The search for ancestrally anoxygenic or non-phototrophic Cyanobacteria

James Hemp

jim.hemp@gmail.com

University of Illinois at Urbana-Champaign

Abstract

Photosynthesis was one of the most important evolutionary innovations in the history of life on Earth. The advent of oxygenic photosynthesis by Cyanobacteria permanently changed the geochemistry of Earth's surface and paved the way for the evolution of oxygen respiring organisms. It is currently unknown how this complex process evolved. Since there is no evidence for the presence of photosynthesis in the last common ancestor, it seems likely that early Cyanobacteria were either non-phototrophic or anoxygenic phototrophs and later acquired the ability to oxidize water into oxygen. The discovery of organisms which retain these ancestral characteristics would be invaluable for understanding the evolution of oxygenic phototrophy. Here we use sequence analyses to search for early branching clades of Cyanobacteria. A cultured early branching Cyanobacteria, Vampirovibrio chlorellavorus, was discovered which is likely non-phototrophic.

Introduction

The evolution of oxygenic phototropy in the ancestors of modern Cyanobacteria during the Archean eon was one of the most important events in the history of life on Earth. Geochemical studies have provided strong evidence that the early Earth was anoxic; however by 2.5 Ga O_2 was present in the surface waters of the ocean, and by 2.4 Ga O_2 had begun to accumulate in the atmosphere in what is known as the great oxidation event (GOE) (Knoll A). The transition from an anoxic to oxic environment had a profound effect on life and Earth's surface. It changed the bioavailability of redox sensitive metals and radically altered Earth's biogeochemical cycles. The availability of O_2 ultimately led to the evolution of aerobic respiration, which increased the productivity of microbial ecosystems and provided life access to many redox coupled reactions that were previously unavailable. It also allowed for the evolution of modern Eukaryotes, and was a prerequisite for the evolution of macroscopic multicellular life.

Chlorophyll based phototrophy can be divided into two types depending on the electron donor used, oxygenic and anoxygenic. Anoxygenic phototrophs are found in five known phyla; the Proteobacteria, Firmicutes, Chlorobi, Chloroflexi and Acidobacteria, whereas the Cyanobacteria are the only organisms known that can perform oxygenic photosynthesis. There are only three molecular components shared by all phototrophs; reaction centers/photosystems, chlorophyll and complex III. Phylogenetic analysis of the reaction centers and chlorophyll biosynthesis pathways demonstrated that these components of photosynthesis have been subject to lateral transfer (Raymond J, *et al*. and Hohmann-Marriott MF and Blankenship RE), making it difficult to determine the order of evolution of the different phototrophic groups.

The high potential electron transfer chains of chlorophyll based phototrophy and aerobic respiration share a common core architecture composed of a number of evolutionarily related proteins (**Figure 1**).

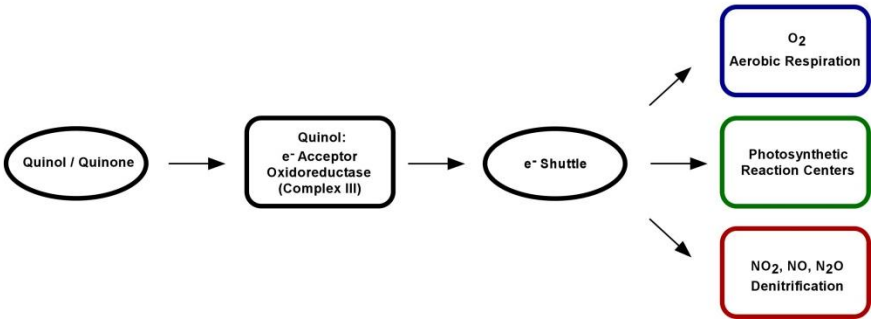

Figure 1. High-redox potential module showing the central role of complex III.

Analysis of the quinol:electron acceptor oxidoreductases (Complex III) and oxygen reductase members of the heme-copper oxidoreductase superfamily (Complex IV) suggests that aerobic respiration originated within early Cyanobacteria after the evolution of oxygenic photosynthesis (Hemp J, *et al*.). All other phyla of Bacteria and Archaea received the ability to perform aerobic respiration via lateral gene transfer. In most currently known phyla early branching groups are anaerobic, implying that at the time of the major diversification of

microbial life oxygenic photosynthesis had not yet evolved. Since complex III is shared with phototrophic pathways it is possible to determine the relative timing of aerobic respiration and phototrophy within a given clade by comparing the evolution of genes for both processes with that of complex III. This demonstrates that the phototrophic electron transfer chains utilized by many extant anoxygenic clades were likely added onto preexisting pathways used for aerobic respiration, suggesting that those anoxygenic clades acquired the ability for phototrophy after the evolution of oxygenic photosynthesis (Hemp J, *et al*.) (**Figure 2**).

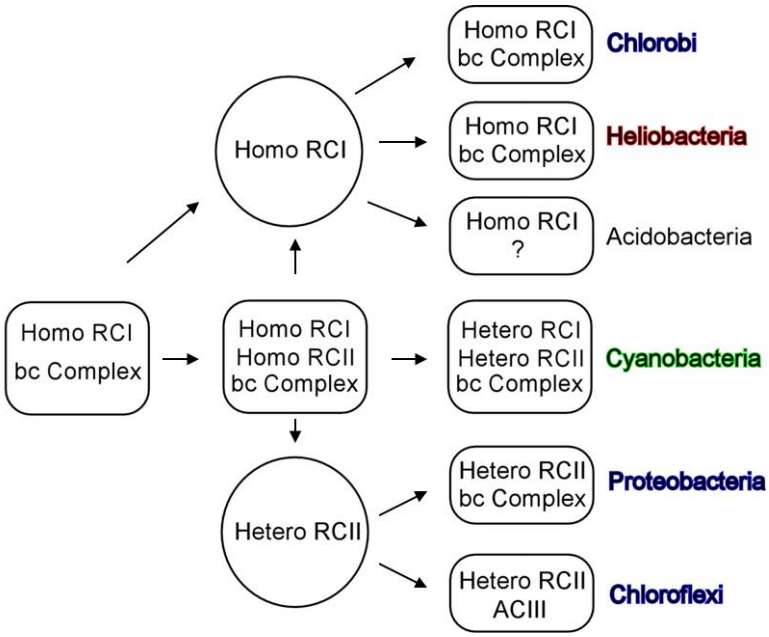

Figure 2. Proposed evolution of chlorophyll based phototrophy based on the evolution of the high-potential redox module.

To test this hypothesis it is critical to identify and characterize any non-phototrophic organisms in clades which diverged before the acquisition of phototrophy in these phyla. Here we attempt to identify ancestrally anoxygenic or non-phototrophic Cyanobacteria.

Materials and Methods

16s analysis of the Cyanobacteria clade

Sequences which were classified as Cyanobacteria within ARB were downloaded via the Silva ribosomal RNA database server (http://www.arb-silva.de/). Sequences from closely related phyla (Chloroflexi, TM7, Actinobacteria, Thermus and OP10) were also retrieved. The sequences were aligned using NAST via the green genes interface (http://greengenes.lbl.gov/cgi-bin/nph-NAST_align.cgi). Hyper-variable regions of the alignment were removed using the Lane mask, producing an alignment with 1287 characters. Sequences with >97% identity were removed using Jalview. Maximum likelihood (RAxML) and Bayesian (Mr. Bayes) phylogenetic trees were calculated utilizing the teragrid at CIPRES. The sequences from closely related phyla were used as an outgroup to root the Cyanobacterial tree.

Identification of putative Cyanobacteria within metagenomic datasets

217 metagenomes within the JGI IMG database (http://img.jgi.doe.gov/cgi-bin/m/main.cgi) were searched for both 16s and RpoB genes. Over 18000 partial and full length 16s and RpoB genes were identified. 16s sequences greater than 500 bp (1612 sequences) were aligned and classified with ARB. RpoB sequences greater than 600 aa (363 sequences) were aligned to an in-house curated database. Maximum likelihood (RAxML) and Bayesian (Mr. Bayes) phylogenetic trees for RpoB were calculated utilizing the teragrid at CIPRES.

Solution CARD-FISH

Sample preparation

- Overnight cultures of Rhodopseudomonas palustris (α Proteobacteria), Bacillus azotoformans (Firmicutes) and Escherichia coli (γ Proteobacteria) were grown. 500 ul from each culture was mixed together in a 2 ml tube by vortexing for 30 seconds. 100 ul of the mixture was used in each of the CARD-FISH reactions.

Fixation

- For cells which were not fixed 100 ul of the sample was added to 900 ul PBS.
- For cells fixed with ethanol (Firmicutes and Eubacteria probes) 100 ul of the sample was added to 400 ul PBS and 500 ul absolute ethanol.
- For cells fixed with PFA (α Proteobacteria , γ Proteobacteria, and the non-probe) 100 ul of the sample was added to 850 ul of PBS and 50 ul of 20% PFA.
- The cells were fixed for 2 hours at -20C.
- After fixation the cells were centrifuged at 15000g for 5 minutes. The cells were then resuspended in 1 ml PBS and sonicated for 30 seconds at 30% power. The cells were then centrifuged at 15000g for 5 minutes.

Permeabilization and endogenous peroxidase inactivation

- The cells were then permeabilized by resuspending in 500 ul 0.1M HCl for 1 minute at room temperature. The cells were then centrifuged at 15000g for 5 minutes.
- The endogenous peroxidases were then inactivated by resuspending in 1 ml 0.15% H_2O_2 (in methanol) and incubating at room temperature for 30 minutes. The cells were centrifuged at 15000g for 5 minutes.

- The cells were then washed by resuspending in 1 ml PBS. The cells were centrifuged at 15000g for 5 minutes.

Hybridization, amplification and DAPI staining

- The cell pellets were not resuspended in the following steps.
- The probes for each condition were hybridized to the cells by adding 150 ul hybridization buffer and 6 ul probe working solution. The cells were then incubated at 46C for 3 hours. The cells were then centrifuged at 15000g for 5 minutes.

Probe list

All bacteria - Eub338
α Proteobacteria - Alf968
γ Proteobacteria - Gam42a + Bet42a competitive probe
Firmicutes - LGC354a
non-probe – NON

- 2 ml of pre-heated washing buffer (48C) was added and the cells were incubated for 15 minutes at 48C.

Washing buffer formamide concentration

Eub338 – 35%
Alf968 – 35%
Gam42a + Bet42a competitive probe – 35%
LGC354a – 20%
NON – 35%

- The cells were then centrifuged at 15000g for 5 minutes and washed with 1 ml PBS at room temperature for 5 minutes. The cells were then centrifuged at 15000g for 5 minutes.
- The amplification of the peroxidase signal was performed by adding 750 ul amplification buffer with Alexa 488 dye. The cells were then incubated at 37C for 20 minutes. The cells were then centrifuged at 15000g for 5 minutes.
- The cells were then washed with 500 ul PBS and incubated at room temperature for 10 minutes. The cells were then centrifuged at 15000g for 5 minutes.
- The washing step was repeated 3 more times.
- The cells were resuspended in 500 ul PBS and and sonicated for 30 seconds at 30% power. At this stage the cells are ready for flow sorting and can be stored at 4C for up to one week.
- 5 ul of each sample was spotted on a .2 um GTTP filter and embedded with DAPI solution.

PCR

25 ul Phusion High-Fidelity polymerase master mix
1 ul of sample
2ul 8f universal 16s primer
2ul 1492r universal 16s primer
20ul H_20

- PCR was preceeded with a 5 minute boil step at 95C to lyse cells.
- 30 cycles of PCR was performed with a melting temperature of 46C and a 1.5 minute extension time.

Results and Discussion

Early branching Cyanobacteria

Cyanobacteria related organisms have been identified in a number of animal gut microbiomes, leading to the suggestion that they may be ancestrally non-phototrophic Cyanobacteria (Ley RE, *et al*. and Warnecke F, *et al*.). In order to identify putative early branching Cyanobacteria clades we built phylogenetic trees of environmental samples using two robust molecular markers: 16s ribosomal RNA and the beta subunit of DNA dependent RNA polymerase (rpoB) (Case RJ, *et al*.). The current SILVA 16s ribosomal RNA release has over 4 million sequences, while rpoB sequences are available mainly from sequenced genomes. The vast majority of 16s sequences in data bases are derived from 16s clone libraries or next generation sequencing platforms which first require the 16s genes to be amplified by PCR. Since there are known PCR biases for 16s amplification, and a limited number of environmental rpoB sequences available, we decided to supplement the datasets with 16s and rpoB genes identified from JGI metagenome projects. This approach worked well for identifying novel 16s and rpoB sequences within the Global Ocean Sampling (GOS) Expedition data set (Wu D, *et al*.).

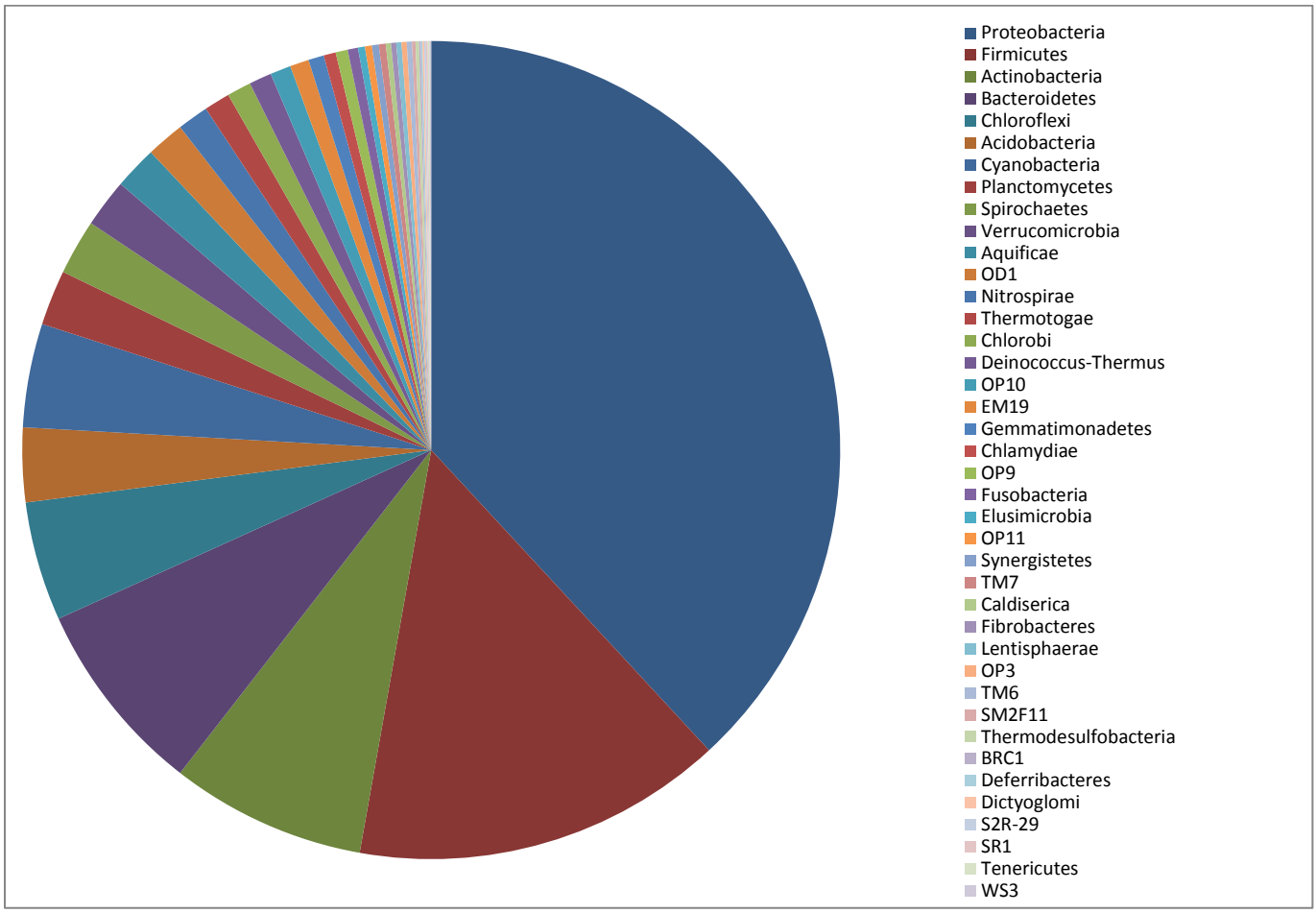

Figure 3. Bacterial 16s diversity within JGI metagenomes.

Analysis of 215 JGI metagenomes identified a large number of 16s (>10000) and rpoB (>9000) full and partial gene sequences. The 16s gene sequences greater than 500 bp in length (1612 total) were characterized using ARB. 90% of the sequences were from Bacteria and 10% were from Archaea. The bacterial sequences were distributed among 40 different phyla (**Figure 3 and Table 1**), and members of 13 candidate phyla were detected. Four of the sequences (IMG Gene Object ID's: 2049510114, 2049460324, 2015346492, and 2093973719) clustered with putative early branching Cyanobacteria clades. The rpoB genes greater than 600 amino acids in length were aligned to an in-house manually curated rpoB dataset. Phylogenetic analysis identified a number of sequences which are likely from candidate phyla (**Figure 4**). Four rpoB sequences were found which clustered at the base of the Cyanobacteria.

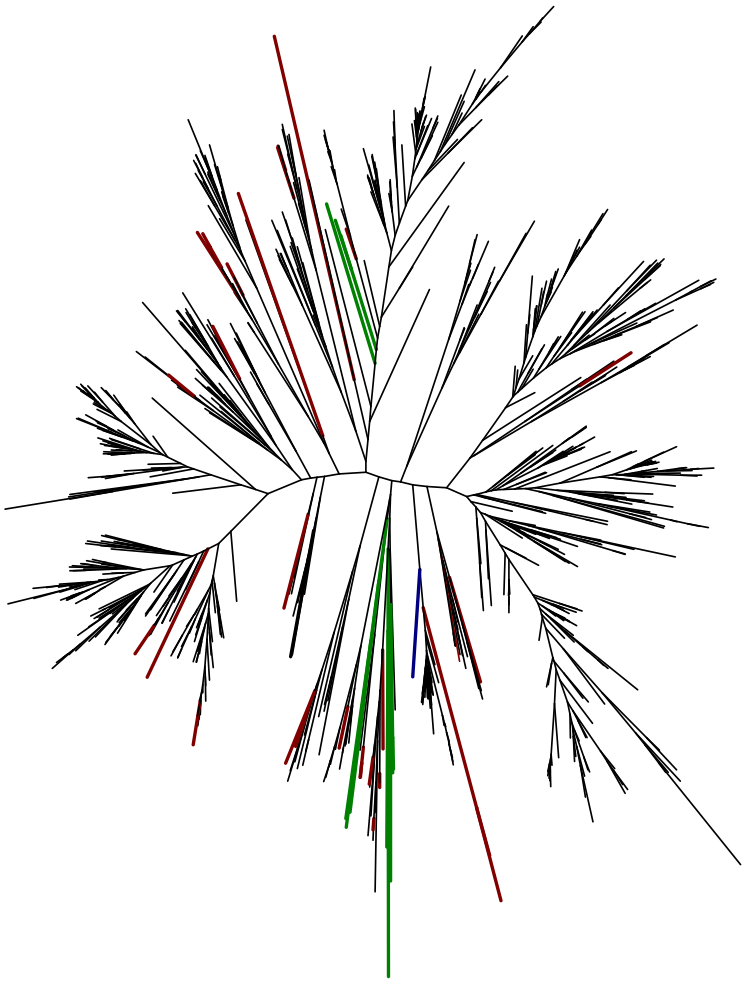

Figure 4. Bacterial rpoB diversity within JGI metagenomes. Branches in red are from sequences with <95% identity to known sequences, branches in green are from putative candidate phyla, and branches in blue are from early branching Cyanobacteria.

The SILVA database recognizes eight clades which appear to branch at the base of the Cyanobacteria (4C0d-2, ML635J-21, MLE1-12, QB36, SHA-109, SM2F09, Vampirovibrio and

WD272). We performed a detailed phylogenetic analysis of the 16s genes from these clades to determine whether they belong to the Cyanobacteria, or that their placement within the Cyanobacteria is an artifact of the algorithm used to classify 16s sequences by SILVA. The 4C0d-2, MLE1-12, ML635J-21, SM2F09 and Vampirovibrio clades appear to be true members of the Cyanobacteria (**Figure 5**). The four early branching Cyanobacteria 16s genes identified in JGI metagenomes belonged to the MLE1-12 and ML635-21 clades. Most importantly the Vampirovibrio clade has a cultured representative within the ATCC strain bank ,Vampirovibrio chlorellavorus (ATCC 29753). Vampirovibrio chlorellavorus appears to be an obligate predator of the algae Chlorella vulgaris (Coder DM and Starr MP). It is unlikely that this organism is a phototroph due to its lifestyle. We attempted to acquire this strain for further characterization and genome sequencing, however it was on backorder from the ATCC and other culture collections have unfortunately lost it.

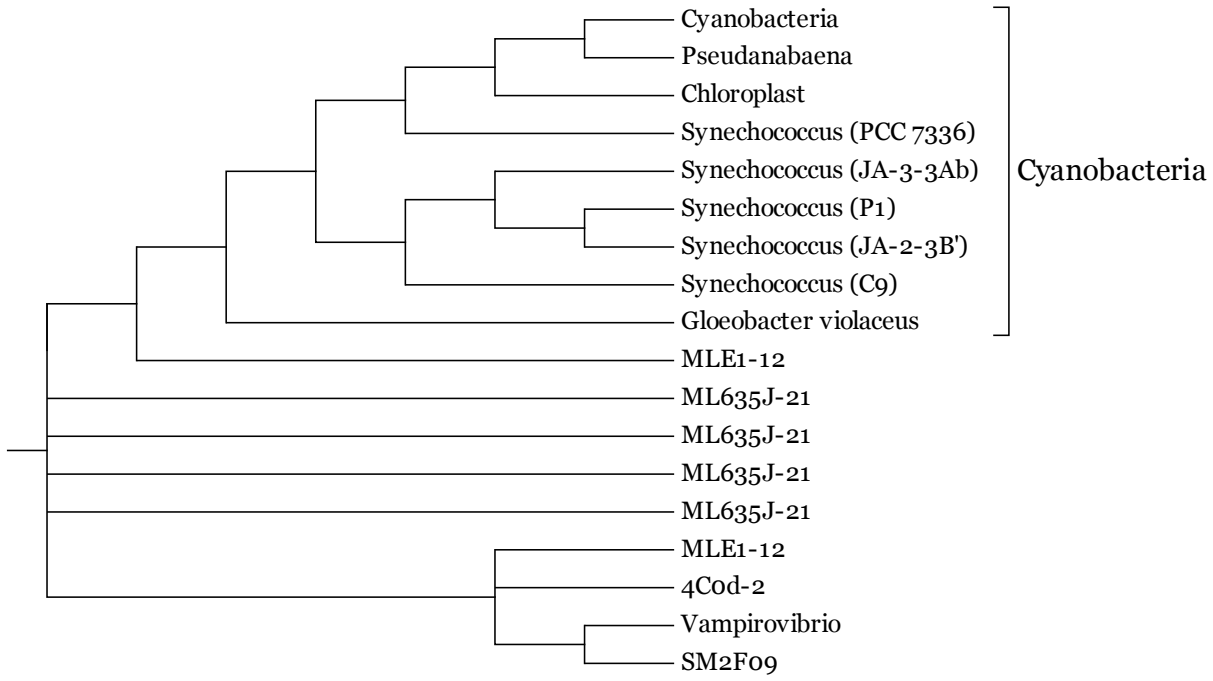

Figure 5. Cyanobacteria 16s phylogenetic tree. The clade labeled Cyanobacteria are all oxygenic phototrophs. The other putative early branching Cyanobacterial clades (4C0d-2, MLE1-12, ML635J-21, SM2F09 and Vampirovibrio) have unknown physiologies.

We also identified what appears to be an early branching Cyanobacteria (both 16s and rpoB genes) within a Poplar biomass decaying microbial community metagenome (**Figure 6**). This community was sequenced to a depth of 350Mb. 1500 protein encoding genes were sorted into the early branching Cyanobacteria bin, none of which function in photosynthesis. In fact no phototrophy genes were found in the entire dataset, suggesting that this organism may be a ancestrally non-phototrophic Cyanobacteria.

Indirect evidence for the presence of early branching Cyanobacteria within gut (4C0d-2 clade) and Poplar biomass decaying microbial communities (MLE1-12 clade), along with a

cultured predatory isolate (Vampirovibrio clade) suggests that non-phototrophic Cyanobacteria are likely to be present in Nature. Characterization of these organisms may provide substantial insight into the evolution of photosynthesis. Future work will include the design of PCR primers and FISH probes for these clades, with the goal of isolating a number these organisms for genome sequencing.

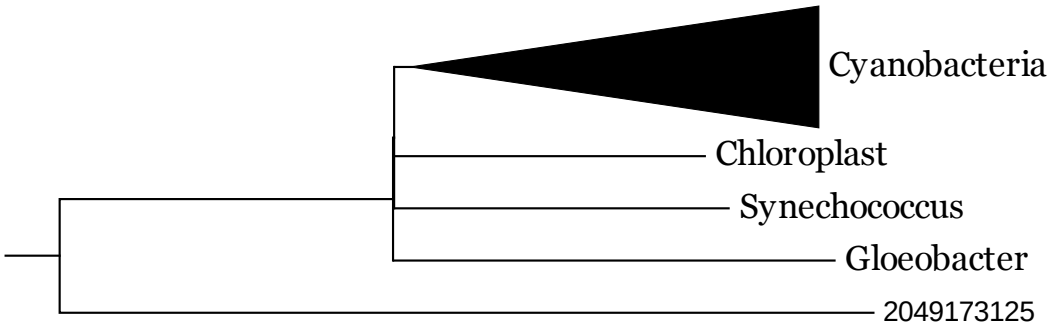

Figure 6. Cyanobacteria rpoB phylogenetic tree. 2049173125 is a full length rpoB gene from a Cyanobacteria found in the Poplar biomass decaying microbial community (blue branch in Figure 4).

Solution CARD-FISH

FISH coupled with cell sorting can be a powerful tool to isolate organisms from the environment for genome sequencing (Podar M., *et al*. and Kalyuzhnaya MG, *et al*.). However the fixation procedures used can damage the DNA so that is useless for multiple displacement amplification (Yilmaz S., *et al*.). We performed solution CARD-FISH using multiple fixation techniques to identify which ones would allow for the PCR amplification of DNA.

A mixture of three different strains was used as a sample for the CARD-FISH procedure; Rhodopseudomonas palustris (α-Proteobacteria), Bacillus azotoformans (Firmicutes) and Escherichia coli (γ-Proteobacteria).

The samples were fixed using three different techniques; no fixation, ethanol, and PFA. During the first experiment non-specific probe binding was observed for all samples. For the second experiment two extra wash steps were added after the amplification step. This reduced non-specific probe binding for samples fixed with PFA, but not for samples fixed with ethanol.

Probe	Cell counts	Fixation
Eubacteria	92%	Ethanol
α-Proteobacteria	37%	PFA
γ -Proteobacteria	46%	PFA
Firmicutes	86%	Ethanol
Non	17%	PFA

PCR of the 16s gene was attempted for each sample. Bands were detected for all three fixation procedures, however the strongest bands were obtained with the no fixation samples. In

future experiments FISH and DOPE-FISH (Stoecker K, *et al*.) will be attempted with different fixation techniques to identify the most promising preparation for cell sorting and sequencing.

Acknowledgements

I thank Dan and Steve for sharing their knowledge and providing such a wonderful learning environment. I also thank Sara for teaching me FISH techniques and Chuck, Carrie and Lizzie for the discussions on Cyanobacteria, photosynthesis and genome evolution.

References

Case RJ, Boucher Y, Dahllöf I, Holmström C, Doolittle WF, Kjelleberg S. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. Appl Environ Microbiol. 2007 73:278-88.

Coder DM and Starr MP. Antagonistic Association of the Chlorellavorus Bacterium ("Bdellovibrio" chlorellavorus) with Chlorella vulgaris. Current Microbiology 1978 1:59-64.

Hemp J, Pace LA, Gennis RB. The evolution of aerobic respiration. Submitted.

Hohmann-Marriott MF, Blankenship RE. Evolution of photosynthesis. nnu Rev Plant Biol. 2011 62:515-48.

Kalyuzhnaya MG, Zabinsky R, Bowerman S, Baker DR, Lidstrom ME, Chistoserdova L. Fluorescence in situ hybridization-flow cytometry-cell sorting-based method for separation and enrichment of type I and type II methanotroph populations. Appl Environ Microbiol. 2006 72:4293-301.

Knoll A. (2003) Life on a Young Planet: The First Three Billion Years of Evolution on Earth. Princeton Science Library

Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A. 2005 102:11070-5.

Podar M, Abulencia CB, Walcher M, Hutchison D, Zengler K, Garcia JA, Holland T, Cotton D, Hauser L, Keller M. Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl Environ Microbiol. 2007 73:3205-14.

Raymond J, Zhaxybayeva O, Gogarten JP, Gerdes SY, Blankenship RE. Whole-genome analysis of photosynthetic prokaryotes. Science. 2002 298:1616-20.

Stoecker K, Dorninger C, Daims H, Wagner M. Double labeling of oligonucleotide probes for fluorescence in situ hybridization (DOPE-FISH) improves signal intensity and increases rRNA accessibility. Appl Environ Microbiol. 2010 76:922-6.

Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottesen EA, Zhang X, Hernández M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P, Leadbetter JR. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature. 2007 450:560-5.

Wu D, Wu M, Halpern A, Rusch DB, Yooseph S, Frazier M, Venter JC, Eisen JA. Stalking the fourth domain in metagenomic data: searching for, discovering, and interpreting novel, deep branches in marker gene phylogenetic trees. PLoS One. 2011 6:e18011.

Yilmaz S, Haroon MF, Rabkin BA, Tyson GW, Hugenholtz P. Fixation-free fluorescence in situ hybridization for targeted enrichment of microbial populations. ISME J. 2010 4:1352-6.

Table 1. Bacterial phylogenetic distribution and abundance of 16s sequences from JGI metagenomes.

Acidobacteria	43
Actinobacteria	113
Aquificae	25
Bacteroidetes	113
BD1-5	
BHI80-139	
BRC1	1
Caldiserica	3
Chlamydiae	7
Chlorobi	14
Chloroflexi	69
Chrysiogenetes	
CK-1C4-19	
Cyanobacteria	60
Deferribacteres	1
Deinococcus-Thermus	13
Dictyoglomi	1
Elusimicrobia	4
EM19	11

Fibrobacteres	3
Firmicutes	215
Fusobacteria	6
GAL08	
Gemmatimonadetes	9
GOUTA4	
HDB-SIOH1705	
Hyd24-12	
JL-ETNP-Z39	
Kazan-3B-28	
LD1-PA38	
Lentisphaerae	3
MVP-21	
Nitrospirae	18
NPL-UPA2	
OC31	
OD1	22
OP10	12
OP11	4
OP3	3
OP9	7
Planctomycetes	32
Proteobacteria	558
RF3	
RsaHF231	
S2R-29	1
SM2F11	2
Spirochaetes	32
SR1	1
Synergistetes	4
TA06	
Tenericutes	1
Thermodesulfobacteria	2
Thermotogae	15
TM6	3
TM7	4
Verrucomicrobia	28
WCHB1-60	
WS3	1
WS6	