© Springer 2005

Microarray profiling of plastid gene expression in a unicellular red alga, *Cyanidioschyzon merolae*

Ayumi Minoda[1,2], Kirina Nagasawa[1], Mitsumasa Hanaoka[1], Makoto Horiuchi[1], Hideo Takahashi[1] and Kan Tanaka[1,*]

[1]*Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, 113-0032, Tokyo, Japan (*author for correspondence; e-mail kntanaka@iam.u-tokyo.ac.jp);* [2]*Department of Plant Biology, Michigan State University, S-338 Plant Biology Building, East Lansing, MI, 48824-1312, USA*

Abstract

Plastid genomes of red algae contain more genes than those of green plant lineages, and it is of special interest that four transcription factors derived from ancestral cyanobacteria are encoded therein. However, little is known about transcriptional regulation of the red algal plastid genome. In this study, we constructed a red algal plastid DNA microarray of *Cyanidioschyzon merolae* covering almost all protein coding genes, and found that plastid genes are differentially activated by illumination. Run-on transcription assays using isolated plastids confirmed that activation takes place at the transcriptional level. In bacteria and plants, sigma factors determine the genes that are to be transcribed, and four plastid sigma factors (Cm_SIG1–4) encoded in the nuclear genome of *C. merolae* may be responsible for differential gene expression of the plastid genome. We found that transcripts for all Cm_*SIG* genes accumulated transiently after a shift from dark to light, whereas only the Cm_*SIG2* transcript was increased after a shift from low to high light, suggesting that Cm_SIG2 is a sigma factor that responds to high light. Phylogenetic analysis of plastid sigma factors suggested that sigma factors of red and green algal plastids and the group 1 sigma factors of cyanobacteria form a monophyletic group.

Abbreviations: ORF, Open reading frame; RT-PCR, Reverse transcriptase polymerase chain-reaction; Ycf, Hypothetical chloroplast open reading frame

Introduction

Plastids of red and green plants are derived from a single endosymbiosis of a cyanobacteria like eubacterium within a non-photosynthetic eukaryote (Moreira *et al.*, 2000). During subsequent evolution, plastids lost their cellular autonomy and became organelles whose metabolic roles are extensively controlled by the nucleus. The first determination of complete plastid genome sequences in liverwort and tobacco strongly supported the cyanobacterial origin of plastids owing to the retention of a number of genes for photosynthesis, transcription and translation, but showed the absence of plastid genes for transcriptional regulators including the sigma subunit of plastid-encoded RNA polymerase (Ohyama *et al.*, 1986; Shinozaki *et al.*, 1986).

Although eubacterial RNA polymerases require the sigma subunit for the recognition of promoters and the initiation of transcription, genes encoding sigma factors are absent from all

known plastid genomes. In general, bacterial cells have a principal sigma factor for housekeeping role and a variable number of alternative sigma factors in the response to stress. Owing to the different promoter recognition properties of these sigma factors, bacteria can choose the strategy to survive in the nature environment (Gruber and Gross, 2003). Several laboratories, including our own, in red algae and green plants, found that sigma factors for plastid RNA polymerase were encoded in the nuclear genome, synthesized in the cytoplasm and imported into plastids (Liu and Troxler, 1996; Tanaka et al., 1996; Isono et al., 1997; Tanaka et al., 1997; Fujiwara et al., 2000). Most transcriptional regulators, other than sigma factors, of the ancestral symbiont have been lost during the evolution of green plant evolution (The Arabidopsis Genome Initiative, 2000), and thus plastid transcriptional regulation in green plants appears to be dominated by nuclear-encoded sigma factors and additional factors acquired after the symbiotic event. Recent studies in *Arabidopsis thaliana* have shown that transcriptional regulation by plastid sigma factors is deeply involved in chloroplast development and in the acclimation of plants to differing environments in a similar way as bacterial systems. For example, SIG2 activates the transcription of a set of plastid tRNA genes, and enhances chloroplast translation and tetrapyrrole biosynthesis. On the other hand, SIG5 responds blue light and various stresses, and activates blue light responsive promoter (BLRP) of *psbD* to support repairing of the PSII reaction center (Kanamaru et al., 2001; Yao et al., 2003; Mochizuki et al., 2004; Nagashima et al., 2004a; Tsunoyama et al., 2004; Ishizaki et al., 2005).

The determination of the plastid genome structure of a red alga, *Porphyra purpurea* showed that its plastids evolved differently than those of green plants (Reith et al., 1995). The plastid of *P. purpurea* retains genes for regulatory factors, including a DNA helicase, a histidine protein kinase, and transcription factors, and for certain metabolic enzymes, which descended from an ancestral cyanobacterium and were lost from the genomes of green plastids genomes. In addition to the nuclear-encoded sigma factors, four transcription factors are encoded in red and related plastid genomes, and believed to play critical roles for the transcriptional regulation. However, only limited information is thus far available for the roles of these transcriptional regulators in plastids of red algae (Reith, 1995; Oikawa et al., 1998).

To elucidate the transcriptional regulatory mechanisms in red algae, we chose *Cyanidioschyzon merolae* as a model system. This primitive, unicellular red alga lives in sulfate-rich acidic hot springs (pH 1.5, 45 °C) (Toda et al., 1995), and recently, complete nucleotide sequences of its nuclear, mitochondrial, and plastid genomes were determined (Ohta et al., 1998; Ohta et al., 2003; Matsuzaki et al., 2004). This alga is advantageous for laboratory experimentation (Minoda et al., 2004). DNA microarray analyses have recently been used to investigate plastid gene expression in green algae and higher plants (Lilly et al., 2002; Nakamura et al., 2003; Nagashima et al., 2004). Therefore, we constructed a DNA microarray representing the plastid genome of *C. merolae* and analyzed transcriptional regulation after changes in illumination. We also conducted a phylogenetic analysis of the four nuclear-encoded sigma factors.

Materials and methods

Algal Culture

C. merolae 10D were grown in Allen's (1959) photoautotrophic medium under continuous white light (30 μE m^{-2} s^{-1}) at 42 °C. Trace elements in Allen's medium were replaced by Arnon's A6 solution (Minoda et al., 2004). Three hundred ml of cells were shaken in a 500-ml flask under air, and harvested at OD$_{750}$ = 0.4–0.6 for nucleic acid extraction. To examine the effect of high light, 50 ml of cells were grown to the OD$_{750}$ = 0.4–0.6 in the in MA medium (Minoda et al., 2004) in a glass vessel at 42 °C, bubbled with 5% CO_2, and the intensity of light was increased from 30 to 1300 μE m^{-2} s^{-1}.

Preparation and analysis of DNA microarrays

DNA and RNA extractions were performed as described previously (Nagashima et al., 2004b). We amplified each protein-coding region of the plastid genome by PCR using specific primers (Supplementary Table 1) and total DNA of *C. merolae* as a template. The specific PCR primers were designed according to the following rules:

1) The 5′-ends of 3′-specific primers were identical to the 3′-termini of each open reading frame (ORF).

2) The lengths of the PCR products were designed to be 1 kb long, but the full length ORFs were amplified when the ORF size was less than 1 kb.

PCR was performed with 35 cycles of 94 °C for 1 min, 52 °C for 1 min and 72 °C for 3 min using ExTaq (Takara, Shiga, Japan). In total, 194 PCR products corresponding to 193 ORFs of the plastid protein-coding genes and the phage lambda Q gene as a control were purified and spotted on a glass slide in three series (Nagashima et al., 2004b).

Each microarray experiment was performed twice, completely independently. For the preparation of fluorescent probes, 20 μg of each RNA sample was mixed with 35 ng of Q-RNA as an external control. This RNA solution was then added to a primer mixture containing 0.5 pmol of the 3′-specific primer for each gene. Fluorescence-labeling and purification were performed using an Atlas Glass Fluorescent Labeling Kit (CLON-TECH CA, USA) and Cy3 or Cy5 Mono-Reactive Dye (Amersham Pharmacia, NJ, USA) according to the instructions of the supplier. Finally, the purified labeled probe was precipitated with ethanol and dissolved in 5 μl of water.

For hybridization to an array, a 5 μl portion of each probe labeled with Cy3 and Cy5 was mixed with 35 μl of ULTRAhyb hybridization buffer (Ambion, TX, USA). The resulting mixture was then incubated at 95 °C for 5 min, cooled to 65 °C, and applied to the microarray. The microarray was covered with Spaced Cover Glass L (TaKaRa, Shiga, Japan) and incubated at 50 °C for 16 h. After incubation, the cover glass was removed from the microarray in $2 \times SSC$. The slide was then washed twice with $0.1 \times SSC$ containing 0.1% SDS for 5 min, and twice in $0.1 \times SSC$ for 5 min. Finally, the slide was dipped in water followed by 99.5% ethanol, and then dried by centrifugation at $2500 \times g$ for 2 min at room temperature. Microarrays were scanned with two wavelengths for Cy3 (560 nm) and Cy5 (675 nm) by a laser fluorescent scanner (GeneTAC LS IV, Genomic Solutions, MI, USA). Gene TAC Analyzer software version 3.0.1 (Genomic Solutions) was used for data analysis. Phage lambda Q-gene spots were used to normalize the two channels with respect to signal intensity. After normalization, the standard deviations (SDs) for each signal intensity ratio (Cy5/Cy3) were within 19% except for 2 genes (accB, 29% in Figure 1A and lpxA, 21% in Figure 1C), and each signal intensity ratio (Cy5/Cy3) and SD is presented in Supplementary Table 2–4. In addition, log ratios are presented as $\log_{10}$ (Cy5/Cy3).

Northern blot and RT-PCR

Northern blot analysis was carried out according to Nagashima et al. (2004). Cm_SIG1, Cm_SIG3 and Cm_SIG4 probes were prepared using the following primer sets: 5′-CTAGAGCTCCGTGAGATA-3′ and 5′-GAAGATGAATCTGCGGAA-3′, 5′-GC-TCAGGTCGCGCGCGGC-3′ and 5′-CTGCGC-TGGTACTTGCGC-3′, and 5′-CAACTTATCG AGCATGAGGAGG-3′ and 5′-TCATCCTCGG-TGGGTTCC–3′, respectively. The probes for the plastid genes were prepared with the same primer sets as for construction of the DNA microarray (Supplementary Table 1). RT-PCR analysis of Cm_SIG2 was performed using ReverTra Dash (TOYOBO, Osaka, Japan) with primers, 5′-CAC-ATACTCGCGAACGAT-3′ and 5′-CGAGCCAC GCAGGACTAC-3′. In each reverse transcription reaction, 0.5 μg of total RNA was used. PCR reactions were carried out in 30 cycles of 94 °C for 1 min, 54 °C for 1 min and 72 °C for 3 min.

Isolation of chloroplasts

Chloroplasts from C. merolae cells were isolated as described previously (Miyagishima et al., 1999) with a few modifications. Cells that were dark-adapted for 12 h (D12) or re-illuminated for either 1 (L1) or 6 (L6) h after dark-adaptation, were collected by centrifugation and suspended in isolation buffer (20 mM HEPES-KOH, pH 7.6, 5 mM $MgCl_2$, 5 mM KCl, 5 mM EDTA, and 1.2 mM spermidine) containing 180 mM sucrose. The cells were then lysed by one passage through a French pressure cell at 276 MPa, and 100 μg/ml DNase I (DN-25; Sigma) was added. The lysate was incubated on ice for 1 h, filtered by passage through two layers of 10-μm nylon mesh and two layers of Miracloth, and then layered on three-step gradients of 80, 60 and 40% Percoll (Amersham) in isolation buffer containing 300 mM sucrose. After centrifugation for 1 h at $28,000 \times g$ in a
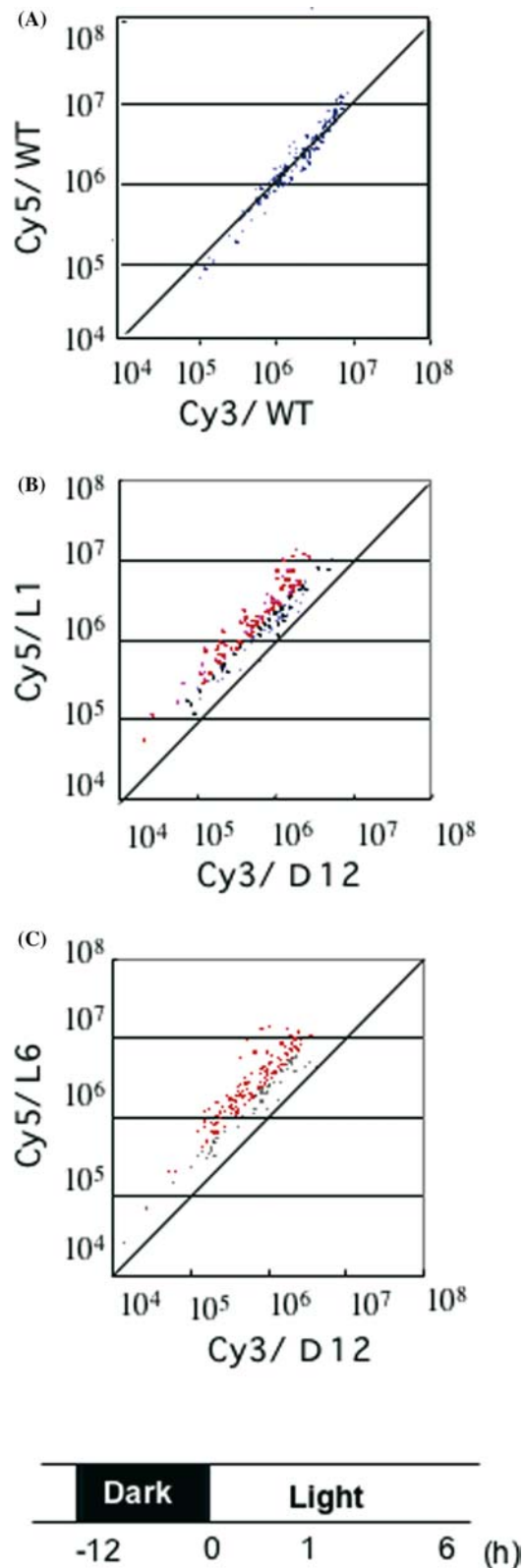
swinging-bucket rotor, a fraction of intact chloroplasts was harvested from the 60–80% Percoll interface. The chloroplasts were washed twice with isolation buffer containing 300 mM sucrose.

Chloroplast run-on transcription

The run-on transcription assay was modified from methods established previously (Orozco *et al.*, 1986; Deng and Gruissem, 1995). The transcription reaction was carried out in a 100 μl reaction mixture containing 5×10^7 chloroplasts, 22 mM HEPES-KOH, pH 7.9, 10 mM $MgCl_2$, 40 mM KCl, 2 mM dithiothreitol, 20 μg/ml heparin, 500 μM each of ATP, GTP, and CTP, and 50 μM UTP containing 100 μCi of [α-^{32}P] UTP (PB10163; Amersham, 400 Ci/mmol). The reaction was started by addition of the chloroplast sample and pipetting up and down 8 times to disrupt the membranes, and was incubated for 10 min at 25 °C. The reaction was then terminated by addition of 20 μl of stop solution (50 mM Tris–HC1, pH 8.0, 25 mM EDTA, 5% sodium sarcosinate), and extracted with phenol/chloroform/isoamylalcohol (25:24:1). To determine the incorporation of [α-^{32}P] UMP into RNA products, one-tenth volume of the phenol-extracted reaction mixture was spotted on DE-81 paper (Whatman), washed five times in 5% Na_2HPO_4 and once in H_2O, and radioactivity incorporated was measured by liquid scintillation (LS6500, Beckman Coulter, Fullerton, CA, USA).

Phylogenetic analysis

The deduced amino acid sequences were aligned using CLUSTAL X (Thompson *et al.*, 1997), and maximum parsimony trees then constructed using PAUP v4.0b10 (Swofford, 2002) (Figure 1, A and B). Maximum likelihood analyses were carried out using TREE-PUZZLE 4.0 based on the JTT model (Jones *et al.*, 1992) with 100 puzzling steps (Strimmer and von Haeseler, 1996). Bootstrap

Figure 1. Scatter plot of the values obtained using the plastid DNA microarray of *C. merolae*. Comparison of a cDNA preparation labeled with Cy3 and Cy5 after growth under the culturing conditions (A), and comparison of cDNA preparations from dark acclimated cells (D12) and from cells illuminated for 1 h (L1) (B) and 6 h (L6) (C). (D) Experimental scheme for the microarray analyses by light shift.

resampling was used to estimate the reliability of the inferred tree.

Results

Construction of the DNA microarray for plastid protein genes

The plastid genome of *C. merolae* contains 243 genes, which consist of 36 RNA genes (3 rRNAs, 31 tRNAs, 1 tmRNA, and 1 ribonuclease P RNA component) and 207 protein genes including hypothetical ORFs (Ohta *et al.*, 2003). To investigate the plastid gene transcription in *C. merolae*, we constructed a DNA microarray using the plastid DNA sequence. At the time of microarray design, 214 protein genes had been mapped on the plastid genome. Based on the information available, we amplified 193 protein-coding regions by PCR, and used the PCR products to construct for micoroarrays. ORFs of less than 50 amino acids that showed no similarity to any other proteins in available databases were eliminated from construction. The *orf340*, *psbX*, *rnpB*, *rpoZ*, *ycf12*, *ycf20* and *ycf33* genes were not included in the microarray, because these genes were identified after construction of the arrays when the sequence-annotation was revised (Ohta *et al.*, 2003). To evaluate quality, the microarray was co-hybridized with Cy3- and Cy5-labeled cDNAs derived from the same RNA preparation from the cells at OD_{750} 0.4–0.6. Almost every hybridization signal was plotted on or near a diagonal line (the diagonal line indicate the signal value ratio was 1.00; Figure 1A and Supplementary Table 2), indicating that the results of the arrays were reliable. Since every fold-change of spots was between 2.5 and 0.4, signals beyond or below these values were taken as significant changes.

Effect of illumination on the expression of plastid genes

We analyzed the effect of illumination on plastid gene expression using these microarrays. Cells at OD_{750} 0.4–0.6 grown as described above were kept in the dark for 12 h and then returned to the light. RNAs were prepared from cells harvested just before the light shift (D12), and after 1 (L1) or 6 h (L6) of illumination (Figure 1D), and the abundance of plastid transcripts was compared between D12 and L1, and D12 and L6. We found that transcripts of about 42 (82/193) and 49% (95/193) of the plastid genes increased above the cut-off value, 2.5, within 1 and 6 h, respectively. Although all plastid transcripts accumulated after illumination, at least two patterns of accumulation were observed (Table 1). Pattern I transcripts accumulated at L1, but then decreased at L6. This class included transcripts for *ftsH, petB, petD, psbC, psbD, psbY, rbcL, rbcS, ycf16*, and *ycf24*. Pattern II transcripts accumulated at L1 and then further accumulated at L6. This class included transcripts for genes such as *apcA, apcB, cfxQ, cpcA, cpcB, cpcG, infB*, and *psbV*.

17.6% and 82.4% of the genes were categorized to patterns I and II, respectively. Thus, plastid transcripts in *C. merolae* appear to be differentially regulated by light regimes, suggesting complex transcriptional regulation. These patterns were confirmed by Northern blot analyses for *rbcL, ftsH, cpcA*, and *cpcG* (Figure 2), which again indicated the reliability of the microarray results.

To examine whether the increase of the transcripts resulted from transcriptional activation, plastids were isolated from *C. merolae* cells at the same time points as the array analyses, and run-on transcription assays were performed. As shown in Table 2, incorporation of the labeled nucleotide increased more than 5-fold and 2.5-fold at L1 and L6, respectively. These results clearly indicated that transcriptional activation was principally responsible for the increase of the plastid transcripts.

Expression analysis of the nuclear-encoded plastid sigma factors of *C. merolae*

To understand the roles of each sigma factor, we analyzed the expression of the four sigma factors after changes in illumination. As in the microarray analysis above, cells were illuminated after dark adaptation, and expression of the sigma factors was monitored periodically by Northern hybridization (Figure 3). The expression of Cm_*SIG2* was monitored by RT-PCR because of the low abundance of the transcripts. Transcripts for all sigma factors accumulated rapidly by 1 h after illumination, and decreased by 3 h. Peak levels of the transcripts appeared to differ among sigma factors; Cm_*SIG1* was the most abundant, and Cm_*SIG2* the least abundant of the four

Table 1. List of genes classified according to transcript accumulation patterns I and II, whose differences of log ratios between L1 and L6 were over 0.2.

	Log ratio (L1)	Log ratio (L6)	Log ratio (L6–L1)
Pattern I genes[a]			
ClpC	0.81	0.33	−0.48
ycf24	0.82	0.39	−0.43
ycf17	0.66	0.26	−0.40
ycf16	0.88	0.52	−0.36
RbcL	0.62	0.27	−0.35
PsbY	0.28	−0.06	−0.33
PsbD	0.80	0.59	−0.21
PetL	0.50	0.30	−0.20
Pattern II genes[b]			
rpl6	0.01	0.28	0.27
rpl24	0.02	0.29	0.27
AtpF	0.20	0.45	0.25
rps19	0.11	0.35	0.24
rps17	0.12	0.36	0.24
rpl13	0.20	0.43	0.23
rpl5	0.02	0.26	0.24
rpl23	0.11	0.35	0.24
rps11	0.12	0.34	0.22
FtrC	0.19	0.41	0.21
AtpG	0.24	0.45	0.21
rpl14	0.13	0.33	0.20
ycf38	0.10	0.30	0.20
rps7	0.27	0.47	0.20

[a] Pattern I genes: their transcripts increased at L1, but then decreased at L6.
[b] Pattern II genes: their transcripts increased at L1 then increased further at L6.

(Figure 3). Subsequently, we investigated expression of the sigma factors in response to high light. Only the transcript of Cm_SIG2 showed a clear increase at 0.5 h after the high light (Figure 4). Thus, Cm_SIG2 appears to be high-light responsive sigma factor.

Phylogenetic analysis of plastid sigma factors

We characterized the four nuclear-encoded sigma factors phylogenetically. Based on the result by maximum likelihood methods (Figure 5A), Cm_SIG1 and Cm_SIG2 are apparently orthologs of SigA and SigB/C of *Cyanidium caldarium* RK-1, respectively (Tanaka *et al.*, 1996; Oikawa *et al.*, 1998). Cm_SIG4 belongs to a newly identified class of red algal sigma factors. Interestingly, Cm_SIG3 showed a close relationship with the RpoD protein encoded by the nucleomorph genome of the cryptophyte *Guillardia theta* (Douglas *et al.*, 2001), supporting the hypothesis that the plastid in that cryptophyte is derived from secondary symbiosis of a red alga (Valentin and

Zetsche, 1990; Douglas *et al.*, 2001) (Figure 5A, C). Furthermore, the algal plastid sigma factors grouped with the cyanobacterial group 1 sigma factors, and only more distally from cyanobacterial group 2 sigma factors (Figure 5B). These results suggest that plastid sigma factors originated from the cyanobacterial group 1 sigma factors. The inclusion of higher plant sigma factors in the phylogenetic analysis made it impossible to solve the tree by maximum likelihood methods due to extensive divergence of these plastid sigma factors during evolution.

Discussion

Light-dependent activation of plastid transcription in C. merolae

In this study we constructed plastid DNA microarrays for *C. merolae*, and analyzed the accumulation of transcripts after changes in illumination. With the use of these microarrays, we successfully
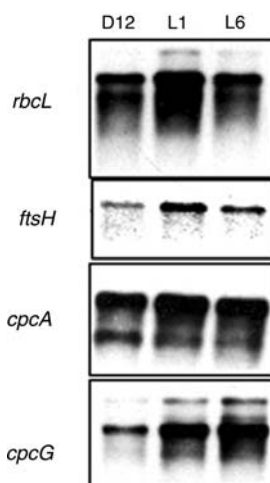
Figure 2. Analysis of changes in gene expression upon transfers of plastids from dark to light. *rbcL*, *ftsH*, *cpcA* and *cpcG* transcripts were monitored by Northern hybridization at D12, L1 and L6. Quantities of 1 μg of RNA for *rbcL* and *cpcA*, and 3 μg for *cpcG* and *ftsH*, respectively, were loaded for each lane. The probes used for Northern blot analysis are indicated on the left.
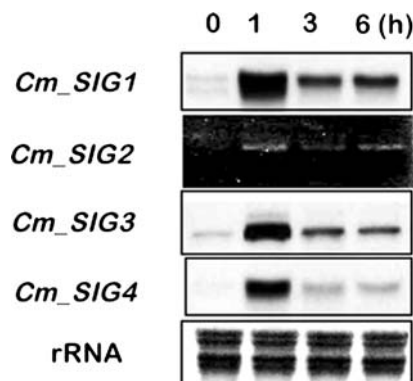

Figure 3. Effects of illumination on the accumulation of sigma factor transcripts in *C. merolae*. Cells acclimated to darkness for 12 h were illuminated (30 μE m^{-2} s^{-1}) for 1, 3 or 6 h. The accumulation of transcripts of Cm_SIG1, 3 and 4 and rRNA were examined by Northern blot analysis. Five micrograms of total RNA was electrophoresed for each lane. The transcript level of Cm_SIG2 was estimated by RT-PCR as described in Materials and methods. Electrophoretic patterns of Northern-blotted rRNA are shown in the lowest panel. Panel-specific probes are indicated on the left.
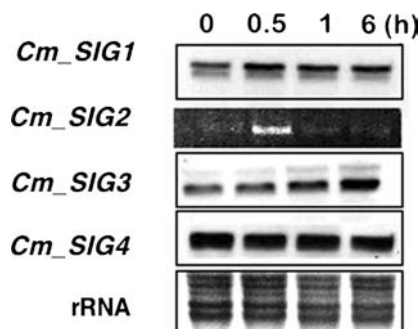
Table 2. Transcriptional activity in isolated *C. merolae* chloroplasts.

Condition	Transcriptional Activity (cpm/5×10^2 chloroplasts)	$\pm$SD
D12	1.031	0.048
L1	5.529	0.094
L6	2.638	0.091

The incorporation of [α-^{32}P] UMP into DEAE-binding material was determined by liquid scintillation counting. Values are the average of four replicate assays.

Figure 4. Accumulation of plastid sigma factor transcripts in *C. merolae*. *C. merolae* cells were grown to OD$_{750}$ = 0.4–0.6 with white light of 30 μE m^{-2} s^{-1} and the light intensity then increased to 1300 μE m^{-2} s^{-1}. RNAs were prepared 0.5, 1 or 6 h after the light shift, and *SIG* and rRNA gene transcripts were analyzed as in Figure 3. Panel-specific probes are indicated on the left.

identified at least two patterns of accumulation of transcripts after a dark-to-light transition (Figure 1, Table 1). The reliability of the microarray results was supported by Northern hybridization (Figure 2). Thus, as in other studies (Lilly *et al.*, 2002; Nakamura *et al.*, 2003; Nagashima *et al.*, 2004b), these microarrays were useful in analyzing plastid gene expression, and will facilitate future analysis of the plastid transcription system in *C. merolae*.

In order to detect transcriptional activity *in vitro*, we also established a run-on transcription system using intact chloroplasts isolated from *C. merolae* cells, and showed that light-dependent accumulation of many plastid transcripts observed

in microarray analyses was due to activation of transcription, rather than a light-dependent increase of RNA stability (Table 2). While overall transcription has a tendency to be activated by the dark-light shift, we found at least two distinct gene expression patterns, suggesting differential transcriptional regulation of plastid genes. One of the patterns of transcript accumulation showed a single peak at L1, but then decreased at L6, which
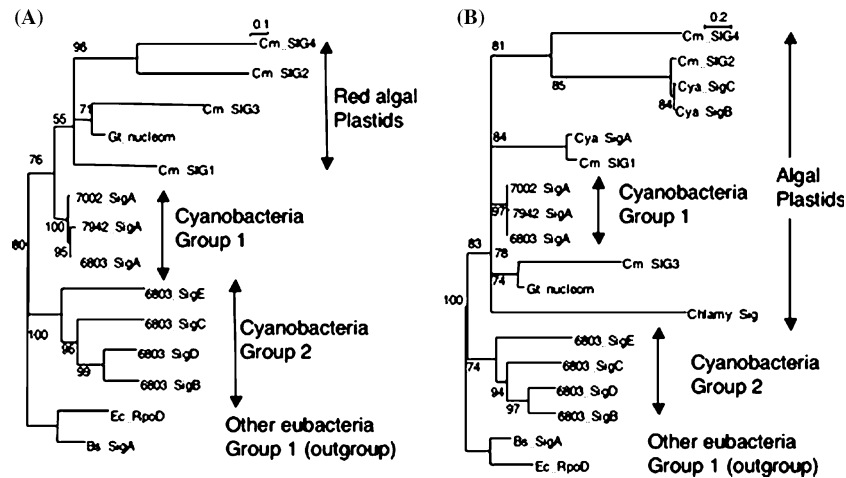
Figure 5. Phylogenetic tree of plastid and cyanobacterial sigma factors. Trees A and B were constructed using the maximum likelihood method, and tree C was constructed by the neighbor-joining method. 255 positions of 11 proteins (A) and 18 proteins (B) were considered using PAUP 4.0b10 (Swofford, 2002) and calculated using TREE-PUZZLE 4.0 based on the JTT model (Jones et al., 1992) with 100 puzzling steps (Strimmer and von Haeseler, 1996). The proteins analyzed were: Bs_SigA (the group 1 sigma factor of *Bacillus subtilis*, X03897), Cm_SIG1 (the *SIG1* gene product of *C. merolae*, CMK044C), Cm_SIG2 (the *SIG2* gene product of *C. merolae*, CMQ213C), Cm_SIG3 (the *SIG3* gene product of *C. merolae*, CMR165C), Cm_SIG4 (the *SIG4* gene product of *C. merolae*, CMM072C), Cyanidium_SigA (the *sigA* gene product of *C. caldarium* RK-1, D83179), Cyanidium_SigB (the *sigB* gene product of *C. caldarium* RK-1, AB006798), Cyanidium_SigC (the *sigC* gene product of *C. caldarium* RK-1, AB006799), Ec_RpoD (the group 1 sigma factor of *Escherichia coli*, J01687), the Guillardia_nucleomorph (the *rpoD* gene product of *Guillaridia theta* nucleomorph, NC_002752), 6803_SigA (the group 1 sigma factor SigA of *Synechocystis* sp. strain PCC 6803, Slr0653), 6803_SigB (the group 2 sigma factor SigB of *S.* sp. PCC 6803, Sll0306), 6803_SigC (the group 2 sigma factor SigC of *S.* sp. PCC 6803, Sll0184), 6803_SigD (the group 2 sigma factor SigD of *S.* sp. PCC 6803, Sll2012), 6803_SigE (the group 2 sigma factor SigE of *S.* sp. PCC 6803, Sll1689), 7942_SigA (the group 1 sigma factor, SigA, of *Synechococcus* sp. PCC 7942, S24172), 7002_SigA (the group 1sigma factor, SigA, of *Synechococcus* sp. PCC 7002, AAB41506) and Chlamy_Sig (the *sig1* gene product, CAF25319).

resembled that of circadian oscillation of the plastid-encoded *psbD* gene transcription in wheat (Nakahira et al., 1998). Although it is presently difficult to distinguish effect between light activation and circadian clock, this observation suggests that expression of some of plastid genes in *C. merolae* could potentially be regulated by circadian clock.

Evolution of the regulatory mechanism for the plastid genes

In plastids of the green plant lineage, such as *Chlamydomonas* and tobacco, accumulation of plastid transcripts is largely controlled at the post-transcriptional level, although some genes such as *psbD* in higher plants are clearly regulated at the transcriptional level (Sexton et al., 1990; Barkan and Goldschmidt-Clermont, 2000; Cahoon et al., 2004). Two kinds of data suggested that, in contrast, transcriptional regulation might

play a greater role in red algal plastids. First, four transcription factors, Ycf27–30, are encoded by the plastid genome. These transcription factors have homologs in ancestral cyanobacteria, and therefore, were presumably retained after endosymbiosis with the cyanobacterium from which they originated. We suggest that these proteins were preserved to maintain rapid transcriptional responses in plastids. Second, red algal plastids encode more proteins than do green plastids, and some of these extra genes were shown to be regulated at the transcriptional level. For example, the plastid *groEL* gene encoding chaperonin 60 is transcriptionally activated during acclimation to high temperature in *Porphyra umbilicalis* and *Cyanidium caldarium* (Reith and Munholland, 1991; Maid et al., 1992). As described above, we found that transcripts of plastid genes of *C. merolae* including *ycf24*, *ycf16*, and *ftsH* that are not present in the plastid genome of higher plants rapidly accumulated with illumination (Figure 1, 2, Table 1), and this accumulation could be

explained by transcriptional activation (Table 2). During the evolution of green plastids, transcriptional regulation, which plays major role in cyanobacteria, appears to have been largely replaced by post-transcriptional regulation. Thus, as to the regulation system of plastid genes, red algal plastids might represent an intermediary stage of evolution between cyanobacteria and plant plastids.

Origin and evolution of plastid sigma factors

As shown in this study, algal sigma factors belong to the same clade as cyanobacterial group 1 sigma factors (Figure 5A, B). Taking the evolutionary relationship between green algae and higher plants into account, we consider it highly likely that plastid sigma factors are also positioned in the same cluster with the algal and the cyanobacterial group 1 sigma factors. Therefore, we suggested that all of plastid sigma factors are monophyletic and originated from the group 1 sigma factors of the primary symbiont. We propose that the principal sigma factors of the cyanobacterial symbiont were transferred from the genome of the ancestral plastid to the host nuclear genome after establishment of symbiosis, and then diversified during subsequent evolution. If there

were other sigma factors of cyanobacterial origin, they appear to have been lost during evolution.

Miniature eubacterial transcription system

Based on the complete genome analysis of *C. merolae*, the plastid transcription system is composed of plastid-encoded subunits of core RNA polymerase (RpoA, RpoB, RpoC, and RpoZ), four nuclear-encoded sigma factors (SIG1–4) and four plastid-encoded transcription factors (Ycf27–30). Among higher plants, another RNA polymerase called NEP is also involved in plastid transcription. However, this type of RNA polymerase is unlikely to be involved in plastid transcription in *C. merolae*, because the homologous gene found in the nuclear genome is a single copy gene, and the gene product was predicted to localize in mitochondria (data not shown). Of the four plastid-encoded transcription factors, Ycf27 and Ycf29 are two-component response regulators that probably require phosphorylation for DNA binding. Only one corresponding candidate gene for histidine protein kinase was found in the nuclear genome, and it is highly likely that this gene product is localized in the plastid to control Ycf27 and/or Ycf29-dependent transcription. Finally, as a potential transcription regulator, a homolog of
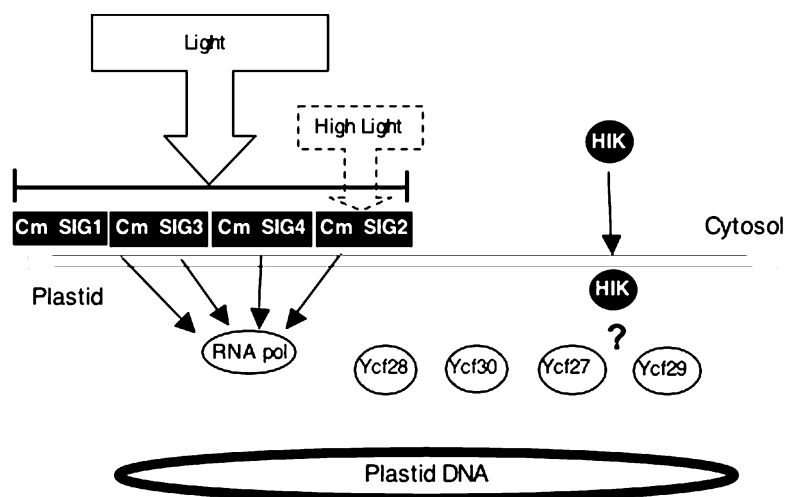

Figure 6. Transcriptional regulatory model of the plastid genome in *C. merolae*. In response to light, four sigma factors appear to activate transcription of plastid genes. Cm_SIG2 appears to respond to an increase in illumination from 30 to 1300 μE m^{-2} s^{-1}. A single presumptive sensory histidine kinase may be transferred from the cytosol to the plastid and together with one or both of two response regulators encoded in the plastid genome, which may constitute a two-component system. Proteins Ycf and RNA polymerase are encoded in the plastid genome, whereas proteins Cm_SIGs and HIK are encoded in the nuclear genome.

the bacterial nucleoid protein, HU, is encoded in the plastid genome (Kobayashi *et al.*, 2002).

As summarized here, the plastid transcription system in *C. merolae* appears to be a miniature eubacterial transcription system, and an excellent model system for understanding the evolution of plastid (Figure 6). The functions of each sigma and each transcription factor remain unknown; however, the differential expression patterns among the Cm_*SIG* genes suggest that the roles of sigma factors have diverged. Moreover, comparative functional analysis of the plastid transcription factors with cyanobacterial homologs should show unique features of plastid transcription system in red algae.

Acknowledgments

References

Allen, M.B. 1959. Studies with *Cyanidium caldarium*, an anomalously pigmented chlorophyte. Arch. Microbiol. 32: 270–277.

Barkan, A. and Goldschmidt-Clermont, M. 2000. Participation of nuclear genes in chloroplast gene expression. Biochimie 82: 559–572.

Cahoon, A.B., Harris, F.M. and Stern, D.B. 2004. Analysis of developing maize plastids reveals two mRNA stability classes correlating with RNA polymerase type. EMBO Rep. 5: 801–806.

Deng, X.W. and Gruissem, W. 1995. Chloroplast run-on transcription: Determination of the transcriptional activity of chloroplast genes. In: P. Maliga, D.F. Klessig, A.R. Cashmore, W. Gruissem and J.E. Varner (Eds.), Methods in Plant Molecular Biology, Cold Spring Harbor Laboratory Press, NY, pp. 191–207.

Douglas, S., Zauner, S., Fraunholz, M., Beaton, M., Penny, S., Deng, L.T., Wu, X., Reith, M., Cavalier-Smith, T. and Maier, U.G. 2001. The highly reduced genome of an enslaved algal nucleus. Nature 410: 1091–1096.

Fujiwara, M., Nagashima, A., Kanamaru, K., Tanaka, K. and Takahashi, H. 2000. Three new nuclear genes, *sigD*, *sigE* and *sigF*, encoding putative plastid RNA polymerase sigma factors in *Arabidopsis thaliana*. FEBS Lett. 481: 47–52.

Gruber, T.M. and Gross, C.A. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu. Rev. Microbiol. 57: 441–66.

Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeba, G., Nakahira, Y. and Shiina, T. 2005. A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. Plant J. 42: 133–144.

Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A. and Kobayashi, H. 1997. Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of σ^{70} factors of bacterial RNA polymerase in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 94: 14948–14953.

Jones, D.T., Taylor, W.R. and Thornton, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8: 275–282.

Kanamaru, K., Nagashima, A., Fujiwara, M., Shimada, H., Shirano, Y., Nakabayashi, K., Shibata, D., Tanaka, K. and Takahashi, H. 2001. *Arabidopsis* sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. Plant Cell Physiol. 42: 1034–1043.

Kobayashi, T., Takahara, M., Miyagishima, S., Kuroiwa, H., Sasaki, N., Ohta, N., Matsuzaki, M. and Kuroiwa, T. 2002. Detection and localization of a chloroplast-encoded HU-like protein that organizes chloroplast nucleoids. Plant Cell 14: 1579–1589.

Lilly, J.W., Maul, J.E. and Stern, D.B. 2002. The *Chlamydomonas reinhardtii* organellar genomes respond transcriptionally and post-transcriptionally to abiotic stimuli. Plant Cell 14: 2681–2706.

Liu, B. and Troxler, R.F. 1996. Molecular characterization of a positively photoregulated nuclear gene for a chloroplast RNA polymerase σ factor in *Cyanidium caldarium*. Proc. Natl. Acad. Sci. USA 93: 3313–3318.

Maid, U., Steinmuller, R. and Zetsche, K. 1992. Structure and expression of a plastid-encoded groEL homolog heat-shock gene in a thermophilic unicellular red alga. Curr. Genet. 21: 521–525.

Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, S., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nshizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N., Nozaki, H., Ogasawara, N., Kohara., Y. and Kuroiwa, T. 2004. Genome sequence of the ultra-small unicellular red alga Cyanidioschyzon merolae. Nature 428: 653–657.

Minoda, A., Sakagami, R., Yagisawa, F., Kuroiwa, T. and Tanaka, K. 2004. Improvement of culture conditions and evidence for nuclear transformation by homologous recombination in a red alga, *Cyanidioschyzon merolae* 10D. Plant Cell Physiol. 45: 667–671.

Miyagishima, S-Y., Itoh, R., Aita, S., Kuroiwa, H. and Kuroiwa, T. 1999. Isolation of dividing chloroplasts with intact plastid-dividing rings from a synchronous culture of the unicellular red alga *Cyanidioschyzon merolae*. Planta 209: 371–375.

Mochizuki, T., Onda, Y., Fujiwara, E., Wada, M. and Toyoshima, Y. 2004. Two independent light signals cooperate in the activation of a plastid *psbD* blue light-responsive promoter in *Arabidopsis*. FEBS Lett. 571: 26–30.

Moreira, D., Guyader, H.L. and Philippe, H. 2000. The origin of red algae and the evolution of chloroplasts. Nature 405: 69–72.

Nagashima, A., Hanaoka, M., Shikanai, T., Fujiwara, M., Kanamaru, K., Takahashi, H. and Tanaka, K. 2004a. The multiple-stress responsive plastid sigma factor, SIG5, directs activation of the *psbD* blue light-responsive promoter (BLRP) in *Arabidopsis thaliana*. Plant Cell Physiol. 45: 357–368.

Nagashima, A., Hanaoka, M., Motohashi, R., Seki, M., Shinozaki, K., Kanamaru, K., Takahashi, H. and Tanaka, K. 2004b. DNA microarray analysis of plastid gene expression in an *Arabidopsis* mutant deficient in a plastid transcription factor sigma, SIG2. Biosci. Biotechnol. Biochem. 68: 694–704.

Nakahira, Y., Baba, K., Yoneda, A., Shiina, T. and Toyoshima, Y. 1998. Circadian-regulated transcription of the *psbD* light-responsive promoter in wheat chloroplasts. Plant Physiol. 118: 1079–1088.

Nakamura, T., Furuhashi, Y., Hasegawa, K., Hashimoto, H., Watanabe, K., Obokata, J., Sugita, M. and Sugiura, M. 2003. Array-based analysis of tobacco plastid transcripts: preparation of a genomic microarray containing all genes and all intergenic regions. Plant Cell Physiol. 44: 861–867.

Ohta, N., Sato, N. and Kuroiwa, T. 1998. Structure and organization of the mitochondrial genome of the unicellular red alga *Cyanidioschyzon merolae* deduced from the complete nucleotide sequence. Nucleic Acids Res. 26: 5190–5298.

Ohta, N., Matsuzaki, M., Misumi, O., Miyagishima, S., Nozaki, H., Tanaka, K., Shin-I, T., Kohara, Y. and Kuroiwa, T. 2003. Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Res. 10: 67–77.

Ohyama k., , Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322: 572–574.

Oikawa, K., Tanaka, K. and Takahashi, H. 1998. Two types of differentially photo-regulated nuclear genes that encode sigma factors for chloroplast RNA polymerase in the red alga *Cyanidium caldarium* RK-1. Gene 210: 277–285.

Orozco, E.M. Jr., Mullet, J.E., Hanley-Bowdoin, L. and Chua, N.-H. 1986. *In vitro* transcription of chloroplast protein genes. Methods Enzymol. 118: 232–253.

Reith, M.E. 1995. Molecular biology of rhodophyte and chromophyte plastids. Annu. Rev. Plant Physiol. Plant Mol. Biol. 46: 549–575.

Reith, M.E. and Munhollad, J. 1995. Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. Plant Mol. Biol. Rep. 13: 333–335.

Reith, M.E. and Munholland, J. 1991. An *hsp70* homolog is encoded on the plastid genome of the red alga, *Porphyra umbilicalis*. FEBS Lett. 294: 116–120.

Sexton, T.B., Christopher, D.A. and Mullet, J.E. 1990. Light-induced switch in barley *psbD–psbC* promoter utilization: a novel mechanism regulating chloroplast gene expression. EMBO J. 9: 4485–4494.

Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi- Shinozaki, K., Ohto, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Tkaiwa, F., Kato, A., Tohdoh, N., Shimada, H. and Shugiura, M. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. 5: 2043–2049.

Strimmer, K. and von Haeseler, A. 1996. Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. Mol. Biol. Evol. 13: 964–969.

Swofford, D.L. 2002. PAUP* 4.0: Phylogenetic analysis using parsimony, version 4.0b 10. Computer program distributed by Sinauer Associates, Sunderland: MA.

Tanaka, K., Oikawa, K., Ohta, N., Kuroiwa, H., Kuroiwa, T. and Takahashi, H. 1996. Nuclear encoding of a chloroplast RNA polymerase sigma subunit in a red alga. Science 272: 1932–1935.

Tanaka, K., Tozawa, Y., Mochizuki, N., Shinozaki, K., Nagatani, A., Wakasa, K. and Takahashi, H. 1997. Characterization of three cDNA species encoding plastid RNA polymerase sigma factors in *Arabidopsis thaliana*: Evidence for the sigma factor heterogeneity in higher plant plastids. FEBS Lett. 413: 309–313.

The Arabidopsis Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882.

Toda, K., Takahashi, H., Itoh, R. and Kuroiwa, T. 1995. DNA contents of cell nuclei in two Cyanidiophycea: *Cyanidioschyzon merolae* and *Cyanidium caldarium* Forma A. Cytologia 60: 183–188.

Tsunoyama, Y., Ishizaki, Y., Morikawa, K., Kobori, M., Nakahira, Y., Takeba, G., Toyoshima, Y. and Shiina, T. 2004. Blue light-induced transcription of plastid-encoded *psbD* gene is mediated by a nuclear-encoded transcription initiation factor, AtSig5. Proc. Natl. Acad. Sci. USA. 101: 3304–3309.

Valentin, K. and Zetsche, K. 1990. Rubisco genes indicate a close phylogenetic relation between the plastids of chromophyta and rhodophyta. Plant Mol. Biol. 15: 575–584.

Yao, J., Roy-Chowdhury, S. and Allison, L.A. 2003. AtSig5 is an essential nucleus-encoded Arabidopsis sigma-like factor. Plant Physiol. 132: 739–747.