

Research article

Open Access

## Chloroplast His-to-Asp signal transduction: a potential mechanism for plastid gene regulation in *Heterosigma akashiwo* (Raphidophyceae)

Melinda R Duplessis<sup>†1</sup>, Kenneth G Karol<sup>†1</sup>, Elinor T Adman<sup>2</sup>,  
Lauren YS Choi<sup>1</sup>, Michael A Jacobs<sup>3</sup> and Rose Ann Cattolico<sup>\*1</sup>

Address: <sup>1</sup>Department of Biology, University of Washington, Box 355325, Seattle, WA 98195-5325, USA, <sup>2</sup>Department of Biological Structure, University of Washington, Box 357420, Seattle, WA 98195-7420, USA and <sup>3</sup>Department of Medicine, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195-2145, USA

Email: Melinda R Duplessis - mdupless@u.washington.edu; Kenneth G Karol - kkarol@u.washington.edu; Elinor T Adman - adman@u.washington.edu; Lauren YS Choi - laurenys@yahoo.com; Michael A Jacobs - mikejac@u.washington.edu; Rose Ann Cattolico\* - racat@u.washington.edu

\* Corresponding author †Equal contributors

Published: 3 May 2007

Received: 15 November 2006

BMC Evolutionary Biology 2007, 7:70 doi:10.1186/1471-2148-7-70

Accepted: 3 May 2007

This article is available from: <http://www.biomedcentral.com/1471-2148/7/70>

© 2007 Duplessis et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Maintenance of homeostasis requires that an organism perceive selected physical and chemical signals within an informationally dense environment. Functionally, an organism uses a variety of signal transduction arrays to amplify and convert these perceived signals into appropriate gene transcriptional responses. These changes in gene expression serve to modify selective metabolic processes and thus optimize reproductive success. Here we analyze a chloroplast-encoded His-to-Asp signal transduction circuit in the stramenopile *Heterosigma akashiwo* (Hada) Hada ex Y. Hara et Chihara [syn. *H. carterae* (Hulburt) F.J.R. Taylor]. The presence, structure and putative function of this protein pair are discussed in the context of their evolutionary homologues.

**Results:** Bioinformatic analysis of the *Heterosigma akashiwo* chloroplast genome sequence revealed the presence of a single two-component His-to-Asp (designated TsgI/TrgI) pair in this stramenopile (golden-brown alga). These data represent the first documentation of a His-to-Asp array in stramenopiles and counter previous reports suggesting that such regulatory proteins are lacking in this taxonomic cluster. Comparison of the 43 kDa *H. akashiwo* TsgI with bacterial sensor kinases showed that the algal protein exhibits a moderately maintained PAS motif in the sensor kinase domain as well as highly conserved H, N, G<sub>1</sub> and F motifs within the histidine kinase ATP binding site. Molecular modelling of the 27 kDa *H. akashiwo* TrgI regulator protein was consistent with a winged helix-turn-helix identity – a class of proteins that is known to impact gene expression at the level of transcription. The occurrence of TrgI protein in actively growing *H. akashiwo* cells was verified by Western analysis. The presence of a PhoB-like RNA polymerase loop in TrgI and its homologues in the red-algal lineage support the hypothesis that TrgI and its homologues interact with a sigma 70 ( $\sigma^{70}$ ) subunit (encoded by *rpoD*) of a eubacterial type polymerase. Sequence analysis of *H. akashiwo rpoD* showed this nuclear-encoded gene has a well-defined 4.2 domain, a region that augments RNA polymerase interaction with transcriptional regulatory proteins and also serves in -35 promoter recognition. The presence/loss of the His-to-Asp pairs in primary and secondary chloroplast lineages is assessed.

**Conclusion:** His-to-Asp signal transduction components are found in most rhodophytic chloroplasts, as well as in their putative cyanobacterial progenitors. The evolutionary conservation of these proteins argues that they are important for the maintenance of chloroplast homeostasis. Our data suggest that chloroplast gene transcription may be impacted by the interaction of the His-to-Asp regulator protein (which is less frequently lost than the sensor protein) with the RNA polymerase  $\sigma^{70}$  subunit.

## Background

One of the simplest mechanisms for controlling gene transcription response to environmental cues is mediated through a "two component" or "His-to-Asp" signal transduction system. In its most minimal configuration, this system is composed of two polypeptides, a sensor kinase and a response regulator protein that communicate via phosphotransfer events [reviewed in [1,2]]. A phosphoryl group is moved from a conserved histidine residue within the sensor kinase protein to an aspartic acid residue on its cognate response regulator. The response regulator impacts transcriptional activity by influencing promoter selection via its interaction with selected DNA targets [3] and with RNA polymerase [4,5]. His-to-Asp signal transduction systems were first characterized in *Escherichia coli* [6-8]. In eubacteria, 0 to more than 300 His-to-Asp proteins have been shown to occur [see [9-11] for discussion]. Recent complete genome sequencing endeavours document that the cyanobacterial genomes of *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120 contain about 80 and 208 general signal transduction proteins, respectively [12,13].

Given the endosymbiotic origin of plastids from a cyanobacterial-like ancestor [14,15], it is not surprising that genes for His-to-Asp signal transducers have been found encoded in chloroplasts. In this study, the wall-less, unicellular alga *Heterosigma akashiwo* is used as a model system for analyzing a chloroplast-encoded, two-component, signal transduction system. Each *H. akashiwo* cell contains approximately 30 discoidal chloroplasts [16]. These organelles are surrounded by four membranes, indicative of their serial endosymbiotic origin from a putative rhodophytic ancestor [14,17]. As an obligate autotroph, *H. akashiwo* is dependent on chloroplast function for survival and responds to changing environmental cues by rapidly altering transcript levels within the plastid [18,19]. Run-on analysis clearly demonstrates that the changes in mRNA abundance are largely due to transcript initiation [19].

This report presents data suggesting that the *Heterosigma akashiwo* response regulator component [20] of a His-to-Asp signal transduction circuit interacts with a nuclear-encoded sigma 70 ( $\sigma^{70}$ ) subunit of a eubacterial-like RNA polymerase to modulate chloroplast gene transcription. We propose that this simplified two-component/ $\sigma^{70}$  - partnership found in *H. akashiwo* may offer insight to a mechanism by which chloroplast gene transcription is controlled in certain algal taxa.

## Results

### *Tsg1* – a sensor kinase protein

The *Heterosigma akashiwo* chloroplast genome [21] contains a single His-to-Asp sensor kinase (transcriptional

sensor gene 1, *tsg1*). The presence of a chloroplast encoded His-to-Asp sensor kinase gene is not universal in stramenopiles or other plastid-containing organisms (Table 1). A *tsg1* homologue is found in the haptophyte *Emiliania huxleyi* (annotated as *dfr*) but not in the chloroplasts of the bacillariophytes *Odontella sinensis*, *Phaeodactylum tricorutum* and *Thalassiosira pseudonana*, the pelagophyte *Aureoumbra lagunensis*, the cryptophyte *Guillardia theta*, and the glaucophyte *Cyanophora paradoxa*. In rhodophytes, a *tsg1* homologue is encoded in the chloroplasts of *Cyanidium caldarium* (*ycf26*), *Gracilaria tenuistipitata* var. *liui* (*dfr*), *Porphyra purpurea* (*ycf26*) and *P. yezoensis* (hypothetical chloroplast protein 26), but is absent in *Cyanidioschyzon merolae*. No *tsg1* genes have been identified in any chlorophyte or charophyte examined to date.

The predicted amino acid sequence of *Heterosigma akashiwo* Tsg1 (43 kDa) has a shorter N terminus than that observed in rhodophytic algal sensor proteins and similar proteins found in cyanobacteria (Figure 1). Tsg1 lacks the transmembrane and HAMP domains [22] observed in some cyanobacterial and all chloroplast-encoded rhodophytic homologues (Figure 1), but it maintains the regions that contain the HisKA and HATPase domains. The Tsg1 protein has a weak PAS domain (Pfam e-value 5.4 and SMART [23,24] e-value  $6.03 \times 10^2$ ) that aligns with the well defined PAS domains of sensor kinases from *Porphyra purpurea*, *Gracilaria tenuistipitata* var. *liui*, *Cyanidium caldarium*, and *Synechocystis* sp. PCC 6803 (Figure 1).

The *Heterosigma akashiwo* Tsg1 kinase domain (Figure 1) displays a highly conserved H box (histidine phosphorylation site) whose consensus amino acid array HELRTP identifies this protein as a Type 1 (subtype B) histidine kinase [25]. Excellent sequence fidelity of the histidine kinase ATP binding site is also maintained for the N, G<sub>1</sub>, F and G<sub>2</sub> motifs, including the conserved glycine G<sub>3</sub> at its terminus [9,25,26]. The H to N distance between the histidine of the H box and the asparagine of the kinase domain is indicative of histidine kinase subtypes [25]. The 116 amino acid residues that lie between H and N in *H. akashiwo* Tsg1 and its homologues in the rhodophytic and chromophytic taxa are consistent with a Type 1 sensor protein identity. Nine lysine and three arginine residues contribute to the net positively charged amino acid sequence between the histidine kinase and ATPase domains in the *H. akashiwo* protein. Tsg1 of both *H. akashiwo* and rhodophytes shows close sequence similarity to the proteins Hik33 and NblS that serve as monitors of environmental stress in cyanobacteria [27,28].

### *Trg1* – a winged helix-turn-helix protein

Two identical copies of *trg1* were found on the chloroplast genome of *Heterosigma akashiwo*, one on each copy of the inverted repeat. The distribution of *trg1* is varied among

**Table 1: Distribution of sensor kinase and response regulatory proteins in photosynthetic plastids.**

Taxon/Organism	Sensor Kinase Copy Number	Sensor Kinase Accession Number	Response Regulator Copy Number	Response Regulator Accession Number
<b>Raphidophyte</b>				
<i>Heterosigma akashiwo</i>	1	<a href="#">EF115378</a>	2	<a href="#">CAB46638</a>
<b>Pelagophyte</b>				
<i>Aureoumbra laguensis</i>	0	NP	0	NP
<b>Bacillariophyte</b>				
<i>Odontella sinensis</i>	0	NP	0	NP
<i>Phaeodactylum tricornutum</i>	0	NP	0	NP
<i>Thalassiosira pseudonana</i>	0	NP	0	NP
<b>Haptophyte</b>				
<i>Emiliana huxleyi</i>	1	<a href="#">YP_277392</a>	1	<a href="#">YP_277359</a>
<b>Cryptophyte</b>				
<i>Guillardia theta</i>	0	NP	1	<a href="#">NP_050679</a>
<b>Glaucophyte</b>				
<i>Cyanophora paradoxa</i>	0	NP	2	<a href="#">NP_043139</a> , <a href="#">NP_043243</a> (partial)
<b>Rhodophyte</b>				
<i>Cyanidioschyzon merolae</i>	0	NP	1	<a href="#">NP_849001</a>
<i>Cyanidium caldarium</i>	1	<a href="#">AAFI2904</a>	2	<a href="#">O9TLQ4</a> , <a href="#">CAA44458</a>
<i>Gracilaria tenuistipitata</i> var. <i>liui</i>	1	<a href="#">YP_063707</a>	1	<a href="#">YP_063539</a>
<i>Porphyra purpurea</i>	1	<a href="#">AAC08278</a>	1	<a href="#">NP_053968</a>
<i>Porphyra yezoensis</i>	1	<a href="#">YP_537073</a>	1	<a href="#">YP_537039</a>
<i>Porphyridium aerugineum</i>	--	NP	1	<a href="#">CAA44464</a>
<i>Rhodella violacea</i>	--	NP	1	<a href="#">AAA62132</a>
<b>Charophyte</b>				
<i>Chlorokybus atmophyticus</i>	0	NP	1	<a href="#">ABM87971</a>

The dashed lines (--) indicate that the complete chloroplast genome for that taxon is not available, so the presence or absence of the gene is unknown. NP indicates not present.

disparate taxa (Table 1). The *trg1* gene has not been found in the pelagophyte *Aureoumbra laguensis*. This gene is also missing from both the nuclear and chloroplast genomes of *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* as well as the chloroplast genome of *Odontella sinensis*. A single copy of the *trg1* gene has been identified in the chloroplast genomes of the haptophyte *Emiliana huxleyi* (*ycf27*), the cryptophyte *Guillardia theta* (*ycf27*), and the rhodophytes *Cyanidioschyzon merolae* (*ycf27*), *Gracilaria tenuistipitata* var. *liui* (*ompR*), *Porphyra purpurea* (*ycf27*), *P. yezoensis* (hypothetical chloroplast protein 27), *Porphyridium aerugineum* (*ompR*) and *Rhodella violacea* (*orf246*; *ompR* homolog), whereas the glaucophyte *Cyanophora paradoxa*, has one full copy (*ycf27*) and a partial fragment of this gene (*orf27*). The rhodophyte *Cyanidium caldarium* chloroplast encodes two complete copies of *trg1* (*ompR* and *ycf27*), though these genes are relatively divergent from each other (E-value  $2.0 \times 10^{-78}$ ). Interestingly, a *trg1*-like gene (*ycf27*) has been identified in the plastid genome of the charophyte *Chlorokybus atmophyticus*.

Western analysis was used to demonstrate that the Trg1 protein is expressed *in vivo*. Data shown in Figure 2 confirmed the presence of Trg1 protein in exponentially growing *Heterosigma akashiwo* cells that were harvested at L3 of a 12 h light:12 h dark growth cycle. An expected pro-

tein band of 27 kDa was present when cell extracts were exposed to post-bleed antiserum (Day 50; Figure 2 lane 4) and was absent in the lane exposed to preadsorption control (Figure 2 lane 2) and prebleed antiserum (Figure 2 lane 6). These data confirmed that the low abundance message for the *trg1* gene [20] is translated into a protein product.

Response-regulatory proteins, such as Trg1, contain a receiver domain and a DNA recognition domain. The DNA recognition domain affects gene expression by interacting with both DNA and with RNA polymerase [29,30]. The inferred amino acid sequence of *Heterosigma akashiwo* Trg1 was compared with Trg1-like sequences from *Thermotoga maritima* (OmpR) and *Escherichia coli* (OmpR and PhoB) for which tertiary structures have been determined [30-33]. A number of conserved regions are shared among these sequences both within the receiver domain and DNA binding domain (Figure 3).

Comparison of *Heterosigma akashiwo* Trg1 and *Thermotoga maritima* OmpR shows good three-dimensional similarity (Figure 4). To gain insight into *H. akashiwo* Trg1-RNA polymerase structure/function relationship as a possible transcriptional regulator of chloroplast genes, we constructed a model of the DNA binding domain for this pro-

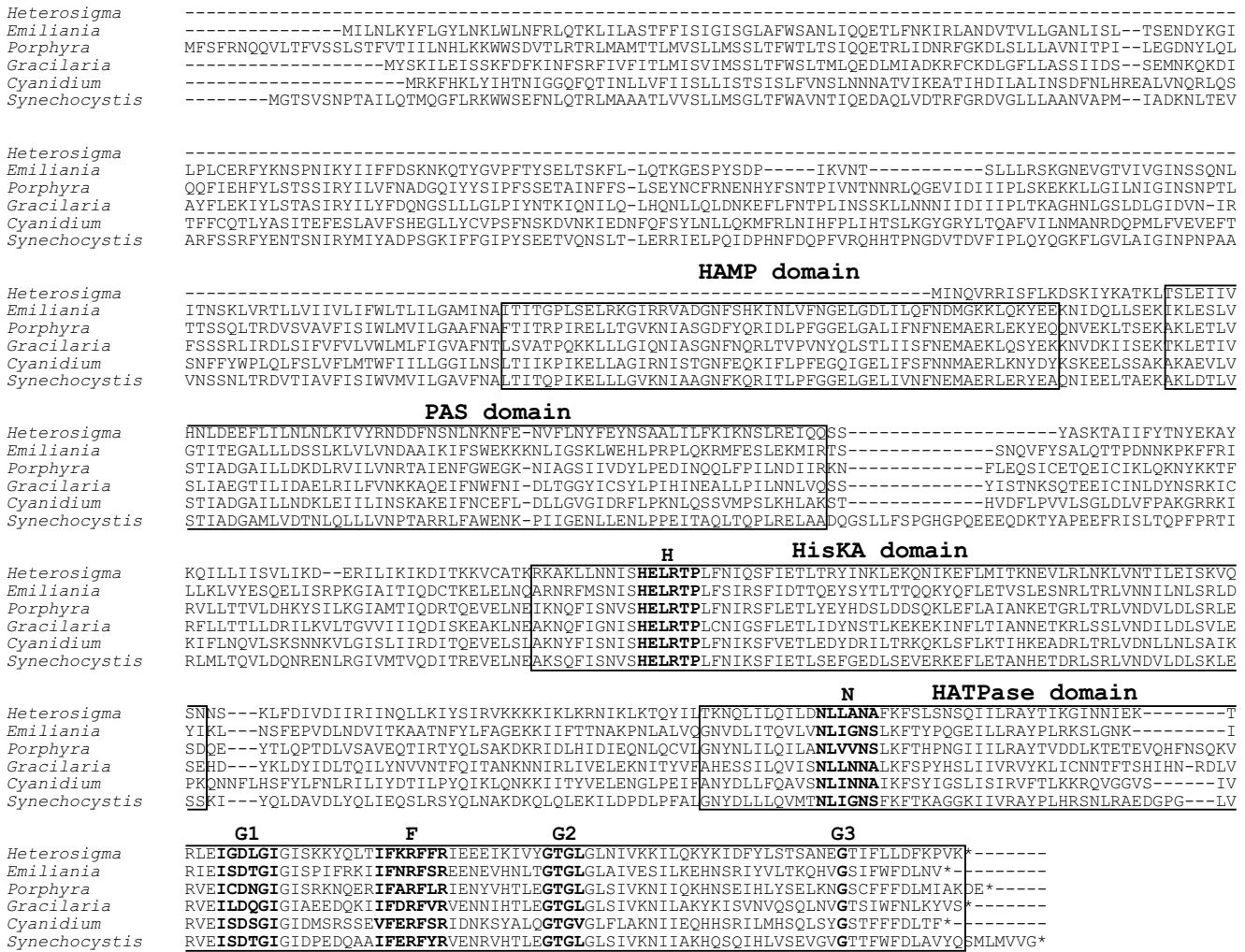


Figure 1

Figure 1

Sequence comparison of *Heterosigma akashiwo* TsgI with putative homologues in rhodophytic plastids and a cyanobacterium. Putative homologues of TsgI from *Heterosigma akashiwo* (accession number [EF115378](#)), *Emiliania huxleyi* (YP\_277392), *Porphyra purpurea* (AAC08278), *Gracilaria tenuistipitata* var. *liui* (YP\_063707), *Cyanidium caldarium* (AAF12904), and *Synechocystis* sp. PCC 6803 (BAA16687) are compared. The HAMP, PAS, HisKA, and HATPase domains are boxed. The H-box which contains the histidine residue, site of phosphorylation, N, G<sub>1</sub>, F, G<sub>2</sub>, and G<sub>3</sub> sites are bolded.

tein. Because *T. maritima* lacks the  $\alpha_2$ - $\alpha_3$  polymerase loop, our model of this region is based on available three-dimensional structures for OmpR and PhoB. OmpR has an X-ray generated structure [32] at 1.95 Å resolution (10 PC) whereas the PhoB report represents an NMR (1QQI) derived structure [30]. Structures of OmpR and PhoB have been previously compared, and our observations are based upon that comparison [31]. The most notable differences between OmpR and PhoB occur in the putative eubacterial-like RNA polymerase contact loop ( $\alpha_2$ - $\alpha_3$  region) and at the loop between helix  $\alpha_3$  and strand  $\beta_5$ . The higher sequence similarity between Trg1 and PhoB in the  $\alpha_2$ - $\alpha_3$  region leads us to believe that PhoB is a better

structure model for residues 130–208 and similarly OmpR for residues 209–231 (Figures 3 and 4). Thus, we constructed a chimeric model from these two portions, which yielded satisfactory internal and external environments for most residues.

Mutational analysis of the PhoB "RNA polymerase loop" region shows that the amino acids W184, G185, V190 and D192 are needed for successful association of this response regulator protein with the  $\sigma^{70}$  subunit of prokaryotic RNA polymerase [30]. The first two of these amino acids are present in Trg1 (as is V183) but V190 is conservatively replaced by I192, and the D192 has been

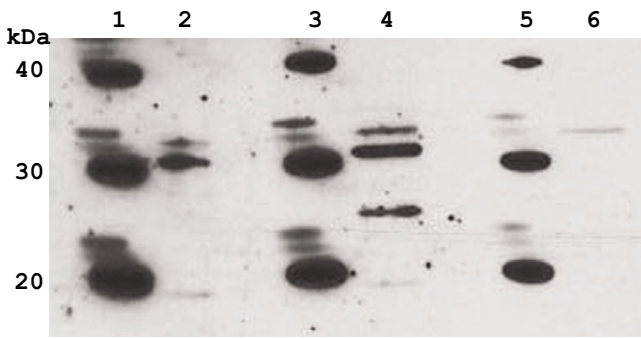


Figure 2

**Figure 2**  
**Western blot showing the expression of Trg1 in *Heterosigma akashiwo*.** Total soluble proteins were separated on NuPAGE Novex Bis-Tris gel and probed with preadsorption control (lane 2), anti-Trg1 peptide antiserum, day 50 (lane 4) and preinjection antiserum, day 0 (lane 6). A band of the expected estimated size of 27 kDa is present in the anti-Trg1 peptide lane 4, but not in the two negative controls (preadsorption and day 0). Lanes 1, 3, and 5 contain "Magic Mark Western" protein standard.

replaced with a non-conservative I194 (Figure 3). Similar to observations in *Heterosigma akashiwo* Trg1, conservation of a PhoB-like domain is found within the response regulator  $\alpha_2$ - $\alpha_3$  loops of some cyanobacteria and proteobacteria, as well as in chloroplasts (Table 2). The observation that the *H. akashiwo* Trg1  $\alpha_2$ - $\alpha_3$  loop is two amino

acids longer than that seen in PhoB and OmpR suggests that either the response regulator may interact with an undescribed RNA polymerase subunit or the RNA polymerase subunit itself may also be modified.

**Presence of rpoD in *Heterosigma akashiwo***

Data presented above support the hypothesis that *Heterosigma akashiwo* Trg1 interacts with a  $\sigma^{70}$ -like subunit of a eubacterial-like RNA polymerase. Sequence analysis of *H. akashiwo* revealed the presence of a nuclear-encoded *rpoD* gene. The presence of both a signal peptide sequence and a putative stromal targeting domain on the amino terminus of the RpoD protein supports the hypothesis that this protein is chloroplast-targeted (Table 3). Although *H. akashiwo* RpoD lacks the autoinhibitory 1.1 region, it retains the highly conserved functional domains (1.2, 2.1 – 2.4, 3.0 – 3.2 as well as 4.1 and 4.3) that have been elucidated for eubacterial homologues ([34]; Figure 5). Most striking in the context of this study is the extensive maintenance of sequence identity within domains 4.1 and 4.2 among phylogenetically diverse organisms (Figure 5). This domain is responsible for interaction with transcriptional regulator proteins and with the -35 promotor array [34-36].

Phylogenetic analyses of *rpoD*/ $\sigma^{70}$  have been published for some eubacteria [4,37-40], and land plants [41,42]. We attempted to generate a phylogeny to place *Heterosigma akashiwo rpoD* among other photosynthetic eukaryotes (bacillariophytes, cryptophytes, glaucophytes, rhodophytes, chlorophytes, and charophytes). Since this

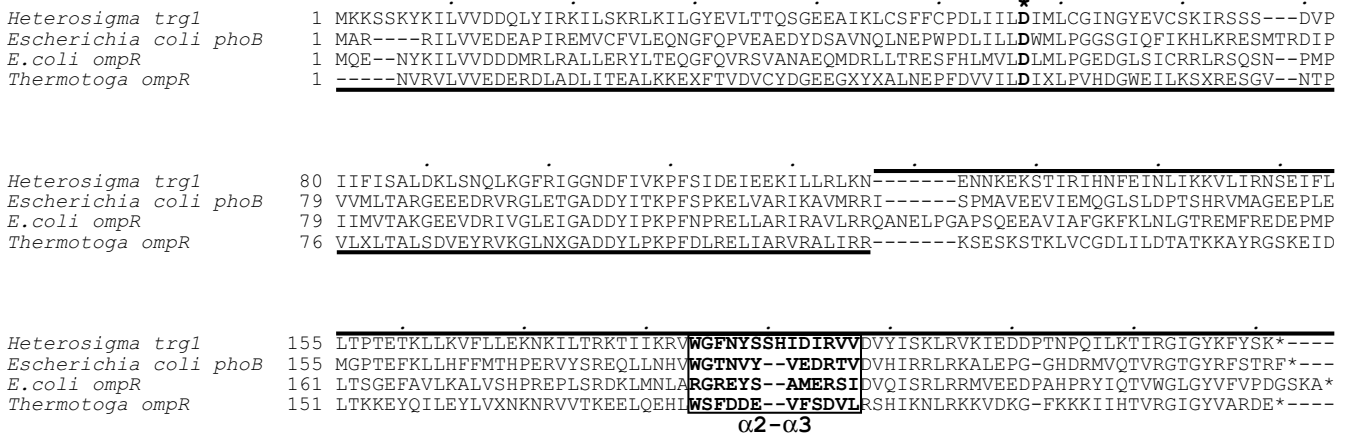


Figure 3

**Figure 3**  
**Sequence comparison of *Heterosigma akashiwo* Trg1 with select bacterial response regulators.** The *Heterosigma akashiwo* Trg1 sequence (accession number CAB46638) was aligned with *Escherichia coli* PhoB (P03025), *E. coli* OmpR (P08402) and *Thermotoga maritima* OmpR (IKGSA). Receiver and DNA binding domains are underlined and overlined, respectively. Within the receiver domain, the phosphorylation site is indicated by an asterisk (\*). Within the DNA binding domain, the RNA polymerase binding site ( $\alpha_2$ - $\alpha_3$  loop) is boxed and bolded.

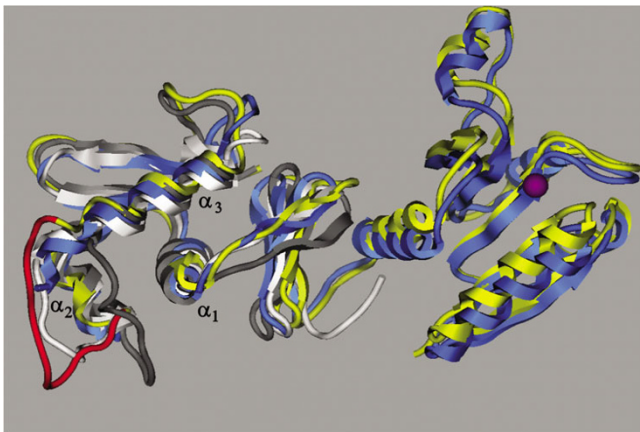


Figure 4

**Figure 4**  
**Molecular model of *Heterosigma akashiwo* TrgI.** Structural model of the DNA recognition domain of TrgI (shown in yellow) based on the defined partial structures of *Escherichia coli* OmpR (grey), PhoB (white), and the complete receiver-regulator structure from *Thermotoga maritima* (blue) reveals important similarities. The predicted TrgI model closely resembles that of OmpR, particularly in the putative DNA binding region ( $\alpha_3$  helix). Notably, the predicted TrgI structure for the putative RNA polymerase interaction site ( $\alpha_2$ - $\alpha_3$  loop, red) more closely matches that of PhoB. The phosphorylation site is shown as a purple sphere.

comparison yielded few regions that could be confidently aligned, the resulting trees were consistently poorly supported as measured by bootstrap values and posterior probabilities (results not shown). A consistent phylogenetic result, however, was the placement of *H. akashiwo* *rpoD* as sister to one of the six *rpoD*/*Sig* sequences that had been annotated from the complete nuclear genome of *Thalassiosira pseudonana* [43].

A number of problems common in phylogenetic inference could explain our difficulty in generating a robust phylogeny. These include problems associated with paralogy, lineage sorting and horizontal gene transfer [for discussion see [44,45]], long-branch attraction [46], and rate heterogeneity among lineages [47].

## Discussion

Most plausibly, the prokaryotic ancestor(s) of eukaryotic chloroplasts were fully equipped with their own multifaceted signal transduction circuitry [14,48], as observed in cyanobacterial and proteobacterial genomes which encode multiple sensor kinase/response regulator pairs [11,12,49]. Since the chloroplast must respond to both extra- and intra-cellular cues, one might anticipate strict conservation of these signal transduction arrays. However, given the observations that fewer sensor/response

circuits exist in intracellular bacterial pathogens than in free-living representatives [50], one may argue that a reduction in the ancestral plastid genome size after endosymbiosis may have driven the loss of chloroplast-encoded His-to-Asp regulatory arrays (see Table 1).

The chloroplast genomes (Table 1) of several rhodophytic algae [51-56], the glaucophyte *Cyanophora paradoxa* [57], the haptophyte *Emiliania huxleyi* [58], the cryptophyte *Guillardia theta* [59], and the charophyte *Chlorokybus atmophyticus* [60] have been shown to encode the response regulator gene and in some, the sensor kinase gene for the His-to-Asp proteins. The presumptive loss of the sensor kinase seen in some chloroplast genomes may suggest that under such a circumstance, the regulatory protein may be governed by nuclear-encoded sensor kinases or by yet undescribed accessory proteins that are either of nuclear or chloroplast origin [see [9] for discussion]. The occurrence of this signalling array is less well documented in stramenopiles. The chloroplast genome of the raphidophyte *Heterosigma akashiwo* encodes a single response regulator and its cognate sensor kinase [[20] and this study], but the chloroplast genomes of representatives within the bacillariophytes and the pelagophytes lack both proteins of this two-component system [61,62]. More than 75 green plant ( $\sim 9$  chlorophyte and  $\sim 66$  charophyte) chloroplast genomes have been sequenced. Of these only the charophyte *Chlorokybus atmophyticus* encodes a response regulator protein. It should be noted, however, that partial footprints of (laterally transferred?) prokaryotic His-to-Asp transduction pairs have been identified in some green-plant nuclear genomes [63-66]. Such truncated His-to-Asp constructs have been shown to signal mitogen-activated protein kinase cascades, which cause the differential regulation of targeted genes [63,67].

The taxonomic distribution of this gene-pair is poorly understood. Complete chloroplast genomes are few, especially in species-rich lineages that are represented by only a small number of complete chloroplast genomes (e.g., the bacillariophytes). Whether the maintenance of the His-to-Asp signal transduction apparatus in the chloroplast corresponds to established phylogenies remains to be determined. One might anticipate that the chloroplasts of chromophytic algae would retain this His-to-Asp array since they are the product of a serial endosymbiotic event, which involved a rhodophytic algal ancestor. Unfortunately, data for the diverse taxonomic assemblage of chromophytes has been both minimal and conflicting. The presence of the His-to-Asp array in *Heterosigma akashiwo* appears to reflect the retention of an ancestral signature. Whether the loss of the His-to-Asp pair in bacillariophyte chloroplast DNA represents a derived genotype, remains an open question. Regardless of phylogenetic profile, the evolutionary retention of all or part of a

**Table 2: Comparison of putative RNA polymerase association loop domain among prokaryotic and plastid-encoded response regulators.**

Organism	Gene Name	Taxon	Putative RNA Polymerase Loop Domain	Accession Number
<b>Heterosigma akashiwo</b>	<b>trg1</b>	<b>raphidophyte</b>	<b>WGFNYSSHIDIRVV</b>	<b>CAB46638</b>
<i>Emiliana huxleyi</i>	<i>ycf27</i>	haptophyte	WGYTPERYLDTRVV	<a href="#">AAX13858</a>
<i>Guillardia theta</i>	<i>ycf27</i>	cryptophyte	WGYTPERHVDTRVV	<a href="#">AAC35613</a>
<i>Cyanophora paradoxa</i>	<i>ycf27</i>	glaucoephyte	WGYTPERHIDTRVV	<a href="#">AAA81319</a>
<i>Cyanidioschyzon merolae</i>	<i>ycf27</i>	rhodophyte	WGYAWPNETRVV	<a href="#">BAC76163</a>
<i>Cyanidium caldarium</i>	<i>ycf27</i>	rhodophyte	WGYKLSKHEPIADTRIV	<a href="#">AAF12879</a>
<i>Cyanidium caldarium</i>	<i>ompR</i>	rhodophyte	WGYTPERHLDTRVV	<a href="#">CAA44458</a>
<i>Gracilaria tenuistipitata</i> var. <i>liui</i>	<i>ompR</i>	rhodophyte	WGYKPERHVDTRVV	<a href="#">YP_063539</a>
<i>Porphyra purpurea</i>	<i>ycf27</i>	rhodophyte	WGYTPERHVDTRVV	<a href="#">AAC08244</a>
<i>Porphyra yezoensis</i>	<i>ycf27</i>	rhodophyte	WGYTPERHVDTRVV	<a href="#">QIXDC9</a>
<i>Porphyridium aerugineum</i>	<i>ompR</i>	rhodophyte	WGYTAERQVDTRVV	<a href="#">CAA44464</a>
<i>Rhodella violacea</i>	<i>ompR</i>	rhodophyte	WGYTPERHIDTRVV	<a href="#">AAA62132</a>
<i>Chlorokybus atmophyticus</i>	<i>ycf27</i>	charophyte	WEYGNDSYIDTRVV	<a href="#">ABM87971</a>
<i>Nostoc</i> sp. PCC7120	<i>ycf27</i>	cyanophyte	WGYTPERHVDTRVV	<a href="#">BAB75521</a>
<i>Synechocystis</i> sp. PCC6803	<i>ycf27</i>	cyanophyte	WGYTPERHVDTRVV	<a href="#">BAA18408</a>
<i>Tolythrix</i> sp. PCC7601	<i>rpaB</i>	cyanophyte	WGYTPERHVDTRVV	<a href="#">AAD30119</a>
<i>Bacillus subtilis</i>	<i>recM</i>	Firmicutes	WGYDYFGDVRTV	<a href="#">BAA05173</a>
<i>Clostridium petringens</i>	<i>ompR</i>	Firmicutes	WGYEYIGETRTV	<a href="#">BAB80348</a>
<i>Lactobacillus sakei</i>	<i>rrp3</i>	Firmicutes	WGYDYFGDVRTV	<a href="#">AAD10263</a>
<i>Listeria innocua</i>	<i>ompR</i>	Firmicutes	WGYDYFGDVRTV	<a href="#">CAC95548</a>
<i>Streptomyces coelicolor</i>	<i>scrA</i>	Firmicutes	WGYRHAADTRLV	<a href="#">AAG15433</a>
<i>Thermotoga maritima</i>	<i>drpA</i>	Thermotogales	WGYDYYGDTRTV	<a href="#">AAD36722</a>
<i>Ralstonia metallidurans</i>	<i>czcR</i>	Proteobacteria	WGVNFDTDTNVV	<a href="#">CAA67086</a>
<b><i>Escherichia coli</i></b>	<b><i>phoB</i></b>	<b>Proteobacteria</b>	<b>WGTNYYVEDRTV</b>	<b>P03025</b>
<b><i>Escherichia coli</i></b>	<b><i>ompR</i></b>	<b>Proteobacteria</b>	<b>RGREYSAMERSI</b>	<b>P08402</b>

The putative RNA polymerase association loop domain was identified in prokaryotic and plastid-encoded response regulators based on molecular modelling for Trg1, PhoB, and OmpR [31]. For easy comparison in the table – the data for Trg1, PhoB, and OmpR are bolded.

His-to-Asp signal transduction circuit in some distantly related algal chloroplasts strongly suggests that this biochemical mechanism must play an important role in the maintenance of chloroplast homeostasis.

We propose that the *Heterosigma akashiwo* Tsg1/Trg1 signal transduction pair, in concert with an RNA polymerase  $\sigma^{70}$  subunit, is involved in regulating chloroplast gene transcription. The environmental stimulus that regulates the signal transduction response remains elusive. The inability to create gene-knockout mutants or perform transformation experiments in chromophytic algae (except diatoms, which lack His/Asp genes) has hampered gene

expression studies that are needed to provide direct evidence for the role of His/Asp systems in chloroplast function. However, one might infer function given the similarity of the chloroplast-encoded *H. akashiwo* Tsg1 and the *ycf26*-encoded proteins of *Emiliana huxleyi*, *Cyanidium caldarium*, *Gracilaria tenuistipitata* var. *liui*, *Porphyra purpurea* and *P. yezoensis* to the cyanobacterial sensor kinases and Nbls. Hik33 has been shown by deletion studies to impact the expression of selected genes in response to osmotic and low temperature stress [27,68,69], while its homologue Nbls is reported to serve as a sensor of nutrient stress and high light intensity [28]. The underlying mechanism driving these physiological

**Table 3: *Heterosigma akashiwo* signal peptides**

Precursor Protein	Accession Number	Signal peptide	Beginning of STD
RpoD	<a href="#">EF115377</a>	VAHCTTFAYKGSNMRSFFHMLWSVVAATAA	<b>F</b> MMPGR...
Fcp1	<a href="#">X99697</a>	MSLKLATLAAALMGASA	<b>F</b> VAPNKM...
Fcp2	<a href="#">EF115376</a>	VVFDLYASTIPSVNQAHKSKMSLKLATFAAALAGASA	<b>F</b> VAPNQM...
PRK	<a href="#">EF115375</a>	MMYKLATLLALLPAVVA	<b>F</b> TTSFNG...
PsbO	<a href="#">AY130990</a>	MKFVAVLVCLMVSVAVVA	<b>F</b> KTQRN...

*Heterosigma akashiwo* RNA polymerase D subunit (RpoD), fucoxanthin chlorophyll binding protein 1 (Fcp1), fucoxanthin chlorophyll binding protein 2 (Fcp2), phosphoribulose kinase (PRK), and 33 kDa oxygen-enhancer 1 protein (PsbO) precursors showing the signal peptide and the start of stromal targeting domain (STD). The conserved phenylalanine (F) indicating the beginning of stromal targeting domain [62, 97, 98] is marked in bold.

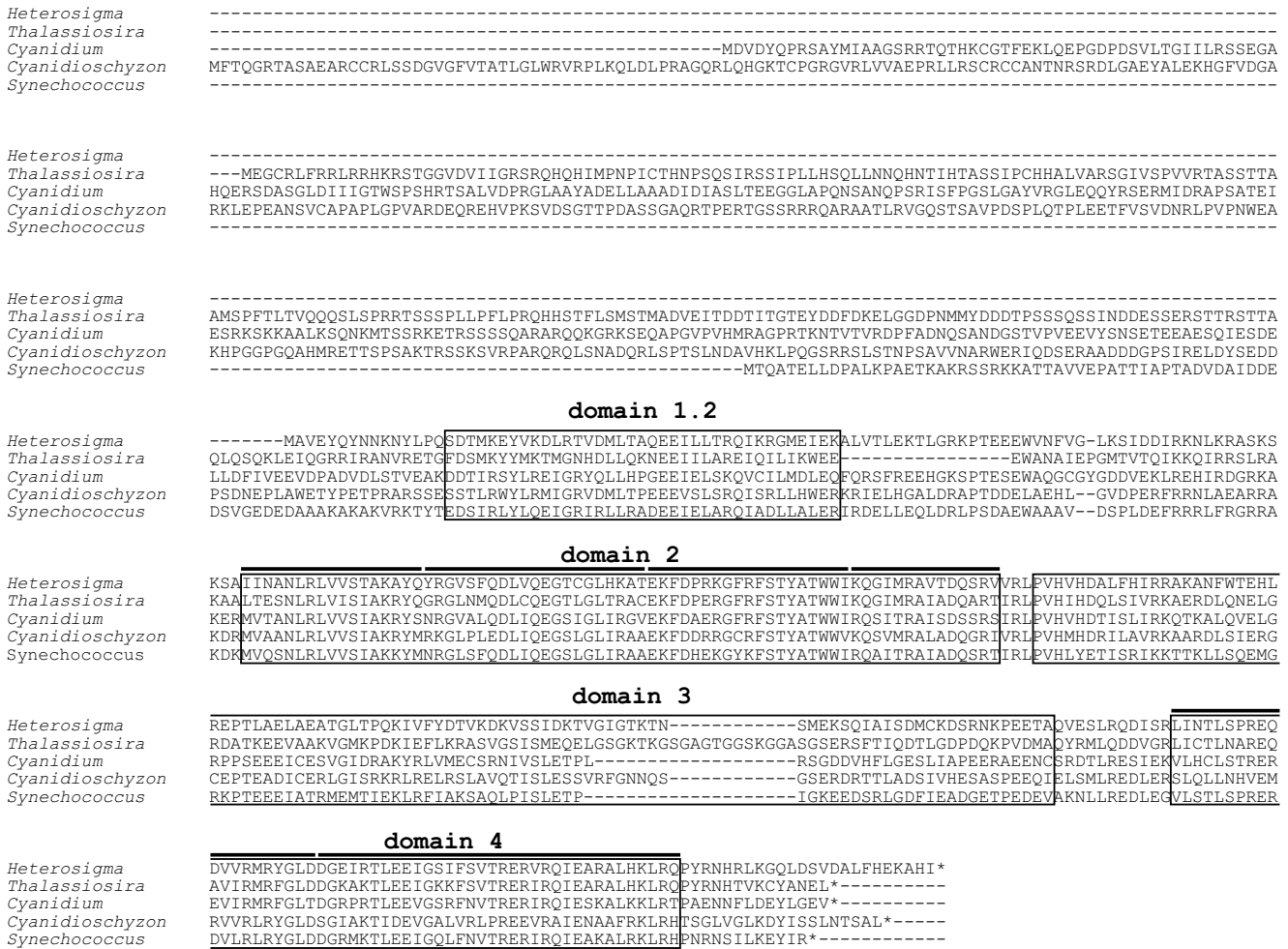


Figure 5

**Figure 5**  
**Sequence comparison of *Heterosigma akashiwo* RpoD with putative homologues in chromophytic and rhodophytic algae, and a cyanobacterium.** Putative homologues of RpoD from *Heterosigma akashiwo* (accession number [EF115377](#)), *Thalassiosira pseudonana* [43], *Cyanidium caldarium* ([BAA11832](#)), *Cyanidioschyzon merolae* [102], and *Synechococcus* sp. PCC 7942 ([BAA01749](#)) are compared. Conserved domains (1.2, 2, 3, and 4) are boxed and sub-domains (2.1–2.4, 4.1, and 4.2) are identified as bars above their respective boxes.



responses may be governed by redox and light signals for two reasons [discussed in [27,28]]: (a) both Hik33 and NblS possess a PAS domain, which is thought to be involved in redox and light sensing [70] and, (b) a large majority of the genes impacted by Hik33 and NblS are related to photosynthesis [27,28]. It has been proposed that redox control of gene expression is a fundamental evolutionary selection mechanism responsible for the maintenance of chloroplast-encoded gene regulation systems [71,72]. As shown in Figure 1, the *H. akashiwo* Tsg1 protein has a putative PAS domain. While the three dimensional structure of PAS domains is conserved among taxa, the primary protein sequences that comprise this motif are often diverse [73]. As more types of PAS domains are characterized, the E-value for the PAS domain in *H. akashiwo* Tsg1 should become more robust.

An important difference between *Heterosigma akashiwo* Tsg1 and its homologues is the absence of a transmembrane region and a HAMP domain. The possibility of a split *tsg1* gene was not supported by detailed analysis of the completely sequenced *H. akashiwo* chloroplast genome. The absence of a transmembrane region implies that Tsg1 is most likely present in the stroma. Though the majority of described Tsg1 proteins are putatively membrane bound (data not shown, SMART database search [23,24]), soluble histidine kinases have been identified [74,75]. Studies are underway using a Tsg1 peptide antibody to verify the location of the protein in the cell.

Contrary to a previous report [76], our data indicate that RpaB in the cyanobacterium *Synechocystis* shares not only similarity to the *ycf27* proteins in red algae but also to *H. akashiwo* Trg1. The premise that *ycf27* homologues are restricted to eukaryotic algae containing phycobilisomes [76,77] is contrary to the description of this protein in the non-phycobilisome containing algae – *Heterosigma akashiwo* (Trg1) [20], *Guillardia theta* (*ycf27*) [59], *Emiliania huxleyi* (*ycf27*) [58], and *Chlorokybus atmophyticus* (*ycf27*) [60]. Nonetheless, the hypothesis that RpaB regulates the synthesis of (unknown) "factors required to couple phycobilisomes to PS1 or PSII" [76,77] is consistent with the possible role of this protein in redox/light sensing.

The assignment of *Heterosigma akashiwo* Trg1 to the "Class 2" (or "ompR" super family) of transcriptional regulators offers additional insight to its function in the plastid. In prokaryotic cells, some "Class 2" proteins (such as PhoB) regulate transcription through interaction with the  $\sigma$  subunit of RNA polymerase [78] while others associate with the  $\alpha$  subunit of this enzyme [29,79]. We have identified a PhoB-like signature for the RNA polymerase recognition domain in Trg1 and a putative chloroplast-targeted  $\sigma^{70}$  subunit in *H. akashiwo*. A comparable transcriptional

mechanism appears to be present in *Cyanidioschyzon mero-lae*, *Cyanidium caldarium*, and *Guillardia theta* as both PhoB-like signatures and  $\sigma$  subunits have been identified (Table 2) [55,80-83].

Sigma factors in concert with core RNA polymerase selectively target chloroplast genes for transcription. For example, the prokaryotic-like, plastid-encoded polymerases with their associated  $\sigma$  factor(s) exclusively transcribe many genes that impact the photosynthetic process including *rbcl*, *psbA*, *psbD*, *petB*, *ndhA*, *atpI*, *atpH* and *rps14* [84-87]. Both plastid-encoded polymerase and the phage-like nuclear-encoded polymerase can transcribe *rrnA*, *atpB*, *clpP*. It should be noted however, that these eubacterial RNA polymerase-associated  $\sigma$  factors often interact with regulatory proteins (such as Trg1) and this association may further influence the transcription of specific genes [3].

How could the transcription of a small, select set of genes impact chloroplast homeostasis? One might propose that a hierarchical assembly of proteins during the formation of molecular complexes could provide an exceptionally efficient mechanism for regulating the quantitative and qualitative production of molecular structures necessary for the maintenance of chloroplast function. In the chloroplast, many large functional complexes that drive oxygenic photosynthesis and carbon fixation are constructed with a definitive stoichiometry that reflects the cooperative interaction between plastid and nuclear genomes. Studies suggest that protein complex formation is regulated by the presence of a "dominant assembly partner" whose presence assures the production of its assembly associates in the correct proportions [88]. For example, a *Chlamydomonas* mutant lacking the D1 protein (encoded by *psbA*) expresses only minimal levels of D2 (*psbB*) as well as CP47 (*psbC*) proteins. Similarly, mutants in CP47 (*psbC*) or D2 (*psbB*) show depressed concentrations of D1 (*psbA*) protein. In contrast, *Chlamydomonas* cells that were mutated in CP43 (*psbD*) were able to assemble D1, D2 and CP47 into a stable complex. The existence of such an assembly cascade is not restricted to photosystem II construction. When cytochrome b6 (*petB*) or subunit IV (*petD*) are not present, cytochrome f (*petA*) synthesis drops to 10% of that in wild type cells. A similar synthesis cascade appears also to occur in the biogenesis of ATP synthetase, which is comprised of five subunits ( $\alpha, \beta, \gamma, \delta, \epsilon$ ). Mutants lacking  $\beta$  will not accumulate any other core peptide. In effect, a minimal signal transduction system (encoded in the chloroplast) in conjunction with the  $\sigma$  subunit (encoded in the nucleus) may have given the ancestral eukaryotic cell a simple and efficient method to integrate chimeric gene sets.

## Conclusion

We have identified a His-to-Asp signal transduction array in the secondary endosymbiotic chloroplast of the stramenopile *Heterosigma akashiwo*. These proteins are similar to those found in bacteria and in the chloroplast genomes of several, though not all, algae.

This study generated a number of interesting questions. For example, why have several of the red lineage chloroplasts retained all or part of the His-to-Asp signal transduction system while only a single green chloroplast lineage retained this system? Will the sensor kinase protein be more frequently missing from this pair as additional stramenopile chloroplast genomes are analyzed? Are there undescribed chloroplast-encoded proteins that can substitute for sensor kinases? Are all sensor kinases triggered by the same or by different environmental cues? What are those cues? When more than one RpoD protein is present in an organism (e.g., *Thalassiosira*), does one member of this  $\sigma^{70}$  factor family interact with the regulator protein or do different sigma/response regulators form partnerships that control the expression of specific gene sets? With the advent of high throughput genome sequencing, more effective bioinformatics/evolutionary analysis as well as extensive molecular studies of cellular processes are now possible. From the questions posed above, it is evident that events associated with chloroplast gene regulation will continue to provide a challenging field for future research.

## Methods

### Cell culture

All algal cultures were maintained on a 12 hr light:12 hr dark (diel) photoperiod using cool white light (60 to 80  $\mu\text{Em}^{-2}\text{s}^{-1}$ ) with continuous rotary shaking at 60 rpm. *Heterosigma akashiwo* (Hada) Hada ex Hara et Chihara (strain CCMP-452) was originally obtained from Sarah Gibbs (McGill University, Canada). One liter vegetative cultures were axenically maintained on an artificial sea water medium as previously described [20,89]. Cells were counted using a Coulter counter (model Z2 Coulter Particle Count and Size Analyzer; Beckman Coulter, Fullerton, CA, USA) equipped with a 100  $\times$  120  $\mu\text{m}$  aperture.

### RNA isolation

Axenic *Heterosigma akashiwo* cultures were harvested at a density of 5  $\times$  10<sup>5</sup> cells/mL and a modified protocol [89] was used to obtain total RNA from these cells. All steps were carried out with chilled reagents and performed at 4°C. *H. akashiwo* cells were collected by centrifugation at 1,000  $\times$  g for 5 min. Pelleted cells were resuspended to a concentration of 2  $\times$  10<sup>7</sup> cells/mL in a pH 7.5 buffer that contained 25 mM KCl, 25 mM MgCl<sub>2</sub>, and 25 mM Tris (KMT buffer). KMT-saturated phenol was added to the lysed cell mixture at a ratio of (1.5:1). The phenol:super-

natant mixture was inverted slowly using a rotating wheel for 20 min, followed by centrifugation at 9,750  $\times$  g for 15 min in a fixed angled rotor. An equal volume of KMT-saturated phenol:chloroform:isoamyl alcohol (75:24:1) was added to the retrieved supernatant and the inversion-centrifugation steps repeated. The supernatant was removed and extracted twice more as described above but with chloroform:isoamyl alcohol (24:1) at a 1:1 ratio. RNA was precipitated by adding 2 volumes of -20°C ethanol (95%) and chilling overnight at -20°C after which it was pelleted at 9,750  $\times$  g for 15 min. The ethanol was decanted and the RNA was air dried at room temperature. One mL of RNase free water was added to the RNA and left to dissolve on ice for 15 min. To remove DNA contamination from total RNA used for RACE, 40  $\mu\text{g}$  of total RNA was incubated with 10 U of DNase I in a 100  $\mu\text{L}$  reaction volume according to the manufacturer's instructions (Amplification Grade, Invitrogen, Carlsbad, CA, USA). After the 15 min incubation period with DNase I, total RNA was purified using the Qiagen RNeasy kit (Qiagen, Valencia, CA, USA) following the RNeasy Mini Protocol for RNA cleanup. The removal of DNA contaminants was monitored by visualizing purified RNA on a 1.5% agarose gel containing 2.2 M formaldehyde and 1X MOPS [90].

### Amplification of 550 bp from rpoD

Primers were designed to conserved regions of the *rpoD* gene by aligning  $\sigma$  factors from *Cyanidium caldarium* SigC (accession number [BAA25788](#)), *Guillardia theta* ([BAB87262](#)) and *Thalassiosira pseudonana* [43]. PCR amplification of the partial *rpoD* fragment was carried out in a 50  $\mu\text{L}$  reaction that contained 100 ng of template DNA, 10 pmole of each primer ORAC 1104 and ORAC 1105 (Table 4), 1.5 Units of *Taq* Polymerase, 1X *Taq* polymerase buffer, 1.5 mM MgCl<sub>2</sub> (Promega, Madison, WI, USA), and 0.2 mM of each dNTP (Invitrogen). Initial denaturation of DNA was at 94°C for 3 min followed by

**Table 4: Primers used for PCR, RACE, and genome-walking to generate *rpoD* and *tsgI* gene sequences.**

Primer name	Sequence
<b>PCR</b>	
M13 forward	5'-GTAAAACGACGGCCAG-3'
M13 reverse	5'-CAGGAAACAGCTATGAC-3'
ORAC 1104	5'-AGGCTGACGAAGCTTGTGCAA-3'
ORAC 1105	5'-TTCTCTACCTACGCAACATGGTGG-3'
<b>RACE</b>	
ORAC 240	5'-GATACGGTGAAGGACAAGGTCTCATC-3'
ORAC 242	5'-CAGTGGCTTCTGCAAGCTCCGCGAG-3'
<b>Genome walking</b>	
ORAC 231	5'-TACCCAAAGTCTTCTCCAGTGTAACTAGAG-3'
ORAC 230	5'-GAATTTCTTCTGTGCCGTCAACATATCG-3'

30 cycles of 94°C for 1 min, 50–60°C for 1 min, and 72°C for 1 min. A final extension was carried out at 72°C for 7 min. The PCR product was separated on a 1% TBE agarose gel and the dominant 550 bp fragment purified using the QIAquick Gel Extraction Kit (Qiagen). The purified 550 bp product was sequenced at the University of Washington Biochemistry Sequencing Facility using ABI 3730XL high-throughput capillary DNA Analyser (Applied Biosystems, Forest City, CA, USA).

### 5' and 3' *rpoD* RACE

BD SMART RACE cDNA amplification kit (BD Biosciences Clontech, Palo Alto, CA, USA) was used to extend the known *rpoD* sequence. 5' and 3' RACE primers were designed using the sequence generated from the amplification of 550 bp *rpoD* fragment by PCR (see above). Total RNA was used to synthesize 5' RACE cDNA using 5'-CDS primer and BD SMART II A oligo primer and the 3' RACE cDNA using 3'-RACE CDS Primer A according to the manufacturer's instructions. A negative control to verify the absence of genomic DNA was performed by the exclusion of the reverse transcriptase. The 5' RACE cDNA and negative control were amplified with Advantage 2 PCR kit using the universal primer mix (UPM) from the SMART-RACE kit and ORAC 242 primer (Table 4). 3' RACE cDNA and negative control were amplified using UPM and ORAC 240 primers (Table 4). PCR protocol included an initial 3 min denaturation at 94°C, and then 30 cycles at 94°C for 30 sec, 60°C for 30 sec, and 72°C for 1 min. At the end of these cycles an additional 5 min at 72°C was performed to complete DNA synthesis. PCR products were purified from the reaction mixture using the QIAquick PCR purification kit and cloned into pCR-Blunt II-TOPO vector (Invitrogen). Randomly selected clones grown on Luria broth kanamycin (25 µg/mL) plates were checked for inserts using PCR. To screen for inserts, 20 µL of the 5 mL overnight cultures (25 µg/mL kanamycin in Luria broth) were centrifuged at 16,000 × g for 5 min and each resuspended separately in 100 µL of ddH<sub>2</sub>O. For the PCR, 5 µL of the cell suspension was added to a 50 µL volume reaction that contained 1X PCR buffer for KOD Hot Start DNA polymerase (EMD Biosciences, Novagen Brand, Madison, WI, USA), 0.2 mM dNTP, 1 mM MgSO<sub>4</sub>, 20 pmole of M13 forward and reverse primers (Table 4), and 1 U KOD Hot Start DNA polymerase. Plasmid DNA containing inserts was isolated using Qiaprep spin Miniprep columns (Qiagen) according to the manufacturer's directions. The inserts sequences were determined by dye-terminator automated sequencing using M13 forward and reverse primers (SigmaGenosys, The Woodlands, TX, USA) on 5 clones each for both the 5' RACE and 3' RACE.

### Genome walking to obtain the 5' UTR region of the *rpoD* gene

High molecular weight DNA was extracted from *Heterosigma akashiwo* cells grown to a density of  $1.3 \times 10^5$  cells/

mL. Harvested cells were extracted by following an on-line protocol [91]. The cells ( $8.7 \times 10^7$ ) were resuspended in 20 ml of cold lysis buffer (20 mM EDTA, 10 mM TrisCl, pH 8, 1% Triton X, 500 mM Guanidine-HCl, and 200 mM NaCl). The lysate was incubated at 37°C for 1 hour with gentle agitation. The DNA was treated with RNase A (20 µg/mL) for 30 min at 37°C followed by Proteinase K (0.8 mg/mL) for 2 h at 50°C using gentle agitation. To remove cell debris, this lysate was pelleted by centrifugation at 9,750 × g for 20 min and the clear supernatant was removed. Three mL of the lysate were added to each QBT buffer equilibrated Qiagen genomic 100 tip. The columns were washed twice with 10 mL of buffer QC. DNA was eluted from the genomic 100 tip with 5 mL of buffer QF and precipitated by the addition of 0.7 volume room temperature isopropanol. The DNA was pelleted by centrifugation at 9,750 × g for 20 min and the air-dried pellet resuspended in 1 mL of buffer EB (Qiagen). The DNA was stored at 4°C to prevent shearing from freeze thawing. The BD GenomeWalker™ Universal Kit (Clontech) was used to make four libraries by digesting the high molecular weight DNA with *DraI*, *EcoRV*, *PvuII*, and *StuI*, followed by DNA purification, and ligation of genomic DNA to BD GenomeWalker™ adaptors according to the manufacturer's instructions. The Genome Walker libraries were used to amplify DNA upstream of the known *rpoD* sequence by using ORAC 233 primer (Table 4) and AP1 primer provided in the kit by using BD Advantage 2 Polymerase Mix according to BD GenomeWalker™ Universal procedures. A nested PCR was performed using primer ORAC 232 (Table 4) and AP2 primer (kit primer) as suggested in the manual. The PCR products were separated on a 1% TAE gel and the dominant bands were excised and purified using the Qiagen Gel extraction protocol. The purified products were cloned into TOPO TA vector (Invitrogen) and sequenced using M13 forward and reverse primers.

### Domain analysis of *Tsg1* protein and its homologues in a cyanobacterium and in rhodophytes

Similarity searches for *Tsg1* were performed using the NCBI protein-protein BLAST algorithm [92]. Domain architecture analysis was carried out for the four proteins of Figure 1 using the SMART database default parameters. Amino acid sequences were aligned using Clustal W version 1.83 [93] as implemented by GenomeNet Computation Service [94] with default parameters. The domains from the SMART dataset were then used to establish the boundaries for the PAS, HAMP, HisKA and HATPase domains for *Tsg1* in the Clustal W alignment. It was unnecessary to make further refinements since the domains fell into groups and few sequence gaps were present.

### Peptide signalling and stromal targeting domains within *rpoD* transcript

The *rpoD* sequence, and unpublished data generated in our laboratory for fucoxanthin chlorophyll binding protein 2, and phosphoribulose kinase transcripts were submitted to SignalP 3.0 Server [95,96] for identification of the signal peptide. The beginning of the stromal targeting domain was designated by the presence of a conserved phenylalanine residue [62,97,98].

### Trg1 peptide antibody and Western detection of Trg1

An oligopeptide (amino acid sequence-CDEIEEKILLR-LKNENNKEK) unique to the *Heterosigma akashiwo* Trg1 sensor kinase was synthesized based on a hydrophilic region in the amino terminal end of the Trg1 protein. Synthesis and purification (85%) of this oligopeptide was done at Lampire Biologicals (Pipersville, PA, USA). The synthesized peptide was coupled to the carrier protein keyhole limpet hemocyanin (KLH) through the cysteine residue at the N-terminus of the peptide. Polyclonal antibodies (IgG) were directed against Trg1 peptide in three rabbits; each rabbit injected with 0.5 mg of purified peptide in Freund's complete adjuvant. Pre-immune antisera, collected prior to the initial injection, and total purified antisera (day 50), were both supplied by Lampire Biologicals and stored at -80°C until use.

Western blots were performed using X-Cell Sure Lock Mini Cell and NuPAGE MOPS SDS buffer kit (Invitrogen) according to manufacturer's instructions. *Heterosigma akashiwo* cells ( $1.6 \times 10^6$  cells) were harvested at a density of  $1.3 \times 10^5$  cells/mL at L3 in a 12 h light:12 h dark period. The cells were pelleted at  $3,800 \times g$  for 10 min and the supernatant decanted. The pelleted cells were resuspended in 200  $\mu$ L of NuPAGE lysis buffer and heated to 70°C for 10 min. The supernatant was clarified by centrifuging the lysed cells at  $10,200 \times g$  for 5 min at room temperature. The clarified protein samples (10  $\mu$ L) and 2  $\mu$ L of "Magic Mark Western" protein standard (Invitrogen) were separated on a NuPAGE Novex Bis-Tris gel (Invitrogen) at 180 V for 1 hr at 20°C. Proteins were electrophoretically transferred to BioRAD Transblot nitrocellulose membrane (0.45  $\mu$ m, BioRAD Laboratories, Hercules, CA, USA) in NuPAGE transfer buffer with the X-Cell Blot Module for 1 hr at 30 V at 20°C. The blot was blocked at 4°C overnight using blocking buffer (1X PBS, pH 7.0, 3% non-fat dried milk, and 0.05% Tween-20). A preadsorption control was carried out to establish the specificity of the polyclonal antibodies to Trg1 by exposing the post-bleed serum (10  $\mu$ L) to 2 mg of Trg1 oligopeptide overnight at 4°C. After the overnight incubation, the preadsorption control was centrifuged at  $16,000 \times g$  for 20 min and the supernatant removed. Blots were probed with prebleed serum, postbleed serum, and the preadsorption control at a dilution of 1:500 for 1 hr at room temperature. The membranes were treated in wash buffer

(10 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.05% Tween-20) with four buffer changes of 45 min each. The Super Signal West Pico anti-rabbit IgG detection kit (Pierce, Rockford, IL, USA) was used to detect Trg1 protein. The blots were gently agitated at room temperature while submerged in the diluted secondary antibody (1:50,000). Four washes (45 min each) in wash buffer were used to remove the unbound secondary antibody. After washing, 3 mL of the working solution from the Pierce kit (equal parts of stable peroxide solution and luminal/enhancer) were combined and placed on the membranes. The blots were placed on CL-Xposure film (Pierce) and developed in Fischer Model K-Plus Automatic X-Ray film processor (Fischer Industries Inc., Geneva, IL, USA) after 4 min.

### Trg1 protein modelling

Three known structures were used to model *Heterosigma akashiwo* Trg1: *Thermotoga maritima* OmpR (1 KGS), *Escherichia coli* PhoB (1QQI), *E. coli* OmpR (1OPC) [30-33]. Structures were obtained from the Protein Data Bank [99,100]. Modelling was carried out with tools available in the suite of programs in the "Molecular Operating Environment" from the Chemical Computing Group, Montreal [101].

### Abbreviations

*tsg1* – transcriptional sensor gene 1; *trg1* – transcriptional regulator gene 1; *rpoD* – RNA polymerase D subunit; CCMP – Provasoli-Guillard National Center for Culture of Marine Phytoplankton

### Authors' contributions

MRD designed the Trg1 antibody and carried out Western analyses, performed RACE and genome walking experiments for *fcp2* and *rpoD*, performed domain analysis and alignment of Tsg1 and its homologues, analysed signal peptides and stromal targeting domains for RpoD, Fcp2, and PRK, and contributed to the manuscript (text and figures). KGK constructed RNA polymerase alignments, performed phylogenetic analyses, and contributed to the manuscript (text and figures). ETA carried out modelling of the Trg1 protein and contributed to the final manuscript (text and figures). LYSC amplified and sequenced the 550 bp RNA polymerase gene fragment and assisted in the construction of the final *rpoD* sequence. MAJ sequenced *tsg1*, participated in modelling of the Trg1 protein and contributed to the development of the manuscript. RAC conceived of the study, participated in its design, performed bioinformatic analyses of the His-to-Asp proteins, and contributed to the development of the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

RAC dedicates this paper to the memory of Dr. Lee McIntosh, friend and former student. We thank Dr. William H. Hatheway, Dr. Han Chuan Ong and Michael B. Lakeman for their helpful comments on the manuscript. We also thank Sandy A. Simon for his contribution to the RACE experiments, Kun-Lin Lee for his assistance with Western analysis as well as Rachael Hall for sequencing the *prk* gene. This research was supported by an NSF Microbial Post-Doctoral Fellowship to MRD, PHS National Research Grant (T32 GM07270) to MAJ, and grant number NA16RG1044, project RB37/RB42, from the National Oceanic and Atmospheric Administration to Washington Sea Grant Program, University of Washington to RAC.

## References

- Appleby JL, Parkinson JS, Bourret RB: **Signal transduction via the multi-step phosphorelay: Not necessarily a road less traveled.** *Cell* 1996, **86**:845-848.
- Thomason P, Kay R: **Eukaryotic signal transduction via histidine-aspartate phosphorelay.** *J Cell Sci* 2000, **113**:3141-3150.
- Bourret RB, Borkovich KA, Simon MI: **Signal transduction pathways involving protein-phosphorylation in prokaryotes.** *Annu Rev Biochem* 1991, **60**:401-441.
- Paget MSB, Helmann JD: **The sigma 70 family of sigma factors.** *Genome Biol* 2003, **4**:203.
- Dove SL, Darst SA, Hochschild A: **Region 4 of sigma as a target for transcription regulation.** *Mol Microbiol* 2003, **48**:863-874.
- Makino K, Shinagawa H, Amemura M, Nakata A: **Nucleotide-sequence of the *phoB* gene, the positive regulatory gene for the phosphate regulon of Escherichia coli K-12.** *J Mol Biol* 1986, **190**:37-44.
- Ninfa AJ, Magasanik B: **Covalent modification of the *glnG* product, NRI, by the *glnL* product, NRII, regulates the transcription of the *glnALG* operon in Escherichia coli.** *Proc Natl Acad Sci U S A* 1986, **83**:5909-5913.
- Wurtzel ET, Chou MY, Inouye M: **Osmoregulation of gene-expression I. DNA sequence of the *ompR* gene of the *OmpB* operon of Escherichia coli and characterization of its gene-product.** *J Biol Chem* 1982, **257**:13685-13691.
- Stock AM, Robinson VL, Goudreau PN: **Two-component signal transduction.** *Annu Rev Biochem* 2000, **69**:183-215.
- Galperin MY: **Bacterial signal transduction network in a genomic perspective.** *Environ Microbiol* 2004, **6**:552-567.
- Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FSL, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrook-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock REW, Lory S, Olson MV: **Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.** *Nature* 2000, **406**:959-964.
- Mizuno T, Kaneko T, Tabata S: **Compilation of all genes encoding bacterial two-component signal transducers in the genome of the cyanobacterium, Synechocystis sp. strain PCC 6803.** *DNA Res* 1996, **3**:407-414.
- Wang L, Sun YP, Chen WL, Li JH, Zhang CC: **Genomic analysis of protein kinases, protein phosphatases and two-component regulatory systems of the cyanobacterium Anabaena sp. strain PCC 7120.** *FEMS Microbiol Lett* 2002, **217**:155-165.
- Gibbs SP: **The Chloroplasts of Some Algal Groups May Have Evolved from Endosymbiotic Eukaryotic Algae; New York, N. Y..** Volume 361 (FEB). Edited by: Fredrick JF. New York Academy of Sciences; 1981:193-208.
- Margulis L: **Symbiosis in Cell Evolution: Life and Its Environment on the Early Earth.** 1st Edition edition. Freeman, San Francisco, ; 1981.
- Cattolico RA, Boothroyd JC, Gibbs SP: **Synchronous growth and plastid replication in naturally wall-less alga Olisthodiscus luteus.** *Plant Physiol* 1976, **57**:497-503.
- Stiller JW, Reel DC, Johnson JC: **A single origin of plastids revisited: convergent evolution in organellar genome content.** *J Phycol* 2003, **39**:95-105.
- Doran E, Cattolico RA: **Photoregulation of chloroplast gene transcription in the chromophytic alga Heterosigma carterae.** *Plant Physiol* 1997, **115**:773-781.
- Reynolds AE, McConaughy BL, Cattolico RA: **Chloroplast genes of the marine alga Heterosigma carterae are transcriptionally regulated during a light/dark cycle.** *Mol Mar Biol Biotechnol* 1993, **2**:121-128.
- Jacobs MA, Connell L, Cattolico RA: **A conserved His-Asp signal response regulator-like gene in Heterosigma akashiwo chloroplasts.** *Plant Mol Biol* 1999, **41**:645-655.
- The Stramenopile Chloroplast Genome Project** [[http://chloplast.ocean.washington.edu/forum/website\\_bug](http://chloplast.ocean.washington.edu/forum/website_bug)]
- Galperin MY, Nikolskaya AN, Koonin EV: **Novel domains of the prokaryotic two-component signal transduction systems.** *FEMS Microbiol Lett* 2001, **203**:11-21.
- Schultz J, Milpetz F, Bork P, Ponting CP: **SMART, a simple modular architecture research tool: Identification of signaling domains.** *Proc Natl Acad Sci U S A* 1998, **95**:5857-5864.
- SMART: Simple Modular Architecture Research Tool** [<http://smart.embl-heidelberg.de/>]
- Kim DJ, Forst S: **Genomic analysis of the histidine kinase family in bacteria and archaea.** *Microbiology* 2001, **147**:1197-1212.
- Parkinson JS, Kofoid EC: **Communication modules in bacterial signaling proteins.** *Annu Rev Genet* 1992, **26**:71-112.
- Mikami K, Kanesaki Y, Suzuki I, Murata N: **The histidine kinase Hik33 perceives osmotic stress and cold stress in Synechocystis sp. PCC 6803.** *Mol Microbiol* 2002, **46**:905-915.
- van Waasbergen LG, Dolganov N, Grossman AR: **nbIS, a gene involved in controlling photosynthesis-related gene expression during high light and nutrient stress in Synechococcus elongatus PCC 7942.** *J Bacteriol* 2002, **184**:2481-2490.
- Martinez-Hackert E, Stock AM: **Structural relationships in the OmpR family of winged-helix transcription factors.** *J Mol Biol* 1997, **269**:301-312.
- Okamura H, Hanaoka S, Nagadoi A, Makino K, Nishimura Y: **Structural comparison of the PhoB and OmpR DNA-binding/transactivation domains and the arrangement of PhoB molecules on the phosphate box.** *J Mol Biol* 2000, **295**:1225-1236.
- Itou H, Tanaka I: **The OmpR-family of proteins: insight into the tertiary structure and functions of two-component regulator proteins.** *J Biochem (Tokyo)* 2001, **129**:343-350.
- Martinez-Hackert E, Stock AM: **The DNA-binding domain of OmpR: crystal structure of a winged helix transcription factor.** *Structure* 1997, **5**:109-124.
- Buckler DR, Zhou YC, Stock AM: **Evidence of intradomain and interdomain flexibility in an OmpR/PhoB homolog from Thermotoga maritima.** *Structure* 2002, **10**:153-164.
- Gruber TM, Gross CA: **Multiple sigma subunits and the partitioning of bacterial transcription space.** *Annu Rev Microbiol* 2003, **57**:441-466.
- Helmann JD, Chamberlin MJ: **Structure and function of bacterial sigma factors.** *Annu Rev Biochem* 1988, **57**:839-872.
- Dove SL, Hochschild A: **Bacterial two-hybrid analysis of interactions between region 4 of the sigma(70) subunit of RNA polymerase and the transcriptional regulators Rsd from Escherichia coli and AlgQ from Pseudomonas aeruginosa.** *J Bacteriol* 2001, **183**:6413-6421.
- Gruber TM, Bryant DA: **Molecular systematic studies of eubacteria, using sigma(70)-type sigma factors of group I and group 2.** *J Bacteriol* 1997, **179**:1734-1747.
- Huckauf J, Nomura C, Forchhammer K, Hagemann M: **Stress responses of Synechocystis sp. strain PCC 6803 mutants impaired in genes encoding putative alternative sigma factors.** *Microbiol* 2000, **146**:2877-2889.
- Imamura S, Yoshihara S, Nakano S, Shiozaki N, Yamada A, Tanaka K, Takahashi H, Asayama M, Shirai M: **Purification, characterization, and gene expression of all sigma factors of RNA polymerase in a cyanobacterium.** *J Mol Biol* 2003, **325**:857-872.
- Iyer LM, Koonin EV, Aravind L: **Evolution of bacterial RNA polymerase: implications for large-scale bacterial phylogeny, domain accretion, and horizontal gene transfer.** *Gene* 2004, **335**:73-88.
- Hara K, Morita M, Takahashi R, Sugita M, Kato S, Aoki S: **Characterization of two genes, sigI and sig2, encoding distinct plastid sigma factors(I) in the moss Physcomitrella patens: phylogenetic relationships to plastid sigma factors in higher plants.** *FEBS Lett* 2001, **499**:87-91.
- Ichikawa K, Sugita M, Imaizumi T, Wada M, Aoki S: **Differential expression on a daily basis of plastid sigma factor genes from**

- the moss *Physcomitrella patens*. Regulatory interactions among PpSig5, the circadian clock, and blue light signaling mediated by cryptochromes. *Plant Physiol* 2004, **136**:4285-4298.
43. The Joint Genome Institute Eukaryotic Genomics [<http://genome.jgi-psf.org/>]
  44. Doyle JJ: Gene trees and species trees - molecular systematics as one-character taxonomy. *Syst Bot* 1992, **17**:144-163.
  45. Maddison WP: Gene trees in species trees. *Syst Biol* 1997, **46**:523-536.
  46. Felsenstein J: Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 1978, **27**:401-410.
  47. Foster PG: Modeling compositional heterogeneity. *Syst Biol* 2004, **53**:485-495.
  48. Whatley JM: Chloroplast evolution - ancient and modern; New York, N. Y.. Volume 361 (FEB). Edited by: Fredrick JF. New York Academy of Sciences; 1981:154-165.
  49. Mizuno T: Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Res* 1997, **4**:161-168.
  50. Koretke KK, Lupas AN, Warren PV, Rosenberg M, Brown JR: Evolution of two-component signal transduction. *Mol Biol Evol* 2000, **17**:1956-1970.
  51. Kessler U, Maid U, Zetsche K: An equivalent to bacterial ompR genes is encoded on the plastid genome of red algae. *Plant Mol Biol* 1992, **18**:777-780.
  52. Bernard C, de Marsac NT, Thomas JC: A ompR gene in the plastid genome of *Rhodella violacea*. *Plant Physiol* 1994, **106**:795-796.
  53. Reith M, Munholland J: Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. *Plant Mol Biol Rep* 1995, **13**:333-335.
  54. Glöckner G, Rosenthal A, Valentin K: The structure and gene repertoire of an ancient red algal plastid genome. *J Mol Evol* 2000, **51**:382-390.
  55. Ohta N, Matsuzaki M, Misumi O, Miyagishima S, Nozaki H, Tanaka K, Shin-I T, Kohara Y, Kuroiwa T: Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. *DNA Res* 2003, **10**:67-77.
  56. Hagopian JC, Reis M, Kitajima JP, Bhattacharya D, de Oliveira MC: Comparative analysis of the complete plastid genome sequence of the red alga *Gracilaria tenuistipitata* var. *liui* provides insights into the evolution of rhodoplasts and their relationship to other plastids. *J Mol Evol* 2004, **59**:464-477.
  57. Stirewalt VL, Michalowski CB, Löffelhardt VV, Bohnert HJ, Bryant DA: Nucleotide sequence of the cyanelle genome from *Cyanophora paradoxa*. *Plant Mol Biol Rep* 1995, **13**:327-332.
  58. Puerta MVS, Bachvaroff TR, Delwiche CF: The complete plastid genome sequence of the haptophyte *Emiliania huxleyi*: a comparison to other plastid genomes. *DNA Res* 2005, **12**:151-156.
  59. Douglas SE, Penny SL: The plastid genome of the cryptophyte alga, *Guillardia theta*: Complete sequence and conserved syntenic groups confirm its common ancestry with red algae. *J Mol Evol* 1999, **48**:236-244.
  60. Lemieux C, Otis C, Turmel M: A clade uniting the green algae *Mesostigma viride* and *Chlorokybus atmophyticus* represents the deepest branch of the Streptophyta in chloroplast genome-based phylogenies. *BMC Biology* 2007, **5**.
  61. Kowallik KV, Stoebe B, Schaffran I, KrothPancic P, Freier U: The chloroplast genome of a chlorophyll a+c-containing alga, *Odontella sinensis*. *Plant Mol Biol Rep* 1995, **13**:336-342.
  62. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou SG, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kroger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS: The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 2004, **306**:79-86.
  63. Urao T, Yamaguchi-Shinozaki K, Shinozaki K: Two-component systems in plant signal transduction. *Trends Plant Sci* 2000, **5**:67-74.
  64. Kiba T, Taniguchi M, Imamura A, Ueguchi C, Mizuno T, Sugiyama T: Differential expression of genes for response regulators in response to cytokinins and nitrate in *Arabidopsis thaliana*. *Plant Cell Physiol* 1999, **40**:767-771.
  65. Brandstatter I, Kieber JJ: Two genes with similarity to bacterial response regulators are rapidly and specifically induced by cytokinin in *Arabidopsis*. *Plant Cell* 1998, **10**:1009-1019.
  66. Wilkinson JQ, Lanahan MB, Yen HC, Giovannoni JJ, Klee HJ: An ethylene-inducible component of signal-transduction encoded by never-ripe. *Science* 1995, **270**:1807-1809.
  67. Wurgler-Murphy SM, Saito H: Two-component signal transducers and MAPK cascades. *Trends Biochem Sci* 1997, **22**:172-176.
  68. Marin K, Suzuki L, Yamaguchi K, Ribbeck K, Yamamoto H, Kanesaki Y, Hagemann M, Murata N: Identification of histidine kinases that act as sensors in the perception of salt stress in *Synechocystis* sp. PCC 6803. *Proc Natl Acad Sci U S A* 2003, **100**:9061-9066.
  69. Suzuki I, Kanesaki Y, Mikami K, Kanehisa M, Murata N: Cold-regulated genes under control of the cold sensor Hik33 in *Synechocystis*. *Mol Microbiol* 2001, **40**:235-244.
  70. Taylor BL, Zhulin IB: PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev* 1999, **63**:479-+.
  71. Allen JF: The function of genomes in bioenergetic organelles. *Philos Trans R Soc Lond Ser B-Biol Sci* 2003, **358**:19-37.
  72. Allen JF, Nilsson A: Redox signalling and the structural basis of regulation of photosynthesis by protein phosphorylation. *Physiol Plant* 1997, **100**:863-868.
  73. Hefti MH, Francoijs KJ, de Vries SC, Dixon R, Vervoort J: The PAS fold - A redefinition of the PAS domain based upon structural prediction. *Eur J Biochem* 2004, **271**:1198-1208.
  74. Schobert M, Gorisch H: A soluble two-component regulatory system controls expression of quinoprotein ethanol dehydrogenase (QEDH) but not expression of cytochrome c(550) of the ethanol-oxidation system in *Pseudomonas aeruginosa*. *Microbiol* 2001, **147**:363-372.
  75. Hutchings MI, Hoskisson PA, Chandra G, Buttner MJ: Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3(2). *Microbiol* 2004, **150**:2795-2806.
  76. Ashby MK, Houmar J, Mullineaux CW: The ycf27 genes from cyanobacteria and eukaryotic algae: distribution and implications for chloroplast evolution. *FEMS Microbiol Lett* 2002, **214**:25-30.
  77. Ashby MK, Mullineaux CW: Cyanobacterial ycf27 gene products regulate energy transfer from phycobilisomes to photosystems I and II. *FEMS Microbiol Lett* 1999, **181**:253-260.
  78. Makino K, Amemura M, Kim SK, Nakata A, Shinagawa H: Role of the sigma-70 subunit of RNA polymerase in transcriptional activation by activator protein Phob in *Escherichia coli*. *Genes and Development* 1993, **7**:149-160.
  79. Kondo H, Nakagawa A, Nishihira J, Nishimura Y, Mizuno T, Tanaka I: *Escherichia coli* positive regulator OmpR has a large loop structure at the putative RNA polymerase interaction site. *Nat Struct Biol* 1997, **4**:28-31.
  80. Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu XN, Reith M, Cavalier-Smith T, Maier UG: The highly reduced genome of an enslaved algal nucleus. *Nature* 2001, **410**:1091-1096.
  81. Troxler RF, Zhang F, Hu J, Bogorad L: Evidence that sigma-factors are components of chloroplast RNA-Polymerase. *Plant Physiol* 1994, **104**:753-759.
  82. Liu B, Troxler RF: Molecular characterization of a positively photoregulated nuclear gene for a chloroplast RNA polymerase sigma factor in *Cyanidium caldarium*. *Proc Natl Acad Sci U S A* 1996, **93**:3313-3318.
  83. Tanaka K, Oikawa K, Ohta N, Kuroiwa H, Kuroiwa T, Takahashi H: Nuclear encoding of a chloroplast RNA polymerase sigma subunit in a red alga. *Science* 1996, **272**:1932-1935.
  84. Hajdukiewicz PJ, Allison LA, Maliga P: The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J* 1997, **16**:4041-4048.
  85. Link G: Green life: control of chloroplast gene transcription. *BioEssays* 1996, **18**:465-471.
  86. Rajasekhar VK, Sun E, Meeker R, Wu BW, Tewari KK: Highly purified pea chloroplast RNA-polymerase transcribes both ribosomal-RNA and messenger-RNA genes. *Eur J Biochem* 1991, **195**:215-228.
  87. Hess WR, Prombona A, Fieder B, Subramanian AR, Börner T: Chloroplast rps15 and the rpoB/C1/C2 gene-cluster are strongly

- transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA-polymerase. *EMBO J* 1993, **12**:563-571.
88. Choquet Y, Vallon O: **Synthesis, assembly and degradation of thylakoid membrane proteins.** *Biochimie (Paris)* 2000, **82**:615-634.
  89. McIntosh L, Cattolico RA: **Preservation of algal and higher plant ribosomal-RNA integrity during extraction and electrophoretic quantitation.** *Anal Biochem* 1978, **91**:600-612.
  90. Sambrook J, Fritsch EF, Maniatis T: **Molecular Cloning: a Laboratory Handbook.** 2nd edition. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press; 1989.
  91. **User-Developed Protocol: Isolation of genomic DNA** [<http://www.qiagen.com/literature/protocols/pdf/QG08.pdf>]
  92. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.
  93. Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.
  94. **Multiple Sequence Alignment by CLUSTALW** [<http://align.genome.jp/>]
  95. **SignalP 3.0 Server** [<http://www.cbs.dtu.dk/services/SignalP/>]
  96. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340**:783-795.
  97. Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG: **In vivo characterization of diatom multipartite plastid targeting signals.** *Cell Science* 2002, **115**:4061-4069.
  98. Chaal BK, Green BR: **Protein import pathways in 'complex' chloroplasts derived from secondary endosymbiosis involving a red algal ancestor.** *Plant Mol Biol* 2005, **57**:333-342.
  99. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardocki C: **The Protein Data Bank.** *Acta Crystallogr Sect D Biol Crystallogr* 2002, **D58**:899-907.
  100. **RCSB Protein Data Bank** [<http://www.rcsb.org/pdb/home/home.do>]
  101. **Chemical Computing Group** [<http://www.chemcomp.com>]
  102. **Cyanidioschyzon merolae Genome Project** [<http://merolaebiol.s.u-tokyo.ac.jp/>]

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

