

**Current state of the data in the cyanobacterium *Synechococcus* prepared for the**

DOE Genomes To Life project  
**CARBON SEQUESTRATION IN *SYNECHOCOCCUS SP.*:  
FROM MOLECULAR MACHINES TO HIERARCHICAL MODELING**

**Jeff Blanchard, NCGR Santa Fe, NM 87505 [jlb@ncgr.org](mailto:jlb@ncgr.org)**

**Report Objective:** To evaluate the current state of the data and literature for the marine cyanobacterium *Synechococcus* WH8102. This overview will be used for developing a comprehensive hierarchical systems model to utilize information at different organizational scales, ranging from gene mutation and expression to metabolic pathways and external environmental response.

**Summary:** *Synechococcus* is a diverse, polyphyletic taxonomic category, thus little functional or evolutionary relatedness can be inferred simply from the commonality of its name. Despite its ubiquitous oceanographic presence, it is also a poorly studied: there are no published experimental molecular genetic or biochemistry papers on the *Synechococcus/Prochlorococcus* group. Genomically, there is a draft sequence of *Synechococcus* WH8102 at [http://genome.ornl.gov/microbial/syn\\_wh/1/syn\\_wh\\_1.html](http://genome.ornl.gov/microbial/syn_wh/1/syn_wh_1.html). Annotation includes cross-linking to external databases and functional assignment, though not the assignment of gene names. Metabolically, databases such as KEGG and WIT do have pathway information on cyanobacteria, but not specifically on *Synechococcus*. Similarly, other information is also sparse, though there have been for example, two published microarray studies on *Synechocystis* (Hihara *et al.* 2001 and Suzuki *et al.* 2002). There is no published, direct experimental data on carbon sequestration in *Synechococcus* WH8102 or on the marine *Synechococcus/Prochlorococcus* group.

Analysis of the sequence data of the closely related taxa *Prochlorococcus sp.* did not identify any genes with homology to known inorganic carbon transporters (Hess *et al.* 2001). The *Prochlorococcus* strains are also missing *ndh* genes typically associated with carbon uptake in other cyanobacteria and with carbonic anhydrase. Furthermore *rbcL*, *rbcS* and other genes associated with the carboxysome are very different from the other cyanobacterial carboxysome genes. The data on *Prochlorococcus* is likely to be indicative of what we will find in *Synechococcus*: already it appears that *Synechococcus* WH8201 is missing a key RUBISCO transcription factor *ccbR*, and has a novel regulatory protein, *cbbx*, that has not been previously found in cyanobacteria. Thus the examination of the current state of the data leads one to conclude that assumptions on the operations of the relevant carbon sequestering molecular machines are likely to be dubious if based solely on comparative approaches with model cyanobacterial systems. Preliminary analysis of the data leads us to hypothesize a role of horizontal gene transfer in the evolution of *Synechococcus* carbon sequestration, and thus a physiological and ecological overview of the diversity of the process and the organism will be important.

**Taxonomy:** It is important to keep in mind that the name *Synechococcus* has little taxonomic or evolutionary value. There are at least 5 major distinct *Synechococcus* groups that are distributed all over the cyanobacterial rRNA trees. Thus although *Synechococcus* PCC7942 and PCC7002 are used as model systems, *Synechococcus* WH8102 and others in this marine *Synechococcus* group are more closely related to *Prochlorococcus* (Hess *et al.* 2001).

**Summary of Published Literature:** Most of the published reports are related to taxonomy and ecological characteristics (photosynthetic pigments, light-dependent physiology, herbivory, phage

susceptibility). Excellent recent reviews are available on the physiology and molecular ecology of the *Synechococcus/Prochlorococcus* group (Scanlan and West 2002, Partensky 1999). The molecular genetics and biochemistry of this group is just beginning to be explored. Brahamsha (1996) developed a transformation system for *Synechococcus* WH8102 that has been applied to identify genes involved in non-flagellar swimming towards nitrogen sources and genes essentially to the utilizing of urea as a nitrogen source (Collier et al. 1999). However, these few papers are all that we know from direct experimentation on *Synechococcus* WH8102. We can and should transfer knowledge from the model cyanobacterial systems, but as discussed below we should use caution as we expect there to be major differences in inorganic carbon assimilation pathways and their regulation.

**Genome Sequencing Efforts:** Two institutes, JGI and the Kazusa Research Institute have generated the six cyanobacterial genomes that have been published in print or on a web site (see list below). The published genomes have been annotated by a combination of automated methods with differing degrees of expert knowledge added manually. For the most part they are linked to functional categories and KEGG pathways. Three other genomes are in the latter draft stages at JGI and Kazusa. The JGI is also producing shotgun coverage of *Synechococcus* sp. PCC 7942. I did not receive any reply to inquiries regarding the state or availability of data from any of the other institutes reported to be sequencing a cyanobacteria genome.

**Cyanobacterial Genome Sequencing Progress (sources – JGI, Cyanobase, GOLD, TIGR, NCBI):**

- *Anabaena/Nostoc* sp. PC7002 – Kazusa - published
- *Gleobacter violaceus* PC 7421 – Kazusa (in finishing phase - due 2002)
- *Microcystis aeruginosa* PCC 7806 – Institut Pasteur (incomplete)
- *Nostoc punctiforme* - JGI (draft)
- *Prochlorococcus marinus* MED4 - JGI
- *Prochlorococcus marinus* MIT9313 - JGI
- *Prochlorococcus marinus* SS120 – Genoscope (incomplete)
- *Spirulina platensis*- Human Genome Center, Beijing (incomplete)
- *Synechococcus* sp. PCC 6301 Nagoya Univ incomplete
- *Synechococcus* sp. PCC 7002 Beijing Univ, Penn State Univ (incomplete)
- *Synechococcus* sp. PCC 7942 Texas A&M Univ (incomplete) and JGI
- *Synechococcus* sp. WH8102 - JGI
- *Synechocystis* sp. PCC6813 – Kazusa - published
- *Thermosynechococcus elongatus* BP-1 – Kazusa - published
- *Trichodesmium erythraeum* – JGI (draft)

**General state of *Synechococcus* WH8102 Annotation:** Frank Larimer at ORNL has done a draft analysis ([http://genome.ornl.gov/microbial/syn\\_wh/1/syn\\_wh\\_1.html](http://genome.ornl.gov/microbial/syn_wh/1/syn_wh_1.html)). The predicted genes and proteins have been linked with many of the KEGG reference maps. From the CONTIG link you can get to summary pages for a gene that include PFAM, InterPro, NR-Blast, COG, and KEGG hits, if you know the coordinates. These can be obtained from a text search of the tab files or a blast search. It appears that gene names are not assigned yet and the best blast hit is used as a place holder. I am not sure how manual curation is being done, but they have recently produced Artemis and ACT ready files. Artemis is a Java genome viewer and annotation tool used by Sanger for microbial genome annotation. ACT (Artemis Comparison Tool) is a DNA sequence comparison viewer based on Artemis. There seems to be files with the assembled sequence and predicted proteins at NCBI, although there is no annotation associated with the files (*i.e.* all

protein names are hypothetical). The sequence related files seem to be similar for the two *Prochlorococcus* genomes.

**Metabolic Information:** The established metabolic pathways databases, KEGG, WIT, and BIOCYC all have past and seemingly on going efforts related to cyanobacteria. KEGG in collaboration with Kazusa, has put together several metabolic maps of key cyanobacterial processes like photosynthesis and have a great collection of metabolic maps on which the genome annotation can be overlaid. These maps are not specific to a particular organisms and thus should be used a starting points to evaluate the annotation. The JGI genomes have not yet been included on the public KEGG site, but JGI provides links to reference maps from the JGI draft analyses. JGI/ORNL has also provided EC numbers with the gene annotation for *Synechococcus* WH8102. Argonne National Lab's WIT2 contains the *Synechocystis* genome and is useful for exploring bacterial gene order. Unfortunately no genomes have been added recently to the public site including the other five sequenced cyanobacteria. The ex-ANL folks at Integrated Genomics are probably keeping this up to date, but there is now a charge for using the Integrated Genomics database and tools. No cyanobacterial organisms are listed in the current version of BIOCYC, but a group at Stanford is using BIOCYC to curate *Synechocystis* PCC6308 (<http://aracyc.stanford.edu/~jshrager/lab/PSBCyc/index.html>). However this link has not been functional this past month. Currently these databases are mostly useful as information resources as a starting point for collecting data for modeling. Since chloroplasts are derived from the cyanobacteria, it should be useful to use the chloroplast pathways present in PathDB and Aracyc to help provide a first pass estimation of cyanobacterial functions. The supplementary tables in Martin *et al.* 2002 (also see commentary by Palenik 2002) are a good starting point for mapping genes between cyanobacteria and *Arabidopsis*. However, much work will need to be done to create pathways for doing computation on networks.

**Other Types of Interaction and Association Data:** Charles Delisi's group at BU has put together a database, Predictome, that contains both experimental (yeast two-hybrid, immunoprecipitation, correlated expression) and computational (gene fusion, chromosomal proximity, gene co-evolution). Unfortunately there is very little of this type of experimental data available for cyanobacteria, although we will be generating quite a bit as part of our experimental effort. The database does contain predictions from several computational methods for *Synechocystis* PCC6308. There is no electronic source of signaling data available for bacteria and databases of regulatory information like Regulon do not contain any cyanobacterial information.

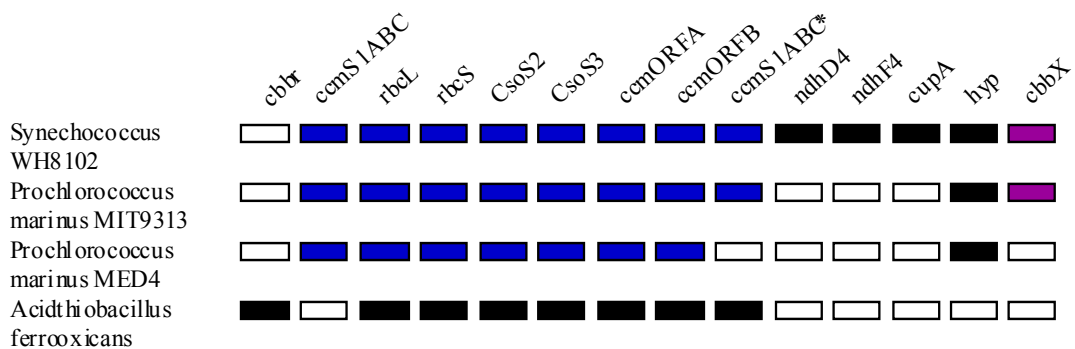
**Microarray studies:** There have been two published microarray studies of cyanobacteria, both on *Synechocystis* (Hihara *et al.* 2001 and Suzuki *et al.* 2002). Hihara *et al.* 2001 noted that genes associated with CO<sub>2</sub> fixation were up regulated during acclimation from low to high light intensity. Arthur Grossman's lab has a web site with some unpublished *Synechocystis* microarray data (<http://aracyc.stanford.edu/~jshrager/lab/cyanoarray/tuchip>). It seems likely that much more data will be cropping up in the coming years.

**Preliminary sequence analysis of mechanisms involved in carbon sequestration:** Since we have no direct experimental evidence on carbon sequestration in *Synechococcus* WH8102 or on the marine *Synechococcus/Prochlorococcus* group, most of what we know at present is inferred from other cyanobacteria or derived from analysis of the genome sequence data. Hess *et al.* 2001 analyzed the sequence data in the two sequenced *Prochlorococcus* *sp.* in the context of the assimilation of inorganic carbon. Quite surprisingly there were not able to identify any genes with homology to known inorganic carbon transporters. The *Prochlorococcus* strains are also missing *ndh* genes typically associated with carbon uptake in other cyanobacteria and with carbonic anhydrase. Furthermore *rbcL*, *rbcS* and other genes associated with the carboxysome

are very different from the other cyanobacterial carboxysome genes and are probably derived from horizontal transfer from a proteobacterium. There are also several other genes involved in the Calvin Cycle that do not appear to phylogenetically congruent with the rest of the cyanobacteria.

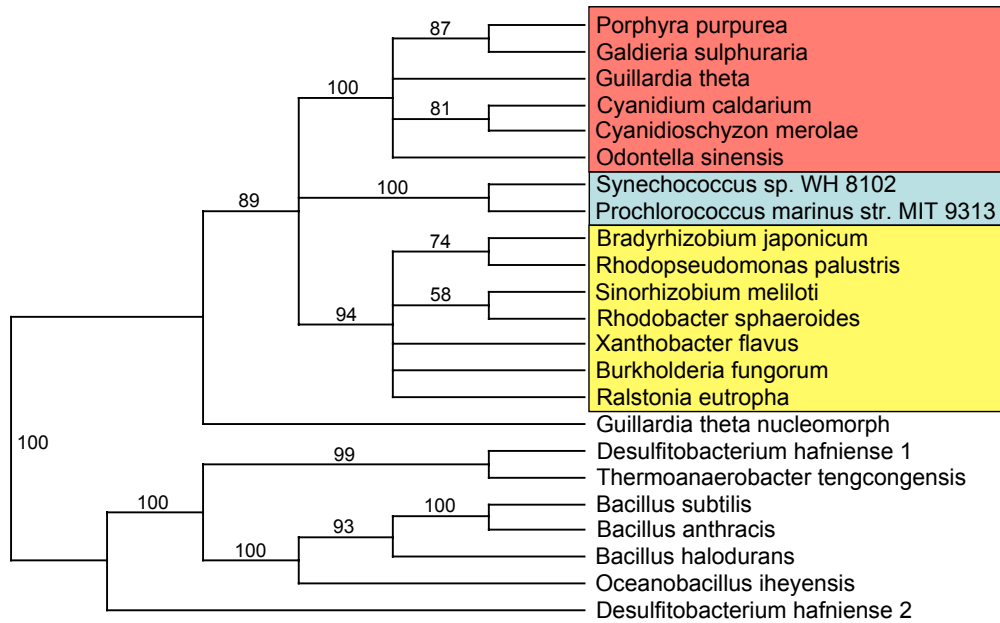
The horizontal gene transfer (HGT) event from which the odd carboxysome genes in *Prochlorococcus* are derived appears to have occurred before the divergence of the marine *Synechococcus* and *Prochlorococcus* based on phylogenetic analysis (please contact the author for preliminary tree). In *Synechococcus* WH8201 there are genes similar to carbonic anhydrase and at least one of the *ndh* gene subcomplexes (*ndhD4-ndhF4-cupB*) is present. (Shibata *et al.* 2001) However, another carbon concentrating system containing *ndhD3-ndhF3-cupA* is missing.

In preliminary analysis of the sequence data we identified a regulator of RUBISCO, *cbbx*, that has not been previously identified in cyanobacteria. Mutations in *cbbx* resulted in impaired photoautotrophic growth (Gibson *et al.* 1997) in an alpha proteobacteria. The reported distribution to date has been limited to alpha proteobacteria and non-green algae (*i.e.* red and brown algae) (Maier *et al.* 2000). In *Prochlorococcus* MIT9313 *cbbx* is located downstream of the carboxysome shell proteins in a similar position as found in *Synechococcus* WH8102 (Figure

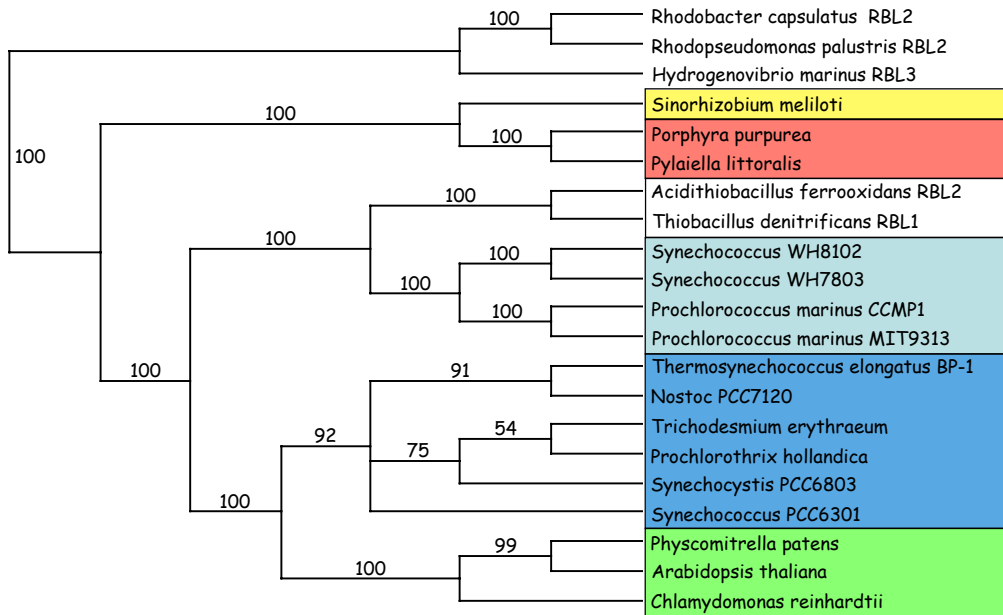


**Figure 1 – Structure of the carboxysome gene array.** Filled boxes indicate shared genes. Blue and purple boxes indicate separate horizontal gene transfer events. *ccmS1ABC* are three separate proteins in bacteria that shared considerable sequence identity. The asterisk indicates that only the second half of the *Synechococcus* WH8102 and *Prochlorococcus* MIT9313 are similar to three *ccmS1* proteins in bacteria. The hypothetical protein (*hyp*) does not show similar to any other known protein outside of the marine *Synechococcus/Prochlorococcus* group. Empty boxes in *Synechococcus* and *Prochlorococcus* indicate they are missing from this cluster and from the genome assembly.

1). Interestingly, *cbbx* is missing in *Prochlorococcus* MED4, which has a very reduced genome size. We then did a search of Genbank for genes similar to *cbbx* and identified candidates in *Nostoc* sp. 7012, but not in other bacteria. The gene is *Nostdoc* is not chromosomally linked with the RUBISCO and carboxysome genes as in other bacteria. A candidate was also identified in the *Arabidopsis* nuclear genome. A straight forward understanding of this gene appears to be complicated again by horizontal gene transfer as preliminary analysis indicates the *Synechococcus* WH8012/*Prochlorococcus* *cbbx* gene is more similar to the alpha proteobacteria *cbbx* and red algal chloroplast genes than to the *Nostoc* or *Arabidopsis* genes.



**Figure 2 – Neighbor-joining phylogeny of the *cbx* protein sequence.** Numbers correspond to bootstrap values (100 replicates).



**Figure 3 – Neighbor-joining phylogeny of the *rbcL* protein sequence.** Numbers correspond to bootstrap values (100 replicates).

The *cbbx* and *rbcL/S* genes have been puzzling to many of us interested in chloroplast evolution. There several explanations including; horizontal gene transfer between alpha proteobacteria and the red chloroplast lineage, the red chloroplast lineage is the result of a separate endosymbiont event than the green lineage (multiple origins of chloroplast) and the original cyanobacterial endosymbiont harbored two divergent copies of *rbcL* and *rbcS*.

For my perspective I think the evidence for a single origin of the chloroplast is better than the multiple origins data, but the *cbbx* gene is making me rethink this. I have attached a phylogeny of this gene. It does group with the red cp lineage and with alpha proteobacteria (Figure 2), but the *Synechococcus* and *Prochlorococcus* gene either group with the red cp lineage or with alpha proteobacteria depending on the method (not shown). Confidence is not great in either case. However, statistically the red cp lineage and alpha proteobacteria do not group together when *Synechococcus* and *Prochlorococcus* are included. Thus it looks like the chloroplast *cbbx* gene was probably in the original cyanobacterial endosymbiont and not derived from horizontal gene transfer.

Since we do not know of any cyanobacteria with multiple divergent copies of *rbcL* and *rbcS* (although we may not have look hard enough yet) the *cbbx* data seems to provide a piece of data supporting multiple origins of the chloroplast.

The regulatory networks are also likely to be different in *Synechococcus* WH8201. A homolog of *rbcR*, a transcriptional regulator of genes in the carbon assimilation pathway is present, but another key member of this *LysR* family of transcription factors, *ccbR*, is missing in *Synechococcus* WH8201 and the two *Prochlorococcus* strains, but present in the model cyanobacterial systems. Interestingly this is the same situation as in algae, where *rbcR* is found in the chloroplast genome. Surprisingly *Arabidopsis* is missing both transcription factors. On a cautionary note the annotation of *LysR* transcription factor family members is extremely misleading and likely to be wrong in many cases. We are working on clarifying these relationships and the gene family nomenclature (Please contact us for preliminary phylogenetic analysis and the linking of gene families to direct experimental evidence).

Based on this preliminary analysis and the *Prochlorococcus* genome analysis of Hess *et al.* 2001 we have identified missing genes and HGT events involving key genes. Thus we expect there to be considerable diversity in the regulatory and mechanistic approaches used for inorganic carbon assimilation both within the marine *Synechococcus/Prochlorococcus* group and between this group and the rest of cyanobacteria. To better understand this diversity, we support Brian Palenik's recommendation for sequencing additional marine *Synechococcus* genomes (perhaps according to the groups identified by Rocap *et al.* 2002). We also recommend to sequence key genes including *cbbx* to estimate levels of variation and to infer adaptive changes. It would also be useful to pinpoint which, if any cyanobacteria outside the marine contain the unusual RUBISCO, carboxysome genes and the *cbbx* regulatory protein. It could be that this horizontal gene transfer event bringing in key inorganic carbon assimilation genes is one of the key determinants in the global success of the *Synechococcus/Prochlorococcus* group. Thus we could possibly have a handle on key innovative features that allow a group to exploit broad environments.

**Other Funded DOE Cyanobacterial Proposals:** In the last year the DOE has funded several large cyanobacterial projects that will generate data relevant to our efforts to model *Synechococcus*. Listed here are the projects, abstracts, and comments on their approaches.

*Brian Palenik (Scripps Inst. of Oceanography). Transport and its Regulation in Marine Microorganisms: A Genomics-Based Approach.* “This work will explore transport proteins and their regulation in *Synechococcus* WH8102 and add to our knowledge of nutrient transport in an environmentally important photosynthetic organism. *Synechococcus* represents an important primary producer, and regulation of transport could have profound environmental importance. The expression of the significant genes that are increased in expression as the organism is put through a series of environmental manipulations, including differing phosphorous, nitrogen, and light regimes, will be probed with microarrays. Then, “knockout mutants” of the transporters will be examined. Comparisons will be carried out with other ocean photosynthetic organisms (whose genome sequences are available) since the activities of these microbes can significantly affect the biosphere.” (<http://www.ornl.gov/microbialgenomes/2001awards.html>)

*Comments:* This is one of the few and possibly the only experimental proposal currently funded that will provide data on our study organism, *Synechococcus* WH8102. The environmental and genetic manipulations will be extremely important for our efforts to relate environmental and genetic changes to cellular phenotypes. It would be very useful if the experiments included different carbon regimes and if corresponding phenotypic data (*i.e.* growth characteristics, cellular energetics) were measured for the knockouts and during the environmental manipulations

*George Church, Sallie Chisholm, Martin Polz, Roberto Kolter, Fred Ausubel, Raju Kucheralapati, Steve Lory, Steve Gygi, Mike Laub (Harvard & MIT) Microbial Ecology, Proteogenomics & Computational Optima.* “We propose to leverage our experience in technology development (in proteomics, selection-genomics, and computational modeling) and experience in the biology/ecology of three microbial systems sequenced by (and of key interest to) DOE to integrate the four GTL goals (below) in the context of the main DOE missions (energy production/carbon sequestration and environmental clean up). The microbial genera proposed include uncultivated isolate plus *Prochlorococcus*, a group of species responsible for a major fraction of the earth's microbial carbon fixation, *Caulobacter*, a species relevant to dilute scavenging and bioremediation as well as cell division, *Pseudomonas*, displaying a broad range of metabolic pathways including chemical/biological toxins and well-studied biofilms. We emphasize the computation theme of optimality and the concept that overconstraining integrative models with comprehensive datasets facilitates examination of inconsistencies for insights into data collection issues and biological discoveries.” ([http://arep.med.harvard.edu/DOEGTL/GTLBOS\\_intro.htm](http://arep.med.harvard.edu/DOEGTL/GTLBOS_intro.htm))

*Comments:* The main organismal focus of this proposal is really *Prochlorococcus*. It would be worthwhile for all members of this group interested in using experimental data in modeling and/or analysis to take a closer look at their proposal. The link is posted on our GTL project home page or go directly to <http://arep.med.harvard.edu/DOEGTL>. Here are a few of the interesting experimental approaches.

- Library of gene knockouts
- Fitness assays of gene knockouts
- Functional diversity assays for gene related to specific processes
- Field assays
- Ecotypic differences at selected loci
- Phage effects on host fitness
- Environmental effects of gene expression (microarray)
- 2D-gel/Mass spec
- Protein-protein interaction traps

*Willem Vermaas (Arizona State University) Genome Sequence-Based Functional and Structural Analysis of a Transformable Cyanobacterium: the Synechocystis sp. PCC 6803 Microbial Cell Project:* “We will study *Synechocystis*, the first photosynthetic organism whose genome was completely sequenced. The main approach will be to exploit the library of targeted gene deletion mutants that are available (with more becoming available) to study photosynthesis and electron transport pathways. This will be combined with other proteomic, metabolic, and intracellular localization information to build a metabolic model.” (<http://microbialcellproject.org/funding>)

*Comments:* Because *Synechocystis* has a well-developed genetic system we may see results appearing more quickly from this project.

**Recommendations:** The regulatory and mechanistic pathways of inorganic carbon assimilation in *Synechococcus* WH8102 are very different than those found in model cyanobacterial systems due to horizontal gene transfer events and the absence of key genes and pathways. Thus, it will be imperative that we understand this diversity at the population and ecotype levels through per genome and per loci sequencing efforts. Their needs to be more of an emphasis on phenotypic analyses in order to effectively capture genetic and biochemical information in models of carbon sequestration behavior. Phenotypic information linked to genetic and environmental data will also be necessary to understand the modularity of processes involved in inorganic carbon assimilation and between these processes and other process contributing to cellular fitness. It is quite likely that the collective effort of this GTL project will define key innovative features, including those involved in carbon sequestration, that allow the marine *Synechococcus* to exploit broad environments. This in turn should lead to better informed engineering decisions.

## References

- Brahamsha, B. (1996). An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus*. *Proc Natl Acad Sci U S A* 93, 6504-9.
- Brahamsha, B. (1996). A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl Environ Microbiol* 62, 1747-51.
- Collier, J. L., Brahamsha, B., and Palenik, B. (1999). The marine cyanobacterium *Synechococcus* sp. WH7805 requires urease (urea amidohydrolase, EC 3.5.1.5) to utilize urea as a nitrogen source: molecular-genetic and biochemical analysis of the enzyme. *Microbiology* 145, 447-59.
- Gibson JL, Tabita FR. Analysis of the cbbXYZ operon in *Rhodobacter sphaeroides*. *J Bacteriol.* 1997 Feb;179(3):663-9.
- Hess HR, Rocap G, Ting CS, Larimer F, Stilwagen S, Lamerdin J, Chisholm SW. The photosynthetic apparatus of *Prochlorococcus*: Insights through comparative genomics. *Photosynthesis Res.* 2001 70:53-71
- Hihara Y, Kamei A, Kanehisa M, Kaplan A, Ikeuchi M. DNA microarray analysis of cyanobacterial gene expression during acclimation to high light. *Plant Cell.* 2001 Apr;13(4):793-806.

- Maeda S, Badger MR, Price GD. Novel gene products associated with NdhD3/D4-containing NDH-1 complexes are involved in photosynthetic CO<sub>2</sub> hydration in the cyanobacterium, *Synechococcus* sp. PCC7942. *Mol Microbiol.* 2002 Jan;43(2):425-35.
- Maier UG, Fraunholz M, Zauner S, Penny S, Douglas S. A nucleomorph-encoded CbbX and the phylogeny of RuBisCo regulators. *Mol Biol Evol.* 2000 Apr;17(4):576-83.
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci U S A.* 2002 Sep 17;99(19):12246-51.
- Palenik B. The genomics of symbiosis: Hosts keep the baby and the bath water. *Proc Natl Acad Sci U S A.* 2002 Sep 17;99(19):11996-7.
- Partensky, F., Hess, W. R., and Vaultot, D. (1999). Prochlorococcus, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* 63, 106-27.
- Rocap G, Distel DL, Waterbury JB, Chisholm SW. Resolution of Prochlorococcus and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. *Appl Environ Microbiol.* 2002 Mar;68(3):1180-91.
- Scanlan, D. J., and West, N. J. (2002). Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microbiology Ecology* 40, 1-12.
- Shibata M, Ohkawa H, Kaneko T, Fukuzawa H, Tabata S, Kaplan A, Ogawa T. Distinct constitutive and low-CO<sub>2</sub>-induced CO<sub>2</sub> uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms. *Proc Natl Acad Sci U S A.* 2001 Sep 25;98(20):11789-94.
- 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 Apr;40(1):235-44.
- Ting CS, Rocap G, King J, Chisholm SW. Phycobiliprotein genes of the marine photosynthetic prokaryote Prochlorococcus: evidence for rapid evolution of genetic heterogeneity. *Microbiology.* 2001 Nov;147(Pt 11):3171-82.
- Uchino Y, Yokota A. "Green-like" and "Red-like" RubisCO *cbbL* genes in *Rhodobacter azotoformans*. *Mol. Biol. Evol.* 2003 20(5):821-830.