances the tolerance against photo-stress (Havaux et al., 2009). Reactive oxygen species (ROS) generated by photo-stress can degrade proteins of PSII, and vitamin B6 can protect against destruction of PSII as an antioxidant defense. The amount of phycobiliproteins is higher in NyEMS than in NyWT. However, the expression of pigment-related transcripts and proteins is less in NyEMS than in NyWT (Lee and Choi 2018), which might be due to insufficient phycobilisome in the algae because the degradation of the phycobiliproteins is inhibited in NyEMS.

In a previous transcriptome study for another *N. yezoensis* mutant (Py2K), expression levels of genes related to nitrogen availability and capacity of antioxidants were increased (Park and Choi 2020). However, in this mutant NyEMS, the expression of genes was altered to decrease the level of ACC and to increase proteasome and vitamin B6 biosynthesis. These differences could provide a variety of genetic engineering strategies for developing high-growth-rate mutants.

This transcriptome analysis will aid in further research into the genetic analysis and the production of biomarkers for *N. yezoensis*. Moreover, it could be useful in understanding the mechanisms involved in high-growth-rate in this study to develop mutant strains in seaweeds, including *N. yezoensis*

Ethical Statement

Not applicable

Funding Information

This work was supported by the National Research Foundation of Korea (NRF) grant, funded by the Korean government (MSIT; NRF-2018R1D1A1B07049359), and the Golden Seed Project Grant, funded by the Ministry of Oceans and Fisheries (213008-05-5-SB910).

Author Contribution

All authors are responsible for the general design of the manuscript. SP analyzed the data and wrote the manuscript. JC contributed to data analysis and revised the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data.* http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Anmarkrud, J.A, Kleven, O., Bachmann, L., & Lifjeld, J.T. (2008). Microsatellite evolution: Mutations, sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. BMC Evol. Biol. 8:1-10, http://doi.org.10.1186/1471-2148-8-138.

Beier, S., Thiel, T., Münch, T., Scholz, U., Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. Bioinformatics. 33:2583-2585, http://doi.org.10.1093/bioinformatics/btx198

Bolger, A.M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114-2120, http://doi.org.10.1093/bioinformatics/btu170

Bryant, D.M., Johnson, K., Ditommaso, T., Tickle, T., Couger, M.B., Payzin-Dogru, D., & Whited, J.L. (2017). A tissue-mapped axolotl *de novo* transcriptome enables identification of limb regeneration factors. Cell Rep. 18:762-776, http://doi.org.10.1016/j.celrep.2016.12.063.

Buchfink, B., Xie, C., & Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods. 12:59, http://doi.org.10.1038/nmeth.3176.

Burg, S.P. (1973). Ethylene in plant growth. PNAS. 70:591-597, http://doi.org.10.1073/pnas.70.2.591.

Bushnell, B., (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. Conference: 9th Annual Genomics of Energy & Environment Meeting. United States.

Bustin et al., (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin. Chem. 55: 611–622, http://DOI: 10.1373/clinchem.2008.112797

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics. 10:421, http://doi.org.10.1186/1471-2105-10-421.

Carvalho, E.L., Maciel, L.F., Macedo, P.E., Dezordi, F.Z., Abreu, M.E., Victoria, F.D.C., & Pinto, P.M. (2018). *De novo* assembly and annotation of the Antarctic alga *Prasiola crispa* transcriptome. Front. Mol. Biosci. 4:89, http://doi.org.10.3389/fmolb.2017.00089.

Chu, Y., Corey, & D.R. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic Acid Ther. 22:271-274, http://doi.org.10.1089/nat.2012.0367.

Colinas, M., Eisenhut, M., Tohge, T., Pesquera, M., Fernie, A.R., Weber, A.P., Fitzpatrick, T.B. (2016). Balancing of B6 vitamers is essential for plant development and

metabolism in *Arabidopsis*. Plant Cell. 28:439-453, http://doi.org.10.1105/tpc.15.01033.

Davison, I.R., & Pearson, G.A. (1996). Stress tolerance in intertidal seaweeds. J. Phycol. 32:197-211, http://doi.org.10.1111/j.0022-3646.1996.00197.x.

Eric, N. (2011). Question: How To Remove Certain Sequences From A Fasta File. https://www.biostars.org/p/4881/ (accessed 14 September 2022).

FAO. (2021). *Cultured Aquatic Species Information Programme Porphyra spp.*, http://www.fao.org/fishery/culturedspecies/Porphyra_spp/en.

Feldmesser, E., Rosenwasser, S., Vardi, A., & Ben-Dor, S. (2014). Improving transcriptome construction in non-model organisms: integrating manual and automated gene definition in *Emiliania huxleyi*. BMC Genomics. 15:148, http://doi.org.10.1186/1471-2164-15-148.

Finn, R.D., Clements, J., & Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39:W29-W37, http://doi.org.10.1093/nar/gkr367.

Ganal, M.W., Altmann, T., & Röder, M.S. (2009). SNP identification in crop plants. Curr. Opin. Plant. Biol. 12:211-217, http://doi.org.10.1016/j.pbi.2008.12.009.

Gao, F., Roy, S.W., & Katz, L.A. (2015), Analyses of alternatively processed genes in ciliates provide insights into the origins of scrambled genomes and may provide a mechanism for speciation. MBio. 6:e01998-14, http://doi.org.10.1128/mBio.01998-14.

Garber, M., Grabherr, M.G., Guttman, M., & Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. Nat. Methods. 8:469-477, http://doi.org.10.1038/nmeth.1613.

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907.

Gerdes, S., Lerma-Ortiz, C., Frelin, O., Seaver, S.M., Henry, C.S., de Crécy-Lagard, V., & Hanson, A.D. (2012). Plant B vitamin pathways and their compartmentation: a guide for the perplexed. J. Exp. Bot. 63:5379-5395, http://doi.org.10.1093/jxb/ers208.

González, E., Danehower, D., & Daub, M.E. (2007). Vitamer levels, stress response, enzyme activity, and gene regulation of *Arabidopsis* lines mutant in the pyridoxine/pyridoxamine 5'-phosphate oxidase (PDX3) and the pyridoxal kinase (SOS4) genes involved in the vitamin B6 salvage pathway. Plant Physiol. 145:985-996, http://doi.org.10.1104/pp.107.105189.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29:644-652, http://doi.org.10.1038/nbt.1883.

Grossman, A.R., Schaefer, M.R., Chiang, G.G., Collier, J.L. (1993). The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol Mol. Biol. Rev., 3:725-749, http://doi.org.10.1128/mr.57.3.725-749.1993

Grossman, A.R., Schaefer, M.R., Chiang, G.G., & Collier, J.L. (1993). The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol. Mol. Biol. Rev. 57:725-749, PMID: 8246846; PMCID: PMC372933.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., & Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8:1494-1512, http://doi.org.10.1038/nprot.2013.084.

Havaux, M., Ksas, B., Szewczyk, A., Rumeau, D., Franck, F., Caffarri, S., & Triantaphylidès, C. (2009). Vitamin B6 deficient plants display increased sensitivity to high light and photo-oxidative stress. BMC Plant Biol. 9:130, http://doi.org.10.1186/1471-2229-9-130.

Huang, L., & Yan, X. (2019a). Construction of a genetic linkage map in *Pyropia yezoensis* (Bangiales, Rhodophyta) and QTL analysis of several economic traits of blades. PLoS One. 14:e0209128, http://doi.org.10.1371/journal.pone.0209128.

Huang, L., & Yan, X. (2019b). Development of simple sequence repeat markers in *Pyropia yezoensis* (Bangiales, Rhodophyta) by high-throughput sequencing technology. Aquac. Res. 50:2646-2654, http://doi.org.10.1111/are.14222.

Im, S., Choi, S., Hwang, M.S., Park, E-J., Jeong, W-J., & Choi, D-W. (2015). *De novo* assembly of transcriptome from the gametophyte of the marine red algae *Pyropia seriata* and identification of abiotic stress response genes. J. Appl. Phycol. 27:1343-1353, http://doi.org.10.1007/s10811-014-0406-3.

Jiang, H., Zou, D., Lou, W., Deng, Y., & Zeng, X. (2018). Effects of seawater acidification and alkalization on the farmed seaweed, *Pyropia haitanensis* (Bangiales, Rhodophyta), grown under different irradiance conditions. Algal Res. 31:413-420, http://doi.org.10.1016/j.algal.2018.02.033.

Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. Genome Res. 12:656-664, http://doi.org.10.1101/gr.229202.

Kim, S., & Misra, A. (2007). SNP genotyping: technologies and biomedical applications. Annu. Rev. Biomed. Eng. 9:289-320, http://doi.org.10.1146/annurev.bioeng.9.060906.152037.

Kitade, Y., Asamizu, E., Fukuda, S., Nakajima, M., Ootsuka, S., Endo, H., & Saga, N. (2008). Identification of genes preferentially expressed during asexual sporulation in *Porphyra yezoensis* gametophytes (Bangiales, Rhodophyta) 1. J. Phycol. 44:113-123, http://doi.org.10.1111/j.1529-8817.2007.00456.x.

Kong, F., Cao, M., Sun, P., Liu, W., & Mao, Y. (2015). Selection of reference genes for gene expression normalization in *Pyropia yezoensis* using quantitative real-time PCR. J. Appl. Phycol. 27:1003-1010, http://doi.org.10.1007/s10811-014-0359-6.

Kurepa, J., & Smalle, J.A. (2008). Structure, function and regulation of plant proteasomes. Biochimie. 90:324-335, http://doi.org.10.1016/j.biochi.2007.07.019.

Kurepa, J., Wang, S., Li, Y., & Smalle, J. (2009). Proteasome regulation, plant growth and stress tolerance. Plant Signal Behav. 4:924-927 http://doi.org.10.4161/psb.4.10.9469.

Lee, H-J., & Choi, J-I. (2018). Isolation and characterization of a high-growth-rate strain in *Pyropia yezoensis* induced by ethyl methane sulfonate. J. Appl. Phycol. 30:2513-2522, http://doi.org.10.1007/s10811-018-1426-1.

Li, B., & Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a

reference genome. BMC Bioinformatics. 12:323, http://doi.org.10.1186/1471-2105-12-323.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997.

Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 22:1658-1659, http://doi.org.10.1093/bioinformatics/btl158.

Li, W., Riday, H., Riehle, C., Edwards, A., Dinkins, R. (2019). Identification of single nucleotide polymorphism in red clover (Trifolium pratense L.) using targeted genomic amplicon sequencing and RNA-seq. Front Plant Sci., 10:1257, https://doi.org/10.3389/fpls.2019.01257

Liu, X., Bogaert, K., Engelen, A.H., Leliaert, F., Roleda, M.Y., & De Clerck, O. (2017). Seaweed reproductive biology: environmental and genetic controls. Bot. Mar. 60:89-108, http://doi.org.10.1515/bot-2016-0091.

Livak, K.J., & Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods. 25:402-408, http://doi.org.10.1006/meth.2001.1262.

Love, M.I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15:550, http://doi.org.10.1186/s13059-014-0550-8.

Nakamura, Y., Sasaki, N., Kobayashi, M., Ojima, N., Yasuike, M., Shigenobu, Y., & Ikeo, K. (2013). The first symbiont-free genome sequence of marine red alga, Susabi-nori (*Pyropia yezoensis*). PLoS One. 8:e57122, http://doi.org.10.1371/journal.pone.0057122.

Nakatsuka, A., Shiomi, S., Kubo, Y., & Inaba, A. (1997). Expression and internal feedback regulation of ACC synthase and ACC oxidase genes in ripening tomato fruit. Plant Cell Physiol. 38:1103-1110, http://doi.org.ss10.1093/oxfordjournals.pcp.a029094.

Ngan Tran, K., & Choi, J-I. (2020). Comparative transcriptome analysis of high-growth and wild-type strains of *Pyropia yezoensis*. Acta Bot. Croat. 79, http://doi.org.10.37427/botcro-2020-020.

Niwa, K., Kikuchi, N., & Aruga, Y. (2005). Morphological and molecular analysis of the endangered species *Porphyra tenera* (Bangiales, Rhodophyta)1. J. Phycol. 41:294-304, http://doi.org.10.1111/j.1529-8817.2005.04039.x.

Park, S-J., & Choi, J-I. (2020). *De novo* transcriptome analysis of high growth rate *Pyropia yezoensis* (Bangiales, Rhodophyta) mutant with high utilization of nitrogen. Acta Bot. Croat. 79, http://doi.org.10.37427/botcro-2020-026.

Park, Y-J., Lee, J.K., & Kim, N-S. (2009). Simple sequence repeat polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. Molecules. 14:4546-4569, http://doi.org.10.3390/molecules14114546.

Provasoli, L. (1963). Organic regulation of phytoplankton fertility. In Hill, M.N. (Ed.) *The composition of sea-water: comparative and descriptive oceanography* (pp. 165–219). Interscience Pub.

Rama Reddy, N.R., Mehta, R.H., Soni, P.H., Makasana, J., Gajbhiye, N.A., Ponnuchamy, M., Kumar, J. (2015) Next generation sequencing and transcriptome analysis predicts biosynthetic pathway of sennosides from Senna (Cassia angustifolia Vahl.), a non-model plant with potent laxative properties. PLoS ONE, 10(6):e0129422, https://doi.org/10.1371/journal.pone.0129422

Singh, N., Choudhury, D.R., Singh, A.K., Kumar, S., Srinivasan, K., Tyagi, R.K., Singh, N.K., & Singh, R. (2013). Comparison of SSR and SNP Markers in Estimation of Genetic Diversity and Population Structure of Indian Rice Varieties. PLoS One. 2013; 8(12): e84136. https://doi: 10.1371/journal.pone.0084136

Sun, J., Liu, T., Guo, B., Jin, D., Weng, M., Feng, Y., & Wang, B. (2006). Development of SSR primers from EST sequences and their application in germplasm identification of *Porphyra lines* (Rhodophyta). Eur. J. Phycol. 41:329-336, http://doi.org.10.1080/09670260600740906.

Sun, P., Mao, Y., Li, G., Cao, M., Kong, F., Wang, L., & Bi, G. (2015) Comparative transcriptome profiling of *Pyropia yezoensis* (Ueda) MS Hwang & HG Choi in response to temperature stresses. BMC Genomics. 16:463, http://doi.org.10.1186/s12864-015-1586-1.

Tambasco-Studart, M., Titiz, O., Raschle, T., Forster, G., Amrhein, N., & Fitzpatrick, T.B. (2005). Vitamin B6 biosynthesis in higher plants. PNAS. 102:13687-13692, http://doi.org.10.1073/pnas.0506228102.

Tanaka, K. (2009). The proteasome: overview of structure and functions. Proc. Jpn. Acad. B Phys. Biol. Sci. 85:12-36, http://doi.org.10.2183/pjab.85.12.

Tang, W., Wu, T., Ye, J., Sun, J., Jiang, Y., Yu, J., Tang, J., Chen, G., Wang, C., Wan, J. (2016) SNP-based analysis of genetic diversity reveals important alleles associated with seed size in rice. BMC Plant Biol., 16:93, https://doi.org/10.1186/s12870-016-0779-3

Thakur, O. & Randhawa, S.S. (2018). Identification and characterization of SSR, SNP and InDel molecular markers from RNA-Seq data of guar (*Cyamopsis tetragonoloba*, L. Taub.) roots.BMC Genomics ,19: 951, https://doi.org/10.1186/s12864-018-5205-9

Titiz, O., Tambasco-Studart, M., Warzych, E., Apel, K., Amrhein, N., Laloi, C., Fitzpatrick, T.B. (2006) PDX1 is essential for vitamin B6 biosynthesis, development and stress tolerance in Arabidopsis. Plant J, 48(6):933-946, http://doi.org.10.1111/j.1365-313X.2006.02928.x.

Titiz, O., Tambasco-Studart, M., Warzych, E., Apel, K., Amrhein, N., Laloi, C., & Fitzpatrick, T.B. (2006). PDX1 is essential for vitamin B6 biosynthesis, development and stress tolerance in *Arabidopsis*. Plant J. 48:933-946, http://doi.org.10.1111/j.1365-313X.2006.02928.x.

Uji, T., Endo, H., & Mizuta, H. (2020). Sexual Reproduction via a 1-Aminocyclopropane-1-Carboxylic Acid-Dependent Pathway Through Redox Modulation in the Marine Red Alga *Pyropia yezoensis* (Rhodophyta). Front Plant Sci. 11:60, http://doi.org.10.3389/fpls.2020.00060.

Uji, T., Matsuda, R., Takechi, K., Takano, H., Mizuta, H., & Takio, S. (2016). Ethylene regulation of sexual reproduction in the marine red alga *Pyropia yezoensis* (Rhodophyta). J. Appl. Phycol. 28:3501-3509, http://doi.org.10.1007/s10811-016-0904-6.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., & Rozen, S.G. (2012). Primer3—new capabilities and interfaces. Nucleic Acids Res. 40:e115-e115, http://doi.org.10.1093/nar/gks596.

Vanderschuren, H., Boycheva, S., Li, K-T., Szydlowski, N., Gruissem, W., & Fitzpatrick, T.B. (2013). Strategies for vitamin B6 biofortification of plants: a dual role as a micronutrient and a stress protectant. Front Plant Sci. 4:143, http://doi.org.10.3389/fpls.2013.00143.

Wang, D., Yu, X., Xu, K., Bi, G., Cao, M., Zelzion, E., & Mao, Y. (2020). *Pyropia yezoensis* genome reveals diverse

mechanisms of carbon acquisition in the intertidal environment. Nat. Commun. 11:1-11, http://doi.org.10.1038/s41467-020-17689-1.

Wang, L., Mao, Y., Kong, F., Li, G., Ma, F., Zhang, B., Sun, P., Bi, G., Zhang, F., & Xue, H. (2013). Complete sequence and analysis of plastid genomes of two economically important red algae: *Pyropia haitanensis* and *Pyropia yezoensis*. PLoS One. 8:e65902, https://doi.org/10.1371/journal.pone.0065902

Wang, Z., Gerstein., &Snyder M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 10(1): 57–63., http://doi: 10.1038/nrg2484

Watanabe, M., Semchonok, D.A., Webber-Birungi, M.T., Ehira, S., Kondo, K., Narikawa, R., & Ikeuchi, M. (2014). Attachment of phycobilisomes in an antenna–photosystem I supercomplex of cyanobacteria. PNAS. 111:2512-2517, http://doi.org.10.1073/pnas.1320599111.

Wei, W., Qi, X., Wang, L., Zhang, Y., Hua, W., Li, D., & Zhang, X. (2011). Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. BMC Genomics. 12:451, http://doi.org.10.1186/1471-2164-12-451.

Xie, C., Li, B., Xu, Y., Ji, D., & Chen, C. (2013) Characterization of the global transcriptome for Pyropia haitanensis (Bangiales, Rhodophyta) and development of cSSR markers. BMC Genomics. 14: 107, http://doi: 10.1186/1471-2164-14-107

Xu, K., Yu, X., Tang, X., Kong, F., & Mao, Y. (2019). Organellar genome variation and genetic diversity of Chinese *Pyropia yezoensis*. Front Mar. Sci. 6:756, http://doi.org.10.3389/fmars.2019.00756.

Yong, H-Y., Zou, Z., Kok, E-P., Kwan, B-H., Chow, K., Nasu, S., & Nishio, T. (2014) Comparative transcriptome analysis of leaves and roots in response to sudden increase in salinity in *Brassica napus* by RNA-seq. Biomed. Res. Int. 2014, http://doi.org.10.1155/2014/467395.

You, F.M., Huo, N., Gu, Y.Q., Luo, M.C., Ma, Y., Hane, D., & Anderson, O.D. (2008). BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. 9:253, http://doi.org.10.1186/1471-2105-9-253.