

RESEARCH ARTICLE

# Transcriptome analysis of the typical freshwater rhodophytes *Sheathia arcuata* grown under different light intensities

Fangru Nan, Jia Feng, Junping Lv, Qi Liu, Shulian Xie\*

School of Life Science, Shanxi University, Taiyuan, China

\* [xiesl@sxu.edu.cn](mailto:xiesl@sxu.edu.cn)



## Abstract

The Rhodophyta *Sheathia arcuata* is exclusively distributed in freshwater, constituting an important component in freshwater flora. This study presents the first transcriptome profiling of freshwater Rhodophyta taxa. A total of 161,483 assembled transcripts were identified, annotated and classified into different biological categories and pathways based on BLAST against diverse databases. Different gene expression patterns were caused principally by different irradiances considering the similar water conditions of the sampling site when the specimens were collected. Comparison results of gene expression levels under different irradiances revealed that photosynthesis-related pathways significantly up-regulated under the weak light. Molecular responses for improved photosynthetic activity include the transcripts corresponding to antenna proteins (LHCA1 and LHCA4), photosynthetic apparatus proteins (PSBU, PETB, PETC, PETH and beta and gamma subunits of ATPase) and metabolic enzymes in the carbon fixation. Along with photosynthesis, other metabolic activities were also regulated to optimize the growing and development of *S. arcuata* under appropriate sunlight. Protein-protein interactive networks revealed the most responsive up-expressed transcripts were ribosomal proteins. The *de-novo* transcriptome assembly of *S. arcuata* provides a foundation for further investigation on the molecular mechanism of photosynthesis and environmental adaption for freshwater Rhodophyta.

## OPEN ACCESS

**Citation:** Nan F, Feng J, Lv J, Liu Q, Xie S (2018) Transcriptome analysis of the typical freshwater rhodophytes *Sheathia arcuata* grown under different light intensities. PLoS ONE 13(5): e0197729. <https://doi.org/10.1371/journal.pone.0197729>

**Editor:** Y-h. Taguchi, Chuo University, JAPAN

**Received:** December 7, 2017

**Accepted:** May 8, 2018

**Published:** May 29, 2018

**Copyright:** © 2018 Nan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data files are available from the GenBank database (accession numbers PRJNA421565, PRJNA421415, PRJNA421429 and PRJNA421431).

**Funding:** We are grateful to the National Natural Science Foundation of China (grant number 31670208 to Shulian Xie) and the Fund for Shanxi “1331 Project” Key Innovative Research Team for funding this project.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

The Rhodophyta constitutes an ancient derived monophyletic eukaryotic lineage. As a member of archaeplastida, Rhodophyta originated from the primary photosynthetic endosymbiosis and subsequently spread plastid through secondary endosymbiosis to a diverse array of photosynthetic lineages [1, 2]. They are primarily marine in distribution, with less than 3% of the over 6500 species occurring in truly freshwater habitats [3, 4]. Though owning a relatively low diversity compared with the marine group, freshwater rhodophytes are usually important constituents of stream floras, either in terms of abundance or distribution from local scale to biomes [5]. Genus *Sheathia* is a typical freshwater Rhodophyta and inhabited exclusively in streams or rivers. It belongs to the Florideophyceae, growing as gelatinous gametophyte

filaments, with beaded appearance, varying from blue-green, olive, violet, and gray to brownish. *Sheathia* can be found the year round but most abundant in late winter and spring, with the growing rate accelerating in December and decreasing in June throughout a year [6, 7]. Species of *Sheathia* are reported worldwide, and numerous species have been collected from different continents. *S. arcuata* is one of the most widespread species in the genus and has recently been reported from numerous localities [8].

Light is one of the important environmental factors regulating photosynthesis, growth and reproduction of photosynthetic organisms. Physiological responses to changing light intensity have been examined extensively [9, 10]. Variation of growth rate, pigment content and photosynthetic characteristics in response to irradiance have been investigated in freshwater red algae [10, 11]. However, little is currently known regarding the molecular mechanisms affecting the regulatory and biochemical pathways of freshwater red algae *Sheathia* in response to irradiance. Previous report has confirmed that *Sheathia* was typically shade-adapted plants, whereas some species can tolerate high irradiances and have mechanisms to avoid photo damage [10]. Thus, analyzing the gene expression patterns in response to different irradiance will provide a molecular basis for their environment adaption. Transcriptome analysis using next-generation sequencing is a powerful tool for examining complex molecular mechanisms. It provides a complete reference profile to understand genome content, gene function, gene expression under various conditions [12]. High-throughput RNA-sequencing (RNA-Seq) provides new perspectives for analyzing functional complexity of transcriptomes [13–15]. It has been used to analyze different gene expression patterns of different morphological types or under different conditions in higher plants [16, 17]. Whereas the transcription profiling is still unknown for freshwater Rhodophyta.

In this study we presented the transcriptome profile of the typical freshwater taxa *S. arcuata*, analyzed the coding gene contents and function annotations based on BLAST against multiple databases. The significantly different expressed genes under different irradiances were analyzed, thus laying a foundation for investigation on the molecular mechanism for the environmental adaption of freshwater Rhodophyta.

## Materials and methods

### Sample collection and preparation

Samples of *S. arcuata* were collected in Nanlaoquan, Jinci Park, Shanxi province, China (37° 42' 24.02"N; 112° 26' 31.76"E) on June 20th and December 22nd, 2015. The park where the samples were collected is open to public and no specific permissions are requested for field sampling, and we confirm that the field studies did not involve endangered or protected species. According to the statistical data of Monthly Averaged Clear Sky Insolation Incident On A Horizontal Surface in Taiyuan (<https://eosweb.larc.nasa.gov/cgi-bin/sse/grid.cgi?email=zhenhuawan@gmail.com>), the collection dates were selected when strongest and weakest light intensities occurred empirically. Both collection days were sunny and the light intensities were measured using a curing radiometer. Specimens of *Sheathia* growing at the same location with similar wet weight were collected at different dates. Other parameters related to the water conditions were measured with pH & EC waterproof (HANNA instruments, Woonsocket RI USA) when samples were collected, and the results were showed in Table 1. Physiochemical factors including temperature, pH, current velocity, total dissolved solids and electrical conductivity of the underground water in the sampling site were relative stable, except for the considerable different light intensities.

The thalli were washed using distilled water and frozen in liquid nitrogen as soon as possible after collection at the sampling site. Total RNA of each specimen was extracted according

**Table 1. Average water parameters over multiple samplings of the sampling site in this study.**

Sampling date	strains	water temperature	pH	Current velocity	TDS (total dissolved solids)	EC (electrical conductivity)	light intensity ( $\mu\text{mol photons/m}^2/\text{s}$ )
June 20th, 2015	12	16.5°C	6.23	13 cm/s	998	498	1462
December 22nd, 2015	12	16.0°C	6.23	13 cm/s	996	486	274

<https://doi.org/10.1371/journal.pone.0197729.t001>

to Holmes and Bonner [18]. After the sample were treated with DNase, RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was checked using the Nano-Photometer® spectrophotometer (IMPLEN, CA, USA). Concentration of RNA was quantified using Qubit (Thermo Fisher Scientific) and integrity of RNA was tested using Agilent 2100 (Agilent technology). RNA samples used for subsequent analyses were with values of A260/A280 ratios between 1.9 and 2.1, RNA 28S:18S ratios higher than 1.0, and RNA integrity numbers (RINs)  $\geq 6.8$ . The extracted RNA samples of each group (high light intensity and low light intensity) were pooled from 3 individual thalli before subsequent handling.

### Library preparation for transcriptome sequencing

A total amount of 1.5  $\mu\text{g}$  RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H). Second strand cDNA synthesis was subsequently performed using DNA polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3  $\mu\text{l}$  USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37°C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform and paired-end reads were generated.

### Transcriptome analysis

Clean reads were produced by removing reads containing adapter, reads containing ploy-N and low quality reads from raw data. At the same time, Q20 (corresponding to sequencing quality with 99% accuracy rates) and GC-content of the clean data were calculated. All the downstream analyses were based on clean data with high quality. Transcriptome assembly was accomplished based on the clean data using Trinity [19] with min\_kmer\_cov set to 2 and all other parameters were set default. Gene function was annotated based on BLAST search

against the following seven databases: Nr (NCBI non-redundant protein sequences); Nt (NCBI non-redundant nucleotide sequences), Swiss-Prot (a manually annotated and reviewed protein sequence database) with  $10^{-5}$  e-value cutoff and KOG/COG (Clusters of Orthologous Groups of proteins) with  $10^{-3}$  e-value cutoff. Automatic annotation ServerKO (KEGG Ortholog database) was conducted with  $10^{-10}$  e-value cutoff. GO (Gene Ontology) annotation was conducted using Blast2GO v2.5 [20] with  $10^{-6}$  e-value and customized script. Pfam (Protein family) annotation was based on hmmscan in the HMMER 3.0 package with e-value 0.01 [21, 22].

### Quantification of gene expression levels and differential expression analysis

Gene expression levels were estimated by RSEM [23] for each sample, with the clean data mapping back onto the assembled transcriptome and read count for each gene was obtained from the mapping results. Prior to differential gene expression analysis, the read counts were adjusted by edgeR program through one scaling normalized factor for each sequenced library, which was designed for datasets with no biological replicates [24]. Differential expression analysis of two samples was performed using the DEGseq package [25, 26]. P-value was adjusted using q-value [27].  $q\text{-value} < 0.005$  and  $|\log_2(\text{foldchange})| > 1$  were set as the threshold for significantly differential expression.

### Quantitative real-time PCR (qRT-PCR) validation

Total RNA extracted for library preparation were used as template of RT-PCR. Reverse transcribed cDNA were used to conduct quantitative real-time PCR with SYBR Green Dye (TakaRa SYBR Premix Ex Taq II). Five genes from differentially expressed gene pools based on the bioinformatics analysis were selected including *psbU*, *LHCA4*, *petH*, *petB* and *petC*. The translation initiation factor 5A (*elf5a*) was selected as internal control gene according to previous literature [28]. Amplification primers for selected genes were shown in S1 Table. The amplification procedures were 95°C for 30 s, 40 cycles of 95°C for 5 s, 53°C for 30 s, followed by dissolution stage of 95°C for 15 s, 53°C for 1 min and 95°C for 15 s. Specificity of the qPCR products was estimated based on melting curve. Expression values of each gene were calculated using the method proposed by Pfaffl [29]. Results of qPCR and RNA-seq data for the selected genes were compared.

### KEGG pathway enrichment analysis and Protein-protein interactive network construction

To understand the different functional pathways between the two samples, we used KOBAS software to test the statistical enrichment of differential expressed genes in KEGG pathways [30, 31]. The sequences of the differentially expressed genes (DEGs) were BLAST to the genome of *Cyanidioschyzon merolae*, which was available of the protein-protein interaction (PPI) in the STRING database (<http://string-db.org/>) to get the predicted PPI of these DEGs. BLAST settings for constructing interaction networks were  $\text{evalue} = 1\text{e-}10$  and  $\text{max\_target\_seqs} = 1$ . Then the PPI of these DEGs were visualized in Cytoscape [32].

## Results

### Transcriptome sequencing and assembly

Two cDNA libraries prepared from samples collected at different seasons were sequenced using the Illumina HiSeq 2000 platform, producing database of 3.9 and 4.5 gigabyte

**Table 2. Statistics of RNA-sequencing quality of *S. arcuata* samples under different light intensities.**

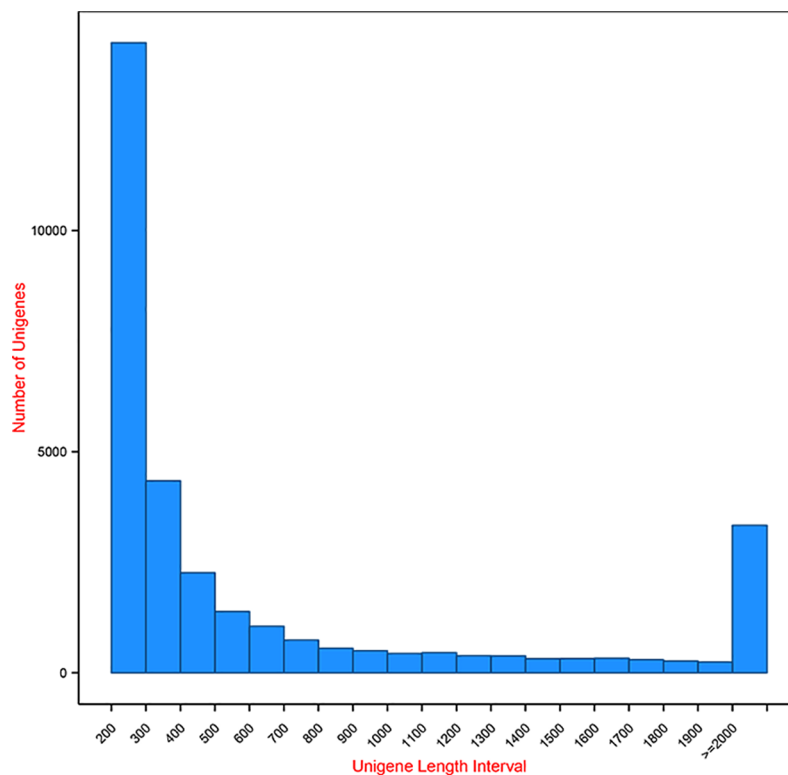
Sample	Raw reads	Clean reads	Clean bases (bp)	Error (%)	Q20 (%)	Q30 (%)	GC content (%)
low light	27499326	26661934	3.33G	0.01	97.49	94.19	54.98
high light	37608294	34837166	4.35G	0.01	97.26	93.95	55.70

<https://doi.org/10.1371/journal.pone.0197729.t002>

respectively. Raw reads were 27.50 and 37.61 million paired-end reads for the algal sample respectively. These reads were 125 bp in length with high quality after reads filtering. After quality control approximately 26.66 and 34.84 million clean reads were obtained with similar GC content (Table 2). The sequence reads generated in this study have been deposited in GenBank under the accession numbers of PRJNA421565, PRJNA421415, PRJNA421429 and PRJNA421431. Mixed assembly of the clean reads generated 161,483 transcripts with the most abundant length interval of 200–300 bp (Fig 1). Transcripts with lengths ranging from 200–500 bp, 500–1000 bp, 1000–2000 bp and ≥ 2000 bp accounted for 87.71% (141,633), 7.62% (12,292), 2.31% (3,732), and 2.37% (3,826) of the total transcripts respectively.

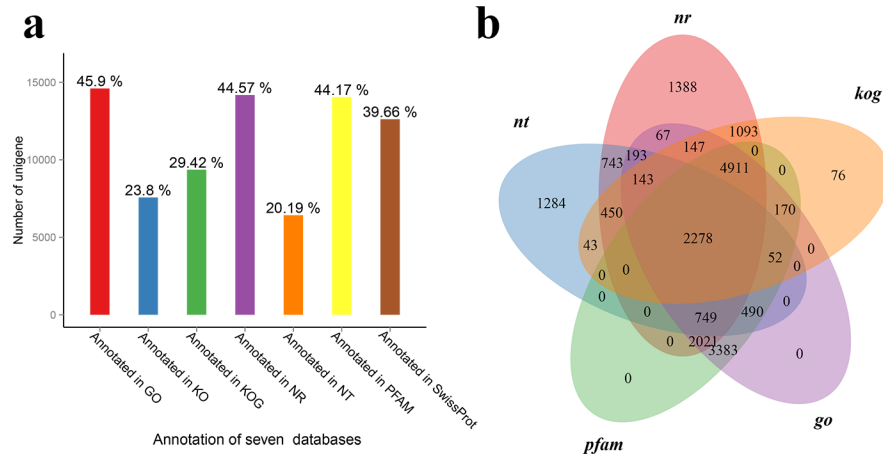
### Gene annotations

All 161,483 assembled transcripts were queried against seven curated databases (Fig 2A). Databases including Nr, Nt, KOG, GO and Pfam were selected to illustrate annotation venn diagram (Fig 2B). 2278 common genes were shared in the five annotation databases. Based on the Nr annotation result (Fig 3), the species with most homologous genes with *Sheathia* was the *Chondrus crispus* (marine Rhodophyta), followed by *Oryza sativa* (green plant), *Galdieria sulphuraria* (thermophilic Rhodophyta) and *Phaeodactylum tricoratum* (Bacillariophyta).



**Fig 1. Annotated transcript-bar from clean reads of *S. arcuata*.**

<https://doi.org/10.1371/journal.pone.0197729.g001>

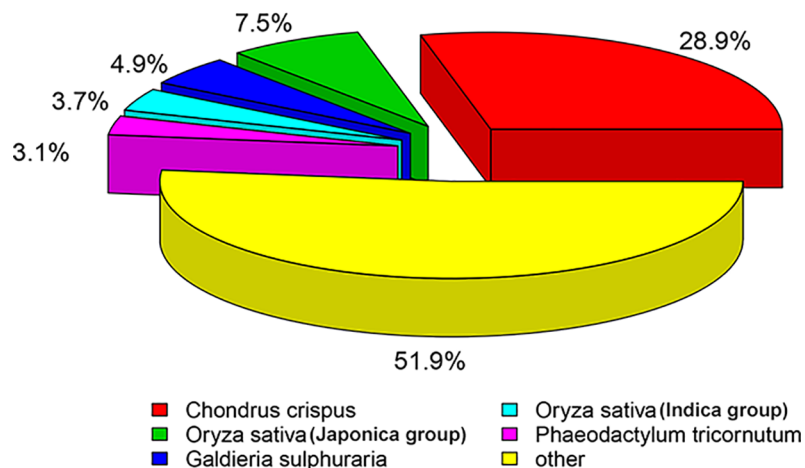


**Fig 2. Function annotations of transcripts of *S. arcuata* based on BLAST against diverse databases.** a. Numbers of transcripts annotated in seven databases; b. Venn diagram of transcripts annotated in five databases.

<https://doi.org/10.1371/journal.pone.0197729.g002>

The predicted *S. arcuata* transcripts were classified according to GO assignments [33]. A total of 62,008 genes (38.4%) were assigned at least one GO term (Fig 4), among which 147,479 were assigned in the biological process category (Level 1), 70,427 in the molecular function category (Level 1) and 95,177 in the cellular component category (Level 1). These transcripts were further classified into functional subcategories. Genes corresponding to the “biological process” group (Level 1) were divided into 24 subcategories, among which “cellular process” (Level 2) comprised 22.6% and was the largest term. Genes corresponding to the “molecular function” group (Level 1) were divided into 10 subcategories, among which “binding” (Level 2) comprised 44.9% and was the largest term. Genes corresponding to the “cellular

### Species classification



**Fig 3. Classification of species with gene homology to *S. arcuata* based on Nr annotation result.**

<https://doi.org/10.1371/journal.pone.0197729.g003>



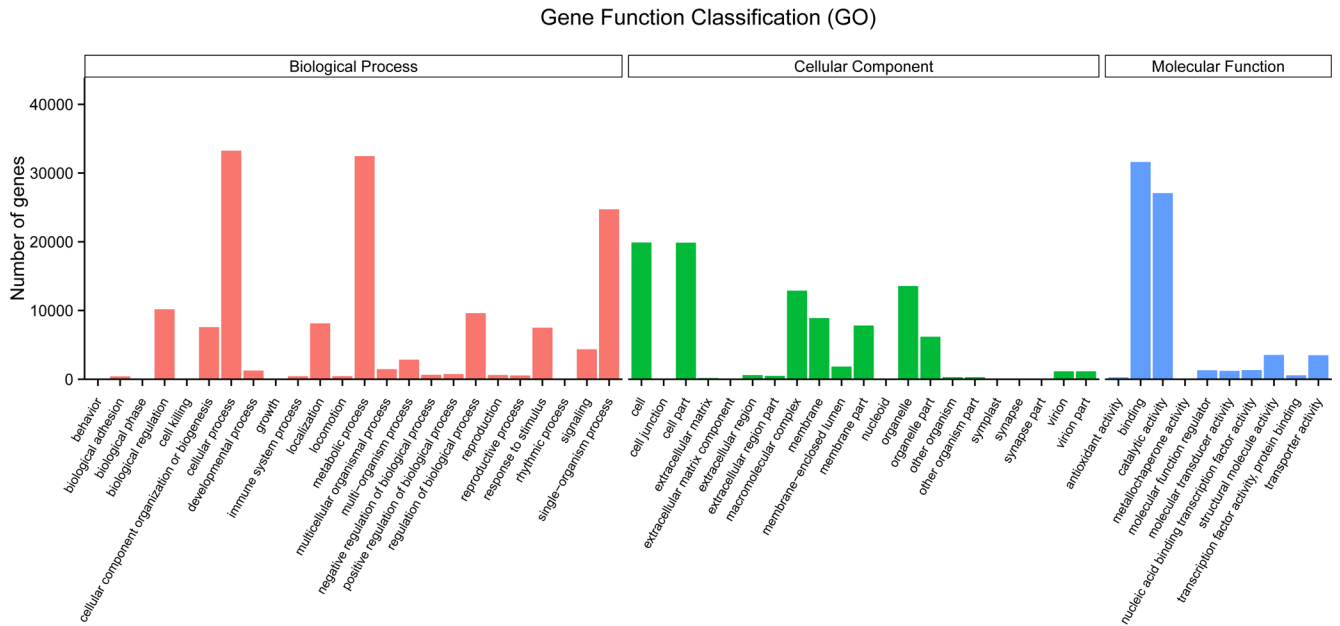


Fig 4. Gene function classification according to GO assignments for the predicted *S. arcuata* transcripts.

<https://doi.org/10.1371/journal.pone.0197729.g004>

component” group (Level 1) were divided into 21 subcategories, among which “cell” (Level 2) comprised 20.9% as the largest term. Based on the KOG annotation result (Fig 5), 40,556 genes belonging to 25 categories were yielded. Among these categories, the largest group was genes for “Posttranslational modification, protein turnover, chaperones” cluster, owning 6,407 (15.8% of the totally annotated transcripts) in number. The biological pathways in *S. arcuata* were identified according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Fig 6). A total of 31,330 transcripts were mapped to 19 KEGG pathways in 5 categories (S2 Table). Among the 5 categories, the pathways represented by most transcripts were metabolism (17,224, 54.98% of the totally annotated transcripts), followed by genetic information (10,893, 34.77% of the totally annotated transcripts) and cellular processes 2285 (7.29% of the totally annotated transcripts).

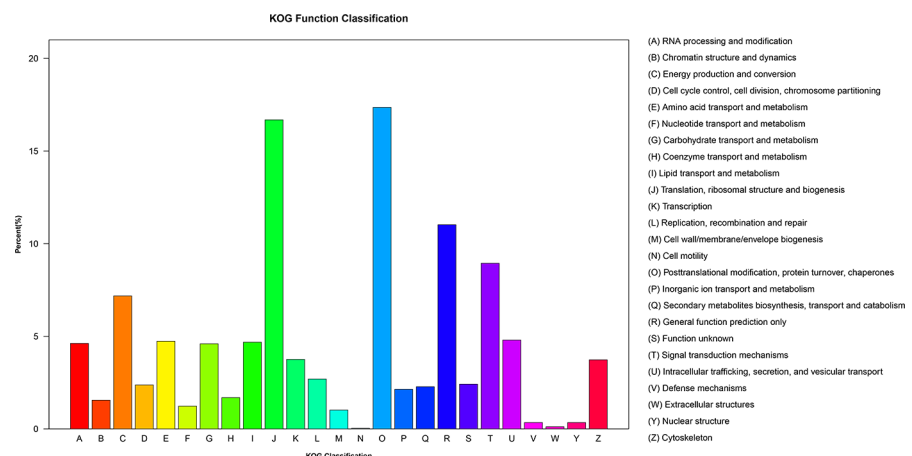
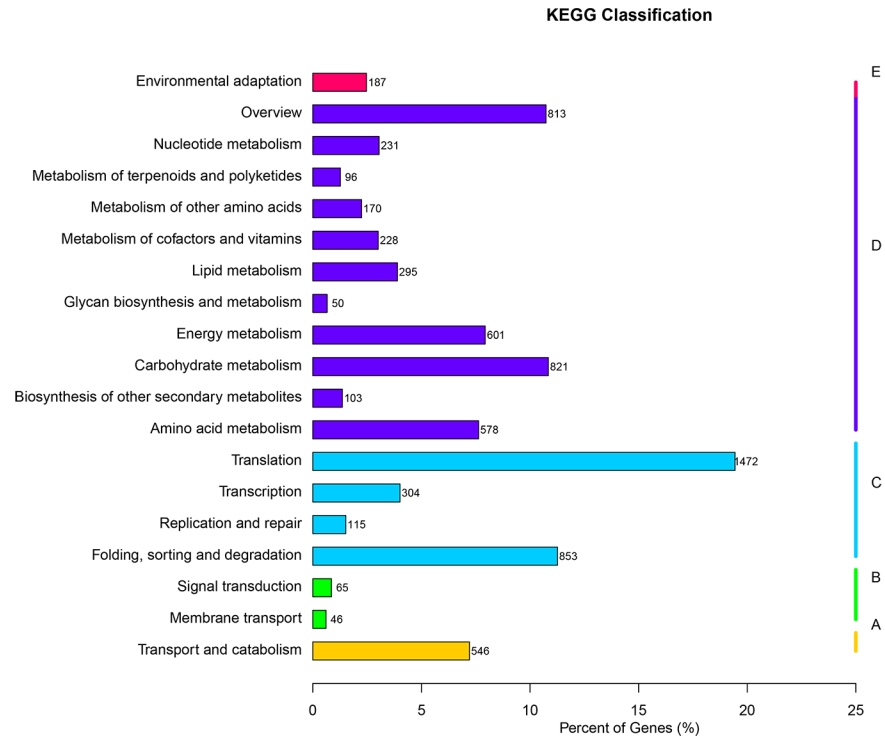


Fig 5. Gene function classification based on the KOG annotation for the predicted *S. arcuata* transcripts.

<https://doi.org/10.1371/journal.pone.0197729.g005>

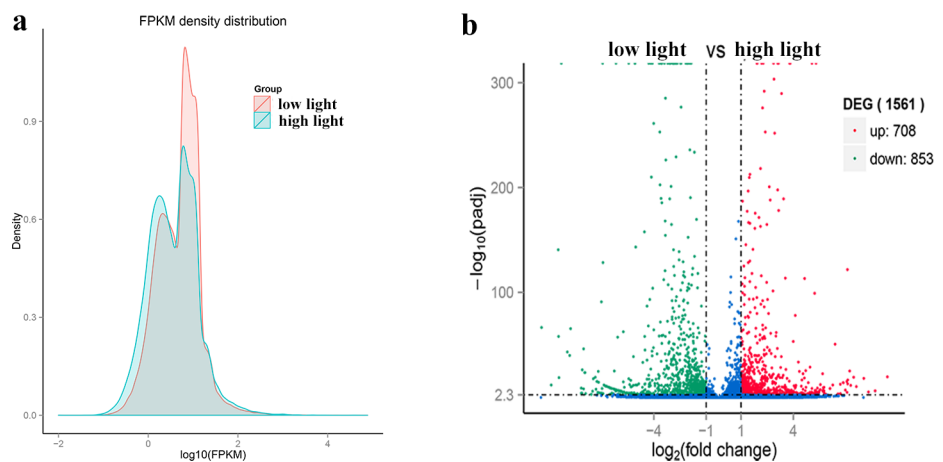


**Fig 6. The biological pathways in *S. arcuata* according to the Kyoto Encyclopedia of genes and genomes (KEGG) database.** A represents cellular processes; B represents environmental information processing; C represents genetic information; D represents metabolism; E represents organismal systems. Numbers on the right margin of each bar represents numbers of transcripts in the corresponding subcategories.

<https://doi.org/10.1371/journal.pone.0197729.g006>

### Differential gene expression analysis

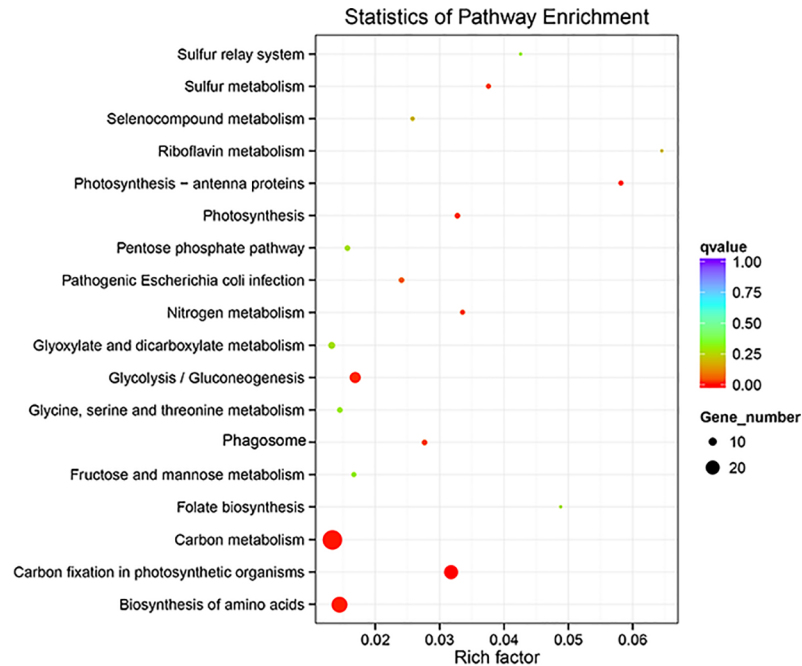
Gene expression levels of each sample were counted using Trinity, with different gene expression patterns observed in the two *S. arcuata* samples illustrated in Fig 7. Considering the similar water conditions of the sampling site when the specimens were collected in this study, the



**Fig 7. Different gene expression patterns for *S. arcuata* specimens collected at low and high light intensities.** a. FPKM density distribution for *S. arcuata* specimens collected at low and high light intensities; b. Volcanoplot showing the up and down regulated genes for *S. arcuata* specimens collected at low and high light intensities.

<https://doi.org/10.1371/journal.pone.0197729.g007>





**Fig 8. Enriched KEGG pathways for up-regulated genes in *S. arcuata* specimens collected under low light.**

<https://doi.org/10.1371/journal.pone.0197729.g008>

different gene expression was caused principally by variant irradiance. Genes with the same expression levels owned different densities in the two samples, revealing variance of gene expression responding to light intensity (Fig 7A). Differentially expressed genes with statistically significance were observed with up-regulated and down-regulated genes mainly in the sample under high and low irradiance respectively (Fig 7B). Gene lists of down-regulated and up-regulated were listed in S3 and S4 Tables.

The top 18 enriched KEGG pathways involving up-regulated genes under low irradiance were illustrated as Fig 8. The enriched pathways of down-regulated genes in *S. arcuata* specimen under low irradiance were not statistically significant, with all the q-values larger than 0.5. Therefore they are not discussed in this study (S5 Table). On the other hand, the enriched pathways corresponding up-regulated genes are all statistically significant, with all the q-values evidently smaller than 0.5 (Table 3). The top 18 significantly up-regulated genes under low irradiance were involved in important metabolism pathways including energy metabolism (photosynthesis, photosynthesis-antenna proteins, carbon fixation in photosynthetic organisms, sulfur metabolism, nitrogen metabolism), carbohydrate metabolism (glycolysis/gluconeogenesis, pentose phosphate pathway, glyoxylate and dicarboxylate metabolism, fructose and mannose metabolism), amino acid metabolism (glycine, serine and threonine metabolism), overview metabolism (carbon metabolism, biosynthesis of amino acids), metabolism of other amino acids (selenocompound metabolism), metabolism of cofactors and vitamins (riboflavin metabolism, folate biosynthesis), cellular processes (phagosome) and genetic information processing (sulfur relay system).

The significantly up-regulated transcripts involved in photosynthesis related pathways in *S. arcuata* specimen under low irradiance were showed in Figs 9–11 [30, 31]. Among photosynthesis—antenna proteins, significantly up-regulated transcripts were the light-harvesting chlorophyll complex LHCA1 and LHCA4, which were associated with the photosystem I (Fig 9). For photosynthesis apparatus, evidently up-expressed transcripts include PSBU in the

**Table 3. The top 18 enriched pathways in *S. arcuata* sample under low light.**

Pathway term	Rich factor	q-value	Gene number
Carbon fixation in photosynthetic organisms	0.031796502	4.37e-07	20
Carbon metabolism	0.01335175	0.00433149	29
Biosynthesis of amino acids	0.014420063	0.005744689	23
Photosynthesis—antenna proteins	0.058139535	0.008377125	5
Glycolysis/Gluconeogenesis	0.016872891	0.011253209	15
Photosynthesis	0.032786885	0.024956933	6
Sulfur metabolism	0.037593985	0.031901757	5
Phagosome	0.02764977	0.040114472	6
Nitrogen metabolism	0.033557047	0.040114472	5
Gap junction	0.024096386	0.066983265	6
Selenocompound metabolism	0.025806452	0.197482769	4
Riboflavin metabolism	0.064516129	0.200305359	2
Folate biosynthesis	0.048780488	0.27698983	2
Pentose phosphate pathway	0.015665796	0.27698983	6
Glyoxylate and dicarboxylate metabolism	0.013245033	0.27698983	8
Sulfur relay system	0.042553191	0.295376788	2
Fructose and mannose metabolism	0.016666667	0.295376788	5
Glycine, serine and threonine metabolism	0.014527845	0.29597947	6

<https://doi.org/10.1371/journal.pone.0197729.t003>

photosystem II, PETB and PETC in cytochrome b6/f complex, PETH in photosynthetic electron transport, and the beta and gamma subunits of F-type ATPase (Fig 10). Moreover, the ATPase, which constituted part of the photosystem, were also up-expressed as a result of increased light absorption and electron transport. As the last step of photosynthesis, pathway of carbon fixation in photosynthetic organisms was the most enriched (Fig 11). The up-expressed transcripts involved in carbon fixation were as followed, phosphoenolpyruvate carboxykinase (EC 4.1.1.49), aspartate transaminase (EC 2.6.1.1), phosphoribulokinase (EC 2.7.1.19), transketolase (EC 2.2.1.1), sedoheptulose-bisphosphatase (EC 3.1.3.37), fructose-bisphosphate aldolase (EC 4.1.2.13), fructose-bisphosphatase (EC 3.1.3.11), Phosphotriose isomerase (EC 5.3.1.1), glyceraldehyde-3-phosphate dehydrogenase (NADP+) (EC 1.2.1.13), glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) and phosphoglycerate kinase (EC 2.7.2.3).

Up-expression of focused genes in *S. arcuata* specimen under low light were validated by qRT-PCR, with the results showed in Fig 12. And the differential expression pattern revealed

**Allophycocyanin(AP)**

**Phycocyanin(PC)/Phycoerythrocyanin(PEC)**

**Phycoerythrin(PE)**

**Light-harvesting chlorophyll protein complex(LHC)**

ApcA	ApcB	ApcC	ApcD	ApcE	ApcF	
CpcA	CpcB	CpcC	CpcD	CpcE	CpcF	CpcG
CpeA	CpeB	CpeC	CpeD	CpeE	CpeR	CpeS
CpeT	CpeU	CpeY	CpeZ			
<b>Lhca1</b>	Lhca2	Lhca3	<b>Lhca4</b>	Lhca5		
Lhcb1	Lhcb2	Lhcb3	Lhcb4	Lhcb5	Lhcb6	Lhcb7

**Fig 9. Up-regulated photosynthesis–antenna transcripts in *S. arcuata* specimen under low light.** The transcripts in red were significantly up-regulated.

<https://doi.org/10.1371/journal.pone.0197729.g009>

<b>Photosystem II</b>	PsbA	PsbD	PsbC	PsbB	PsbE	PsbF		
	PsbL	PsbJ	PsbK	PsbM	PsbH	PsbI	PsbO	PsbP
	PsbQ	PsbR	PsbS	PsbT	<b>PsbU</b>	PsbV	PsbW	PsbX
	PsbY	PsbZ	Psb27	Psb28	Psb28-2			
<b>Photosystem I</b>	PsaA	PsaB	PsaC	PsaD	PsaE	PsaF	PsaG	PsaH
	PsaI	PsaJ	PsaK	PsaL	PsaM	PsaN	PsaO	PsaX
<b>Cytochrome b6/f complex</b>	<b>PetB</b>	PetD	PetA	<b>PetC</b>	PetL	PetM	PetN	PetG
<b>Photosynthetic electron transport</b>	PetE	PetF	<b>PetH</b>	PetJ				
<b>F-type ATPase</b>	<b>beta</b>	alpha	<b>gamma</b>	delta	epsilon	c	a	b

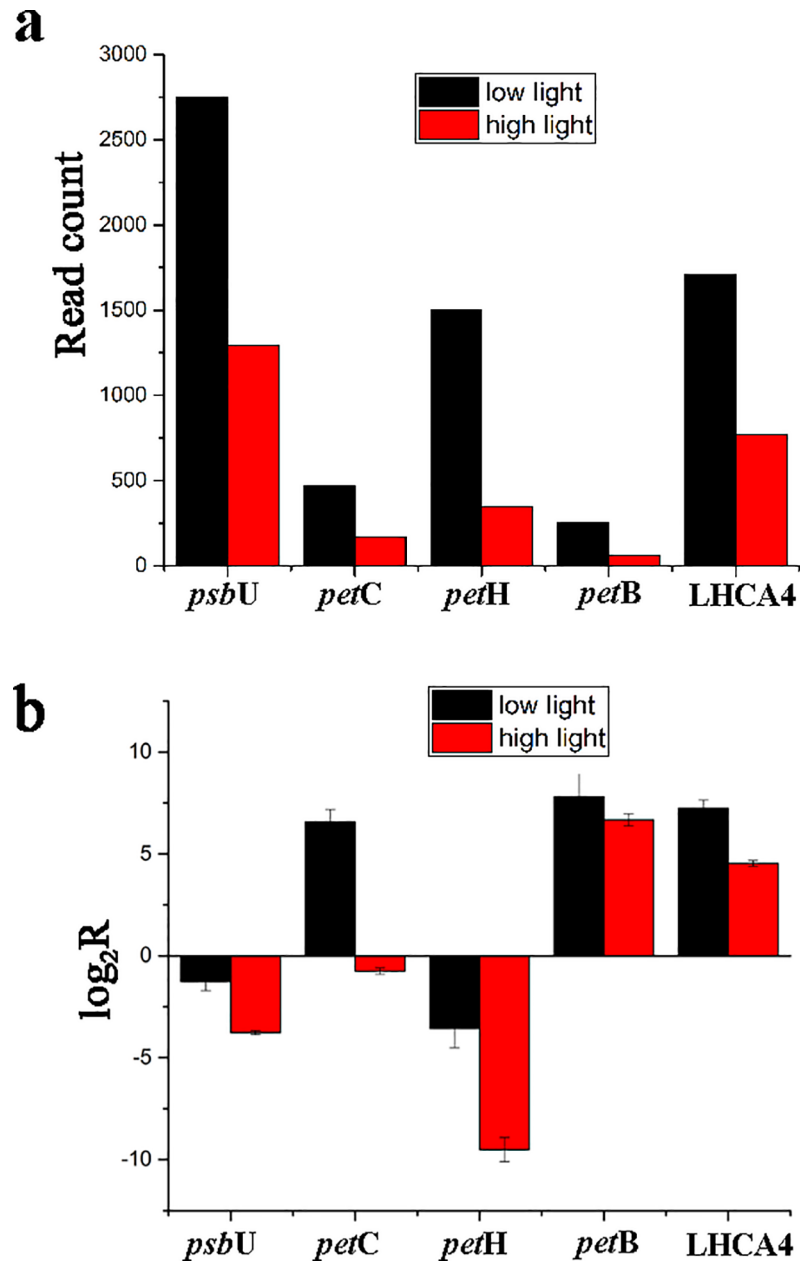
Fig 10. Up-regulated photosynthesis transcripts in *S. arcuata* specimen under low light. The transcripts in red were significantly up-regulated.

<https://doi.org/10.1371/journal.pone.0197729.g010>

<b>C4-Dicarboxylic acid cycle</b>	4.1.1.31	<b>2.6.1.1</b>	1.1.1.82	<b>4.1.1.49</b>	1.1.1.37	2.7.9.1
	2.6.1.2	1.1.1.39	1.1.1.40			
<b>Carbon fixation pathways in prokaryotes</b>	4.1.2.22	<b>4.1.2.13</b>	<b>2.2.1.1</b>	<b>3.1.3.11</b>	<b>3.1.3.37</b>	<b>5.3.1.1</b>
	2.7.1.14	<b>1.2.1.13</b>	1.2.1.59	<b>1.2.1.12</b>	5.3.1.6	<b>5.1.3.1</b>
	4.1.2.9	<b>2.7.1.19</b>	<b>2.7.2.3</b>	4.1.1.39		
<b>CAM</b>	4.1.1.31	2.7.9.1	1.1.1.37	1.1.1.40		

Fig 11. Up-regulated carbon fixation transcripts in *S. arcuata* specimen under low light. The transcripts in red were significantly up-regulated.

<https://doi.org/10.1371/journal.pone.0197729.g011>



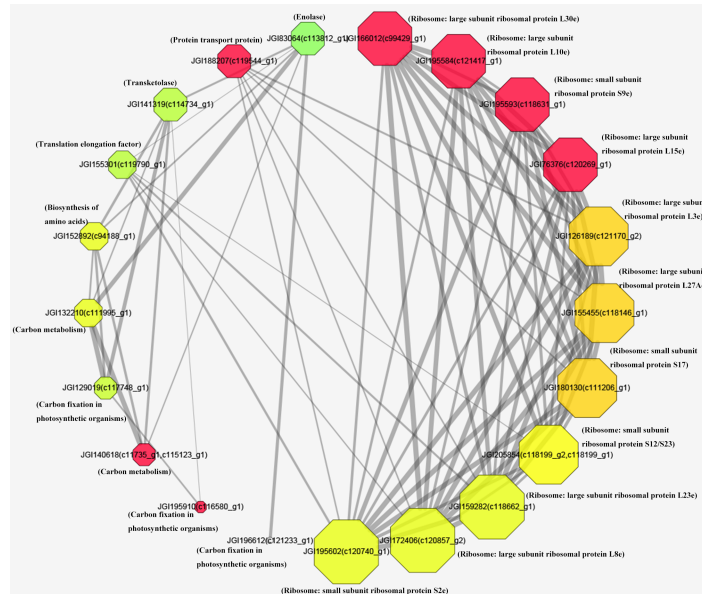
**Fig 12. The qRT-PCR validation of focused genes in specimens collected at high and low light.** a. Read counts of selected genes in each specimen based on high-throughput sequencing data; b. The qRT-PCR analysis of selected gene expression data.

<https://doi.org/10.1371/journal.pone.0197729.g012>

by qRT-PCR of selected genes was consistent with the high-throughput sequencing results, thus enhancing the statistical reliability based on sequencing data.

### Transcription factors

Transcription factors in *S. arcuata* (S6 Table) were identified and classified into different families using the iTAK pipeline (<http://bioinfo.bti.cornell.edu/tool/itak>) [34]. Results showed most abundant transcription factors involved in *S. arcuata* transcription process were



**Fig 13. Protein-protein interactive network showing the relationship among up-regulated genes for *S. arcuata* specimen under low light.** The size of each node represents interactive degree, with larger size corresponding to higher degree; the color of each node represents clustering coefficient, with colors ranging from green to red corresponding to lower to higher coefficients.

<https://doi.org/10.1371/journal.pone.0197729.g013>

regulatory genes including C2H2, C3H and orphan family. Both positive and negative transcriptional regulation of transcription factors by light has been documented. Different members in the families of common transcriptional factors and elongation factors were regulated diversely in the transcriptome profile of *S. arcuata* under weak irradiance (S3 and S4 Tables).

### Protein-protein interactions

Interactive networks involving the up-regulated transcripts of *S. arcuata* in response to weak light intensity were shown in Fig 13. The results revealed that the transcripts in response to light were all cross-linked and in a closely-related network. The nodes with highest degrees were transcripts corresponding ribosomal proteins, followed carbon metabolism, protein transport proteins, translation elongation factors, biosynthesis of amino acids and carbon fixation proteins in photosynthetic organisms.

### Discussion

Despite the widely application of transcriptome sequencing in marine Rhodophyta [28, 35–39], transcriptome data have not been reported for freshwater red algae. This study presents the first transcriptome profiling of *S. arcuata*, which will enrich the repertoire of transcripts of the freshwater rhodophytes and provide more data for the further investigation on this plant lineage. Compared with the marine red algal samples with transcriptomes reported to date (*P. yezoensis* with 18,640 annotated transcripts [28]). Assembled transcripts of freshwater taxa *S. arcuata* was considerably larger. The overrepresented transcripts length obtained in this study is consistent with sequencing result of *P. yezoensis*, with the most abundant length distributed between 200–500 bp, while the GC content of the transcriptome in *S. arcuata* is lower than *P. yezoensis* [28].

Specimens used in this study were collected at the irradiance of 1462 and 274  $\mu\text{mol photons/m}^2/\text{s}$  respectively, which were similar with light intensities used in previous research of genus *Sheathia* by Necchi. Necchi proved the maximal photosynthetic rate ( $8.1 \pm 0.5$ ) occurred in specimens collected at the irradiance of 320  $\mu\text{mol photons/m}^2/\text{s}$  and the minimal rate ( $4.9 \pm 0.6$ ) was under 1510  $\mu\text{mol photons/m}^2/\text{s}$  based on oxygen evolution test [10]. Photosynthetic changing trend revealed by physiological parameters measured in Necchi's research was consistent with the transcriptome regulation observed in our study.

Transcriptome analysis of *S. arcuata* grown under different light intensities in our study also shed light on the molecular mechanisms underlying the shade-adaptation of this taxon. For *S. arcuata* specimen under weak light intensity, the up-expressed photosynthesis-antenna transcripts (LHCA1 and LHCA4, as observed in this study) facilitated more light absorption and thus improving the photosynthetic activity. Light-harvesting complexes (LHCs) associated with both photosystems I and II (in green lineage) and phycobilisomes (in cyanobacteria) served as the primary light-harvesting antenna for photosynthesis [1, 40]. LHCs are important constituents that facilitate photosynthetic function in response to light quantity and quality [41]. Moreover it was found LHCs responded more evidently than the phycobilisome in *S. arcuata* when grown under low light intensity, revealing the improved adaptive ability to surrounding environment and the advanced stage of LHCs in Rhodophyta evolution. It was consistent with previous report that in red algae, the photosynthetic apparatus represented a transitional state between cyanobacteria and chloroplasts of green lineage, with enhanced light-harvesting capacity by owning LHCs (light harvesting complexes) associated with PSI [42]. Another pathway in regulation of adaptive response to weak light was photosynthesis. The up-expressed transcripts including PSBU, PETB, PETC, PETH, the beta and gamma subunits of F-type ATPase contribute to the adaptive response. It was reported that in cyanobacteria and red algae, the PS-II system gene *psbU* encodes protein constituting part of the oxygen-evolving complex (OEC), which was also involved in stabilizing the oxygen-evolving machinery of PSII against high-temperature stress [43]. PETB participated in electron transferring in the photosynthesis, and PETC in the cytochrome b6/f complex was involved in mediating electron transfer between photosystem II (PSII) and photosystem I (PSI) [44]. In combination with the photosynthetic electron transport protein PETH, the transcripts related to protein network involved in the electron transport of photosynthesis were all up-expressed in *S. arcuata* under weak irradiance. Photosynthetic control of electron transport was a fundamental concept in the regulation of photosynthesis [45]. Additionally, regulatory pathway of *S. arcuata* in respond to weak light was carbon fixation. Phosphoenolpyruvate carboxylase (EC 4.1.1.49) was up-regulated in response to weak light in freshwater *S. arcuata*, combined with other enzymes to improve the carbon fixation activity. In marine macroalgae, phosphoenolpyruvate carboxylase (EC 4.1.1.49) was characterized as the only enzyme for dark carbon fixation [46]. The up-regulated transcripts of photosynthesis-antenna and photosynthesis apparatus triggered the increasing rate of carbon assimilation, thus fueling the growth of *Sheathia* specimen under the low light intensity. In contrast with previous report on higher plant including *Bryophyllum fedtschenkoi*, maize and barley, expression of phosphoenolpyruvate carboxykinase (EC 4.1.1.49) was down-regulated in response to decreased light [47]. It is speculated that the up-regulation of this enzyme under weak light in freshwater Rhodophyta *S. arcuata* is relevant to their shade-adaptation. Our study revealed that for freshwater *S. arcuata*, transcripts involved in light harvesting, photosystem II, cytochrome b6 complex, photosynthetic electron transport and ATPase were all up-regulated thus enabling the increased photosynthetic function, which in turn provided sufficient energy and nutrients for fast growing of the plant under weak irradiance, which was consistent with the proposal that freshwater red algae were shade-adapted eukaryotic lineage [10].



Light provided fuel for photosynthetic electron transport and CO<sub>2</sub> fixation. As the primary determinant of ATP levels and carbon metabolites, it served to modulate cellular processes based on complex transcriptional networks [48]. Genus *Sheathia* regulated relative small amount of photosynthetic genes under different light intensities compared with marine diatom *Chaetoceros neogracile*, which exhibited altered expression of most photosynthesis genes (48 out of 70) in response to high light according to previous report [49]. It maybe explain the molecular mechanisms for weak ability in environmental adaption of *Sheathia*, leading to the current situation of strict habitat demand and limited distribution of freshwater Rhodophyta globally.

Along with the increased photosynthesis, other metabolic pathways with up-regulated transcripts were also observed including carbon metabolism, biosynthesis of amino acids, glycolysis / gluconeogenesis, sulfur metabolism, phagosome and nitrogen metabolism. *S. arcuata* displayed sophisticated responses to optimize their photosynthesis and growth under weak light conditions. This finding was in line with previous research on marine red algae and diatom, which indicated that light regulated many important cellular processes, physiological processes and biochemical pathways [37, 50–51]. Among the diverse responsive transcripts, those corresponding to ribosomal proteins and involved in protein synthesis, proved highly-regulated for *S. arcuata* under weak light. In previous research on marine Rhodophyta *Chondrus crispus*, stress treatments caused decreased expression of protein synthesis-related genes [38], which implied indirectly that low light intensity was more appropriate for growth and development of *S. arcuata*.

Both up- and down-regulation of diverse transcription factors in *S. arcuata* in responding to different irradiances revealed the complexity of regulation network. Transcription factors have been enriched in early light-responsive genes according to recent genomic studies [52]. The transcription factors identified in this study can direct adaptive changes in gene expression of freshwater Rhodophyta in response to environmental light signals in further study.

## Conclusions

We present the first transcriptome profiling of freshwater Rhodophyta by conducting high-throughput RNA sequencing on *S. arcuata*. A total of 161,483 assembled transcripts were identified and different gene expression patterns under different irradiances were observed. The results revealed that photosynthesis-related pathways significantly up-regulated under the weak light, revealing the shade-adaption of freshwater red algae *S. arcuata*. Molecular mechanisms underlying shade-adaption are increased expression of transcripts corresponding to antenna proteins (LHCA1 and LHCA4), photosynthetic apparatus proteins (PSBU, PETB, PETC, PETH and beta and gamma subunits of ATPase) and metabolic enzymes in the carbon fixation. The most responsive up-expressed transcripts were ribosomal proteins in *S. arcuata* grown under low light intensity. The *de-novo* transcriptome assembly of *S. arcuata* laid the foundation for further investigation on environmental adaption of freshwater Rhodophyta.

## Supporting information

**S1 Table. Specific primers of each gene used for qRT-PCR experiments.**

(DOCX)

**S2 Table. KEGG\_classification of the annotated transcripts in *Sheathia arcuata*.**

(XLSX)

**S3 Table. Down-regulated gene lists of *Sheathia arcuata* sample under low light intensity.**

(XLSX)



**S4 Table. Up-regulated gene lists of *Sheathia arcuata* sample under low light intensity.**  
(XLSX)

**S5 Table. The top 20 enriched pathways of down-regulated genes in *Sheathia arcuata* specimen under low irradiance.**  
(XLSX)

**S6 Table. Transcription factors identified in *Sheathia arcuata* and the related classification.**  
(XLSX)

## Acknowledgments

We are grateful to the National Natural Science Foundation of China (grant number 31670208 to Shulian Xie) and the Fund for Shanxi “1331 Project” Key Innovative Research Team for funding this project.

## Author Contributions

**Conceptualization:** Fangru Nan, Shulian Xie.

**Data curation:** Fangru Nan.

**Formal analysis:** Fangru Nan.

**Funding acquisition:** Shulian Xie.

**Methodology:** Jia Feng, Junping Lv, Qi Liu.

**Supervision:** Shulian Xie.

**Writing – original draft:** Fangru Nan.

**Writing – review & editing:** Shulian Xie.

## References

1. Moreira D, Le Guyader H, Philippe H. The origin of red algae and the evolution of chloroplasts. *Nature* 2000; 405:69–72. <https://doi.org/10.1038/35011054> PMID: 10811219
2. Gould SB. Evolutionary genomics: Algae's complex origins. *Nature* 2012; 492:46–48. <https://doi.org/10.1038/nature11759> PMID: 23201689
3. Sheath RG. The biology of freshwater red algae. *Prog Phycology Res.* 1984; 3:89–157.
4. Guiry MD, Guiry GM. AlgaeBase. World-wide Electronic Publication, National University of Ireland, Galway. <http://www.algaebase.org> (accessed 15.06.14).
5. Wehr JD, Sheath RG, Kociolek JP. *Freshwater algae of North America: ecology and classification.* Elsevier; 2015.
6. Whitford LA, Schumacher GJ. Communities of algae in North Carolina streams and their seasonal relations. *Hydrobiologia* 1963; 22:133–196.
7. Xie S. Seasonal dynamics of *Batrachospermum gelatinosum* growth and distribution in Niangziguan spring, China. *J. Appl. Ecol.* 2004; 15:1931–1934.
8. Vis ML, Feng J, Chiasson WB, Xie SL, Stancheva R, Entwisle TJ et al. Investigation of the molecular and morphological variability in *Batrachospermum arcuatum* (Batrachospermales, Rhodophyta) from geographically distant locations. *Phycologia* 2010; 49:545–553.
9. Anderson JM. Photoregulation of the composition, function, and structure of thylakoid membranes. *Annu Rev Plant Physiol.* 1986; 37:93–136.
10. Necchi O. Light-related photosynthetic characteristics of freshwater rhodophytes. *Aqua Bot.* 2005; 82:193–209.
11. Zucchi MR, Necchi O. Effects of temperature, irradiance and photoperiod on growth and pigment content in some freshwater red algae in culture. *Phycology Res.* 2001; 49: 103–114.

12. Gohin M, Bobe J, Chesnel F. Comparative transcriptomic analysis of follicle-enclosed oocyte maturational and developmental competence acquisition in two non-mammalian vertebrates. *BMC Genomics* 2010; 11:18. <https://doi.org/10.1186/1471-2164-11-18> PMID: 20059772
13. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* 2008; 18:1509–1517. <https://doi.org/10.1101/gr.079558.108> PMID: 18550803
14. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 2008; 5:621–628. <https://doi.org/10.1038/nmeth.1226> PMID: 18516045
15. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009; 10:57–63. <https://doi.org/10.1038/nrg2484> PMID: 19015660
16. Guo S, Zheng Y, Joung JG, Liu S, Zhang Z, Crasta OR, et al. Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics* 2010; 11:384. <https://doi.org/10.1186/1471-2164-11-384> PMID: 20565788
17. Feng C, Chen M, Xu C, Bai L, Yin XR, Li X, et al. Transcriptomic analysis of Chinese bayberry (*Myrica rubra*) fruit development and ripening using RNA-Seq. *BMC Genomics* 2012; 13:19. <https://doi.org/10.1186/1471-2164-13-19> PMID: 22244270
18. Holmes DS, Bonner J. Preparation, molecular weight, base composition, and secondary structure of giant nuclear ribonucleic acid. *Biochemistry* 1973; 12:2330–2338. PMID: 4710584
19. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011; 29:644–652. <https://doi.org/10.1038/nbt.1883> PMID: 21572440
20. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 2008; 36:3420–3435. <https://doi.org/10.1093/nar/gkn176> PMID: 18445632
21. Eddy SR. Accelerated profile HMM searches. *PLOS Comput Biol.* 2011; 7: e1002195. <https://doi.org/10.1371/journal.pcbi.1002195> PMID: 22039361
22. Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz HR, et al. The Pfam protein families database. *Nucleic Acids Res.* 2008; 36: D281–D288. <https://doi.org/10.1093/nar/gkm960> PMID: 18039703
23. Li B, Dewey C. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011; 12:323. <https://doi.org/10.1186/1471-2105-12-323> PMID: 21816040
24. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616> PMID: 19910308
25. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010; 11: R106. <https://doi.org/10.1186/gb-2010-11-10-r106> PMID: 20979621
26. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 2010; 26:136–138. <https://doi.org/10.1093/bioinformatics/btp612> PMID: 19855105
27. Storey JD. The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann Stat.* 2003; 31:2013–2035.
28. Sun P, Mao Y, Li G, Cao M, Kong F, Wang L et al. Comparative transcriptome profiling of *Pyropia yezoensis* (Ueda) MS Hwang & HG Choi in response to temperature stresses. *BMC genomics* 2015; 16: 463. <https://doi.org/10.1186/s12864-015-1586-1> PMID: 26081586
29. Pfaffl MW. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic acids Res.* 2001; 29: e45–e45. PMID: 11328886
30. Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 2005; 21:3787–3793. <https://doi.org/10.1093/bioinformatics/bti430> PMID: 15817693
31. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 2008; 36:D480–D484. <https://doi.org/10.1093/nar/gkm882> PMID: 18077471
32. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003; 13:2498–2504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
33. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000; 25:25–29. <https://doi.org/10.1038/75556> PMID: 10802651

34. Jin J, Zhang H, Kong L, Gao G, Luo J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 2014; 42:D1182–D1187. <https://doi.org/10.1093/nar/gkt1016> PMID: 24174544
35. Wu S, Sun J, Shan C, Wang L, Wang X, Liu C, et al. Transcriptome sequencing of essential marine brown and red algal species in China and its significance in algal biology and phylogeny. *Acta Oceanologica Sinica* 2014; 33: 1–12.
36. Xie C, Li B, Xu Y, Ji D, Chen C. Characterization of the global transcriptome for *Pyropia haitanensis* (Bangiales, Rhodophyta) and development of cSSR markers. *BMC Genomics* 2013; 14:107. <https://doi.org/10.1186/1471-2164-14-107> PMID: 23414227
37. Ho CL, Teoh S, Teo SS, Rahim RA, Phang SM. Profiling the transcriptome of *Gracilaria changii* (Rhodophyta) in response to light deprivation. *Mar Biotechnol.* 2009; 11:513–519. <https://doi.org/10.1007/s10126-008-9166-x> PMID: 19043658
38. Collén J, Guisle-Marsollier I, Léger JJ, Boyen C. Response of the transcriptome of the intertidal red seaweed *Chondrus crispus* to controlled and natural stresses. *New Phytologist* 2007; 176:45–55. <https://doi.org/10.1111/j.1469-8137.2007.02152.x> PMID: 17803640
39. Lim EL, Siow RS, Rahim RA, Ho CL. Global transcriptome analysis of *Gracilaria changii* (Rhodophyta) in response to agarolytic enzyme and bacterium. *Mar Biotechnol.* 2016; 18:189–200. <https://doi.org/10.1007/s10126-015-9680-6> PMID: 26631182
40. Grabowski B, Cunningham FX, Gantt E. Chlorophyll and carotenoid binding in a simple red algal light-harvesting complex crosses phylogenetic lines. *P Natl Acad Sci USA.* 2001; 98:2911–2916.
41. Staubar EJ, Fink A, Markert C, Kruse O, Johannngmeier U, Hippler M. Proteomics of *Chlamydomonas reinhardtii* light harvesting proteins. *Eukaryot Cell* 2003; 2:978–994. <https://doi.org/10.1128/EC.2.5.978-994.2003> PMID: 14555480
42. Wolfe GR, Cunningham FX, Durnford D, Green BR, Gantt E. Evidence for a common origin of chloroplasts with light-harvesting complexes of different pigmentation. *Nature* 1994; 367:566–568.
43. Enami I, Okumura A, Nagao R, Suzuki T, Iwai M, Shen JR. Structures and functions of the extrinsic proteins of photosystem II from different species. *Photosynth Res.* 2008; 98:349–363. <https://doi.org/10.1007/s11120-008-9343-9> PMID: 18716894
44. Kurisu G, Zhang H, Smith JL, Cramer WA. Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science* 2003; 302:1009–1014. <https://doi.org/10.1126/science.1090165> PMID: 14526088
45. Foyer CH, Furbank RT, Harbinson J, Horton P. The mechanisms contributing to photosynthetic control of electron transport by carbon assimilation in leaves. *Photosynth Res.* 1990; 25: 83–100. <https://doi.org/10.1007/BF00035457> PMID: 24420275
46. Akagawa H, Ikawa T, Nisizawa K. <sup>14</sup>CO<sub>2</sub> fixation in marine algae with special reference to the dark fixation in brown algae. *Botanica Marina.* 1972; 15: 126–132.
47. Hartwell J, Smith LH, Wilkins MB, Jenkins GI, Nimmo HG. Higher plant phosphoenolpyruvate carboxylase kinase is regulated at the level of translatable mRNA in response to light or a circadian rhythm. *The Plant Journal* 1996; 10: 1071–1078.
48. Grossman AR, Bhaya D, He Q. Tracking the light environment by cyanobacteria and the dynamic nature of light harvesting. *J Biol Chem.* 2001; 276:11449–11452. <https://doi.org/10.1074/jbc.R100003200> PMID: 11279225
49. Park S, Jung G, Hwang YS, Jin E. Dynamic response of the transcriptome of a psychrophilic diatom, *Chaetoceros neogracile*, to high irradiance. *Planta* 2010; 231: 349–360. <https://doi.org/10.1007/s00425-009-1044-x> PMID: 19924439
50. Bowler C, Falciatore A, Finazzi G. An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *P Natl Acad Sci USA.* 2010; 107:18214–18219.
51. Depauw FA, Rogato A, d'Alcalá MR, Falciatore A. Exploring the molecular basis of responses to light in marine diatoms. *J Exp Bot.* 2012; 63:1575–1591. <https://doi.org/10.1093/jxb/ers005> PMID: 22328904
52. Lu FH, Yoon MY, Cho YI, Chung JW, Kim KT, Cho MC, et al. Transcriptome analysis and SNP/SSR marker information of red pepper variety YCM334 and Taeon. *Scientia Horticulturae* 2011; 129:38–45.