Chapter 2

Probing Functional Diversity of Thermophilic Cyanobacteria in Microbial Mats

Devaki Bhaya*
Department of Plant Biology, Carnegie Institution for Science, 260 Panama Street, Stanford, CA 94305, USA

Summary

In recent years there has been growing appreciation of the unexpectedly large genetic diversity of microbes in the environment. This diversity has important implications for our understanding of photosynthesis in populations and in the environment. Conventional methodologies often cannot

*Author for correspondence, e-mail: dbhaya@stanford.edu

R. Burnap and W. Vermaas (eds.), Functional Genomics and Evolution of Photosynthetic Systems, Advances in Photosynthesis and Respiration 33, pp. 000–000, DOI 10.1007/978-94-007-1533-2_2,

effectively capture this aspect. This chapter describes new approaches including comparative genomic and metagenomic approaches combined with a more detailed understanding of the metabolism and functionalities of cyanobacteria. This approach, which can be defined broadly as "functional ecogenomics" is partly motivated by the availability of high-throughput sequence data, which are steadily being acquired. The focus is on unicellular cyanobacteria in the hot spring microbial mats of Yellowstone National Park, which are primary producers in this prokaryotic community. We took a three-pronged approach, in which we (a) acquired complete genome sequences from two dominant *Synechococcus* sp. and carried out a comparative genomic analysis to understand the functional differences between these temperature adapted isolates; (b) produced a metagenome dataset that allows us to place genomic information in the context of the community within which these cyanobacteria grow and evolve; and (c) obtained pure isolates of some dominant organisms that allows us to manipulate them in a well-defined laboratory setting. *In situ* transcriptomics also allowed quantification of transcripts during the diel cycle. These diverse approaches and the ability to measure environmental parameters such as light and O_2 levels allows us detailed insight into the microbial mat system. Such an approach could be used to study a wide array of photosynthetic microorganisms as populations and interacting communities. As sequencing capacity, single cell capture techniques, proteomics and imaging techniques become more widely accessible we expect to obtain ever more detailed information about natural communities.

I. Introduction

Fundamental insights into the functional components of oxygenic photosynthesis have been derived from experiments conducted with model cyanobacteria. Cyanobacteria are the progenitor of plastids in all vascular plants and much of the photosynthetic machinery of these vastly separate lineages still function in similar ways (Howe et al., 2008; Archibald, 2009). Thus, the cyanobacterial photosynthetic apparatus is considered to be a useful proxy for the study of photosynthetic processes in all oxygenic photoautotrophs. Crystal structures of photosynthetic proteins and complexes, elucidation of pathways of photosynthetic electron flow, and biophysical features of photosynthetic reactions have used model cyanobacteria (such as *Synechococcus elongatus* PCC 7942 {also known as *Anacystis nidulans* R2}, *Synechocystis* sp. PCC 6803 and *Thermosynechococcus elongatus*). The major advantages in using cyanobacteria for studying photosynthesis and other processes are: (i) they have a much shorter life cycle than most plants and algae, (ii) particularly the unicellular cyanobacteria represent a uniform population of cells that all may respond similarly within the population, (iii) some cyanobacteria are able to grow heterotrophically, so mutants can survive if photosynthesis has been eliminated, (iv) it is easy to grow large volumes of cells for the isolation of cell fractions or for biochemical analyses, (v) many cyanobacteria are easily manipulated genetically (e.g., targeted disruption of specific genes), (vi) they have relatively small genomes, without introns, facilitating the identification of coding regions in comparison with eukaryotic photosynthetic organisms.

In recent years there has been growing appreciation of the unexpected genetic diversity of microbes in the environment (see Ward et al., Chapter 1, *this volume*). This molecular diversity has important implications for functional genomics and particularly for our understanding of photosynthesis in populations and in the environment. Conventional methodologies that rely on isogenic populations of model organisms and gene-by-gene analysis often cannot effectively capture this aspect. Thus, the goal of this chapter will be to describe some new approaches and to emphasize their potential. We have attempted to bring together perspectives from a comparative genomic and metagenomic approach, combined with those generated by an increasingly detailed understanding of the metabolism and functionalities of cyanobacteria in the hot spring microbial mats. This approach, which can be defined broadly as "functional ecogenomics" is motivated by the availability of the vast wealth of information derived from genomics and metagenomics projects and focuses on the dynamic aspects of gene expression.

A. Cyanobacteria: Ubiquitous and Important Members of Communities and Environments

The revolution brought about by using high-throughput genomic sequencing has made an enormous impact on the field of microbiology. Traditionally, microbiology relied on the use of axenic cultures maintained in the laboratory coupled to a variety of approaches ranging from biochemistry to gene expression studies. However, the vast majority of bacteria in the environment cannot be easily cultivated and therefore are not amenable to traditional microbiological methodologies. This is a serious drawback if one wishes to understand microbial function in the context of the environment. Thus, the ability to leapfrog over the arduous and difficult step of getting axenic cultures by simply obtaining DNA sequence information directly from environmental samples has provided major insights. One important new insight is that the microbial universe is almost unbelievably diverse and that we know almost nothing about the range of diversity that exists for microbes in nature. Another important insight is that microbes can be found in significant numbers in almost any environment from the most mundane to the most extreme (Whitman et al., 1998). Thirdly, as microbial diversity is being explored with these powerful new genomic tools it is also becoming clear that most microbes in the environment do not "hack it alone" but exist and thrive as members of communities or consortia (Schloss and Handelsman, 2007; Cardenas and Tiedje, 2008; Wilmes et al., 2009).

Cyanobacteria, in particular, are ubiquitous (Whitton and Potts, 2000) and have the ability to thrive in a range of extreme environmental conditions including thermophilic (Brock, 1978; Ley et al., 2006) and psychrophilic environments (Christner et al., 2003) as well as in desert crusts where they can withstand extreme desiccation (Nagy et al., 2005). Cyanobacteria also form critical associations with other microbes in metabolically integrated consortia, and function as symbionts supplying carbon to fungi, plants and animals (Adams, 2000; Adams and Duggan, 2008). These features make cyanobacteria one of the most versatile and enduring groups of photosynthetic organisms on the Earth. Although freshwater organisms such as *Synechocystis* sp. strain PCC 6803, *Synechococcus elongatus* PCC 7942 and filamentous cyanobacteria such as *Nostoc* sp. (*Anabaena* sp.) strain PCC 7120 have been studied extensively as model organisms for various processes, we have very limited knowledge of genomic diversity and its relationship to photosynthetic potential among terrestrial cyanobacteria.

In environments such as the oceans it has been estimated that cell counts in surface water may exceed 10^5 cells per ml, which would amount to a total of ~3.6×10^{29} microbial cells with a total cellular carbon content of $\approx 3 \times 10^{17}$ *g* (Whitman et al., 1998). In the oceans, cyanobacteria are important and abundant primary producers; as a consequence, they play an important role in primary productivity and the cycling of inorganic carbon in diverse environments (Falkowski et al., 2008). They are also important players in the global biogeochemical cycling of nutrients since they can fix inorganic carbon, reduce molecular nitrogen, ferment sugars, and can alter the redox state of iron and sulfur compounds (Cohen et al., 1975; Partensky et al., 1999; Paerl et al., 2000; Guerrero et al., 2002; Teske and Stahl, 2002; Decker et al., 2005). These activities impact other microbes that associate with the cyanobacteria, and may be important driving forces that shape microbial interactions. Recent genomic sequencing efforts with various ecotypes of marine *Synechococcus* sp. and *Prochlorococcus* sp. have greatly expanded our knowledge of cyanobacterial diversity in the marine environment and demonstrated the potential of a comparative genomic approach. Several recent articles and reviews cover this active area of research (Suzuki and DeLong, 2002; DeLong and Karl, 2005; Coleman et al., 2006; Kettler et al., 2007; Frias-Lopez et al., 2008; Haverkamp et al., 2008).

The focus of this chapter will be on the thermophilic cyanobacterial populations in the microbial mats of Yellowstone hot springs, which have been a test case for the development of some relevant methodologies. Progress in the use of comparative genomic and metagenomic tools to understand evolving functions in environmentally relevant photosynthetic communities will be discussed. This is a relatively young field that has the potential to provide unique insights that are not accessible through the study of model organisms or of isogenic populations. It also holds the promise of understanding evolutionary processes in

bacteria within the context of an environment that has both predictable (e.g., day/night light levels) and unpredictable (e.g., cloud cover or nutrient fluxes) dynamic fluctuations.

B. Cyanobacteria in Microbial Mats and Biofilms

Microbial mats are considered modern-day analogs of ancient ecosystems represented by stromatolites, in which mineral depositions within a copious exopolysaccharide matrix have preserved the stratified cyanobacterial community structure (Schopf, 2000; Stal, 2000; Teske and Stahl, 2002). The oxygenic, photoautotrophic cyanobacteria are believed to have pioneered the formation of these early Earth communities ~2.5 billion years ago, and may have contributed to the oxygenation of the early Earth's atmosphere (Hoehler et al., 2001). Photosynthetic microbial mats occur in many terrestrial and aquatic environments such as hypersaline coastal pools of Guererro Negro (Ley et al., 2006), freshwater ponds, geothermal hot springs of Yellowstone National Park (Ward et al., 2002), cold dry valleys of Antarctica (Vincent, 2000; Vincent et al., 2004; Jungblut et al., 2006), and alkaline and low sulfide hot springs in Russia (Namsaraev et al., 2003). Photosynthetic microbial mats also form crusts on rocks (Stal, 2000; Wynn-Williams, 2000; Arakawa et al., 2006).

Some of these types of mats have extremely high calculated ratios of carbon assimilation to standing biomass, suggesting highly efficient carbon utilization. It has been suggested that microbial mats in marine environments may be much more efficient in nutrient acquisition and utilization than mixed planktonic populations (Paerl, 2000; Guerrero et al., 2002; Decker et al., 2005). Microbial mats can flourish in regions where predation is low, and stratified microbial mats tend to proliferate in diverse environments that are often inhospitable, so they provide a perfect paradigm for studying how moderately complex communities of microbes develop and optimize the utilization of scarce resources. Most but not all mat communities are comprised of a limited number of dominant prokaryotic genera, and the system is less complex than soil or marine ecosystems (Nubel et al. 1999, 2002; Paerl et al., 2000; Guerrero et al., 2002; Ley et al., 2006). In photosynthetic microbial mats, cyanobacteria and other prokaryotes can position themselves at various interfaces (such as the sediment/liquid interface or the air/surface interface) as a function of light, chemical and gas gradients. This leads to the formation and stabilization of biologically stratified layers, within which diverse metabolic processes can occur on a temporal scale (e.g., a diurnal or seasonal cycle). This partitioning of nutrient cycling, niche differentiation and homeostasis within the mat community may promote biological control of the microenvironment. This in turn can support higher survival rates than may be possible for individual species growing alone (Paerl et al., 2000). Finally, this biological partitioning can also influence trapped or underlying sediments and mineral precipitation so that the mats can take on a number of different morphological forms such as hard lithified crusts, laminated structures or loose biofilms (Reid et al., 2000).

C. Microbial Mat Communities in the Hot Springs of Yellowstone National Park

Yellowstone National Park (YNP) is unique in that it contains a vast array of geothermal features, several of which have been studied extensively from a geochemical, biological and historical perspective over many decades. There is an extensive literature pertaining to the geology and biology of YNP (for example, Allen and Day, 1935; Brock, 1978; Keiter and Boyce, 1991; Ward et al., 1998; Reysenbach and Cady, 2001; Reysenbach and Shock, 2002; Teske and Stahl, 2002; Inskeep and McDermott, 2005; Sheehan et al., 2005) (see Ward et al., Chapter 1, *in this volume*). Yellowstone National Park, which was established in 1872 and was the first national park in the USA, is a protected environment, so carrying out experiments requires permission and coordination with the Park authorities. At any one time, many experiments are being carried out within the Park {http://www.nps.gov/yell/naturescience/ynpconditions.htm}. Experimental sites in the Park are carefully maintained, which enables return site visits over several years. This is particularly valuable for any experiments that may require data from a time series, although the

springs and channels themselves have been known to change course or exhibit other alterations over time (Brock, 1978).

The alkaline siliceous springs such as Octopus and Mushroom Springs (Fig. 2.1a) are located in the White Creek drainage area of the Lower Geyser Basin and have been studied over many decades by geochemists, ecologists and microbiologists. Therefore, there is extensive literature on aspects of physiology, biogeochemistry and the identification of species in these environments. Indeed the identification, purification and use of thermostable DNA polymerase from *Thermus aquaticus* (a species first identified and studied from these springs by microbiologist Thomas Brock) was instrumental in improving the efficiency of the polymerase chain reaction, which is a routine and widely used technique in molecular biology for the amplification of DNA. The varied geothermal features and their associated microbes are actively studied by many groups aimed at finding uncharacterized bacterial phyla and enzymes with particularly valuable qualities (Brock, 1997; Botero et al., 2004; Podar and Reysenbach, 2006; Sato et al., 2007; Wheeler et al., 2008). These environments also provide a unique opportunity to study the interactions among different prokaryotes and following the pioneering research of Thomas Brock (Brock, 1978), many of his students and other investigators have made significant contributions to our understanding of photosynthetic communities in these extreme environments.

Variously colored (orange to brownish green), often dense microbial mats are formed in the alkaline siliceous hot spring effluent channels (Fig. 2.1b) and these have been extensively documented, examined and categorized (Brock, 1978; Inskeep and McDermott, 2005). The heated water from the source pool gradually cools in the effluent channels, so stable temperature and flow gradients are formed and a variety of microorganisms flourish between particular temperature ranges. Typically, these mats are formed by stable and simple communities of microorganisms such as cyanobacteria (predominantly *Synechococcus* sp.) and green non-sulfur bacteria (GNSLB), such as *Roseiflexus* and *Chloroflexus* sp. There are also less well-characterized heterotrophic anaerobic and aerobic bacteria that are found in these particular mats (Pierson and Castenholz, 1974; Brock, 1978; Ruff-Roberts et al., 1994; Ward et al., 1998; Ward and Castenholz, 2000). The metabolic activities of these mat-forming organisms create stratified layers, within which steep and fluctuating gradients of light, oxygen and nutrients can exist (Stal, 2000; Ferris et al., 2003). Molecular, microsensor and biochemical approaches have been used to measure the metabolism and diversity of bacteria within the microbial mat communities (Revsbech and Ward, 1984; Stal and Caumette, 1994; Kuhl et al., 1998; Kuhl, 2005).

Fig. 2.1a. Octopus Spring in Yellowstone National Park, Summer 2007 (Photograph courtesy Sheila Ingraham Jensen and Melissa Adams)

Fig. 2.1b. Close up of an effluent channel of an alkaline siliceous hot spring showing green/yellow green mats (Photograph courtesy Ilina Bhaya-Grossman, Summer 2006)

The hot spring mats also provide an ideal setting for understanding how thermophilic microbes have adapted to a particular temperature range as well as delving into the question of what sets the upper temperatures of life. Questions such as why certain phyla or ecotypes within a bacterial species are better adapted to a particular temperature or niches and how these processes are evolving are also of wide interest to ecologists, evolutionary biologists and microbiologists. Despite the growing interest in, and concern with, the effects of global warming on different environments and the macro-fauna and flora, surprisingly little attention has been paid to the possible effects on microbes. These hot spring environments, which have stable temperature gradients and where microbes have evolved to deal with high temperatures, offer an opportunity to witness and investigate some of these effects.

In summary, the advantages of using the hot spring microbial communities to further understand the dynamics and function of photosynthetic microbes are: (i) these mats are stable, simple stratified prokaryotic communities, (ii) there is a rich diversity of micro-organisms present at the temperature range from ~40°C to 70°C although it is not as complex as certain environments such as soil, (iii) there are steep and fluctuating gradients of various environmental parameters such as light, oxygen etc. and microsensor data of these parameters can be acquired over a diel cycle or over different seasons, (iv) there is extensive 16S RNA-based phylogenetic analysis of cyanobacteria in the mats, and (v) the ecophysiology of the mats has been studied extensively for decades. These studies set the stage for a deeper exploration of functional genomics in photosynthetic microbes within the context of the environment to which they have successfully adapted. These attractive features can also be used to understand and integrate across different levels of organization, from regulation at the level of single cells to community dynamics.

II. Cyanobacteria as Primary Producers in Hot Spring Mats

Knowledge of microbial mat physiology is crucial for the development of the functional genomics of thermophilic cyanobacteria. Microbial physiology, in turn, is linked to knowledge about the geochemistry, the major 'players' in the community, and their energy requirements and interactions. As mentioned above, mat communities (both at Yellowstone Park and elsewhere) have been studied in great detail by many groups, so it is impossible to provide a comprehensive bibliography within the scope of this review. In this section only a brief summary is provided. For descriptions of microbial biodiversity and the techniques initially used to probe diversity in hot springs the reader is referred to Brock (1978), Stal and Caumette (1994), Ward et al. (1998), Whitton and Potts (2000), Teske and Stahl (2002), Krumbein et al. (2003), and Inskeep and McDermott (2005). The focus of this review will be on selected new techniques and their use to probe function and diversity in the context of mat physiology, structure and diel changes.

A. Mat Structure and Community Members

The dense, high biomass microbial communities in the effluent channels of the hot springs of YNP contain a diversity of microorganisms that range from phototrophs such as cyanobacteria and green non-sulfur bacteria, to heterotrophic anaerobic and aerobic bacteria (Brock, 1978; Ward and Castenholz, 2000; Ferris et al., 2001). At the lower temperature range (40–50°C) mats may be dominated by *Phormidium* or *Plectonema* species, which are filamentous cyanobacteria (Ward et al., 1998; Ward and Castenholz, 2000; Lau et al., 2005). At higher temperatures these are replaced by the unicellular cyanobacterium *Synechococcus* sp. These highly fluorescent rod-shaped cyanobacteria (~1 μm wide and 4–6 μm long) are typically found in the top green layer of mats where they can carry out photosynthesis, while the lower orange-pigmented layers of the mat contain members of the filamentous anoxygenic phototrophs or green non-sulfur bacteria (GNSLB) as well as other less well-identified heterotrophs. Above ~72°C, cyanobacteria cannot survive and the archaea become common, while below 45°C the cyanobacterial populations are grazed extensively by larger copepods, etc. (Brock, 1978). Thus, between 45°C and 72°C, there is a stable temperature gradient and cyanobacteria, predominantly *Synechococcus* spp., are found along this gradient. No evidence of other dominant filamentous cyanobacteria was noted at these higher temperatures

although in other microbial mats it is quite common to have filamentous cyanobacteria as major mat constituents.

Different groups of microbes have been identified in microbial mat consortia, including (i) oxygenic phototrophs, such as the cyanobacteria, (ii) anoxygenic phototrophs, primarily purple and green bacteria that can use hydrogen sulfide as an electron donor (or potentially H_2), (iii) aerobic heterotrophs that generate energy by consuming O_2 to respire organic carbon, (iv) fermenters that use organic carbon or sulfur compounds as electron donors and acceptors, (v) anaerobic heterotrophs, predominantly sulfate-reducing bacteria (SRB), that respire organic carbon using SO_4^{2-} as an electron acceptor and producing H_2S, and (vi) sulfide-oxidizing bacteria (SOB), many of which are chemolithoautotrophs that oxidize reduced sulfur compounds with O_2 or nitrate as electron acceptors, while fixing CO_2 (Stal and Caumette, 1994; Guerrero et al., 2002; Dupraz and Visscher, 2005). So far, a detailed molecular analysis of the functional pathways in the context of carbon cycling in the mat community has not yet been established. The identification of enzymes that are unique to a pathway or key in an organism is the first step towards developing a thorough understanding of critical pathways within microbial communities; such understanding can be tested and further expanded by techniques such as global microarray analysis (Gentry et al., 2006).

B. Metabolism and Diel Cycling Events in the Mat

The chemical composition and pH of hot springs vary substantially by spring location, making generalizations difficult. Furthermore, fluctuations over diel and seasonal cycles are not easily captured. However, the environment in many mat communities is nutrient-poor and especially deficient in nitrogen (N) and phosphorus (P); levels of iron (Fe) and sulfur (S) compounds vary (Brock, 1978; Stal, 2000; Papke et al., 2003; Ludwig et al., 2006). Mats undergo dramatic changes in metabolic processes over the diel cycle, so the organisms in the mat may have evolved a temporally complementary set of metabolisms (van der Meer et al., 2005). During the day, under conditions of high light, the mat becomes highly oxic, with O_2 levels within the matrix of the mat reaching up to 8 times air saturation. The cyanobacteria fix CO_2 via the reductive pentose-phosphate pathway and export a considerable portion of the fixed carbon that was generated (see below). This sustains other members of the microbial community (including the photo-heterotrophs GNSLB, such as *Chloroflexus* and *Roseiflexus*). They also secrete exopolysaccharides that form part of the dense matrix, within which the microbes survive. This exopolysaccharide matrix plays multiple roles, such as serving as a diffusion barrier, providing substrates for growth of heterotrophs, binding certain heavy metals, and controlling calcium carbonate precipitation or lithification in mats (Paerl et al., 2000; Dupraz and Visscher, 2005).

Fixation of CO_2 in mats can also be achieved by organisms other than the cyanobacteria, and by processes other than the reductive pentose phosphate pathway (see Chapters 3 and 9 in this volume). For example, based on carbon fractionation data, the GNSLB appear to fix CO_2 via the novel cyclic 3-hydroxypropionate pathway (Holo, 1989; van der Meer et al., 2005). This pathway, first discovered in *Chloroflexus*, proceeds in a cyclic manner from acetyl-CoA to 3-hydroxypropionate, which may be released under certain conditions (Strauss and Fuchs, 1993; Herter et al., 2001; Ishii et al., 2004). While the ATP that drives the pathway is likely to come from anoxygenic photosynthesis when light levels are low (the mat is anoxic with photosynthesis being driven by low levels of excitation energy), the source of electrons used in the reductive steps are not known. The dominant carbon compound exported by cyanobacteria under conditions of photoautotrophic growth is acetate, while at night they export fermentation products (mostly ethanol and formate) generated by from the breakdown of polyglucose (Staley, 1997; Teske and Stahl, 2002; van der Meer et al., 2005). The mat reaches different characteristic carbon/energy/redox states during the night and day. As the light intensity declines during the late afternoon, the relative ratio of O_2 evolution by photosynthesis to O_2 consumption by respiration begins to decline. Once O_2 uptake exceeds O_2 evolution, the mat becomes anoxic, so that, over a period of minutes, the mat must transition from a consortium functioning under oxic conditions to one that operates anoxically.

C. Phototaxis, Vertical Migration and Positioning in the Mat Environment

In the mat environment, which is stratified in the vertical dimension and is densely packed with phototrophic organisms such as *Synechococcus* sp. (oxygenic phototrophs), GNSLB (such as *Chloroflexus* and *Roseiflexus* sp.) and green sulfur bacteria (such as *Chlorobium* sp.), access to optimal light conditions is obviously of prime importance. Although this aspect has not been studied intensively, a few interesting concepts and measurements are worth noting regarding ordered structure in the mat and phototaxis. Ramsing et al. (2000) used micro-sensors and observed that oxygenic photosynthesis peaked in the uppermost 100–200 μm region of the mat in the morning, but by afternoon this peak had shifted into the deeper layers. Interestingly this correlated with a vertical shift in discrete bands of auto-fluorescence emanating from the *Synechococcus* populations. They were unable to find strong evidence of any diel migration of particular species within the mat and analysis of cells from vertical thin sections revealed that the rod shaped *Synechococcus* sp. were randomly oriented in various parts of the mat with one exception. Around noon, when light levels are high and may be damaging to the photosynthetic apparatus,, there was a narrow band of cells that assumed an upright position about 400–800 μm below the surface. This could suggest that cells have the ability to change their orientation in response to light levels and raises the question of how this is perceived by the cell and how the light signal is transduced to the cell surface to cause a change in cell orientation. These results also correlate with related work of Ramsing et al. (1997) showing that particular *Synechococcus* sp. isolates show light-dependent motility. This phototactic motility was a complex phenotype, in which cells showed different rates of motility depending on the light intensity, and movement was observed both toward and away from the light source. Different strains may show differing motility phenotypes, although all strains appeared to produce copious exopolysaccharide trails as they moved along the surface. This preliminary report did not describe motility as a function of light quality, but our unpublished results show that *Synechococcus* sp. isolates from the mat can move rapidly towards white or red light (D. Bhaya, unpublished) (Fig. 2.2). Motility within the mat environment may be quite advantageous in terms of optimizing photosynthesis, since light is strongly attenuated by the mat and so gliding up into more lighted areas of the mat may be an advantage; conversely, when light is damaging at certain times of the day, it may be optimal to move deeper into the mat and avoid damage to the photosynthetic apparatus. Thus, one might expect these mat cyanobacteria to be able to sense light direction as well as light quality, but this has not yet been examined.

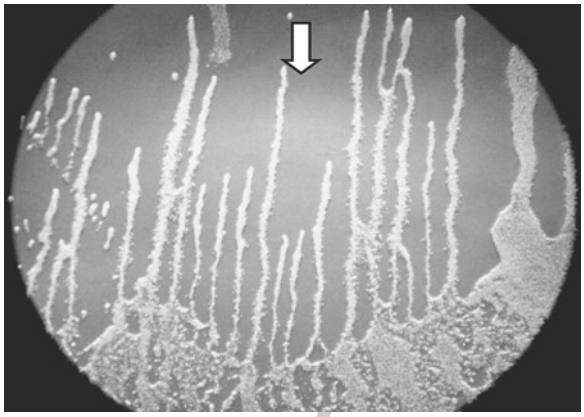

Fig. 2.2. Phototactic motility of *Synechococcus* OS-B′ cells on low concentration agarose plates with a white light source positioned directionally (*arrow*). Note thin fingers of cells that have moved towards the light (Photograph taken after 48 h of placing on agarose surface)

Our preliminary results show that the *Synechococcus* isolates contain all of the genes required for pilus-dependent motility and thus one might expect that motility is a surface-dependent phenomenon in these cyanobacteria. Light-dependent motility has been characterized in unicellular model cyanobacteria and photoreceptors and other components of the motility apparatus have been identified (Bhaya, 2004; Yoshihara and Ikeuchi, 2004). Extending these studies to environmentally relevant cyanobacteria that have evolved to cope with fluctuating light conditions in a mat community is likely to uncover novel features of photo-movement and perhaps of social communication.

Gliding motility, which is strongly influenced by light, has also been characterized in filamentous thermophilic cyanobacteria such as *Oscillatoria* and *Phormidium*, which are components of hot

spring microbial mats and show vertical migration patterns. Other filamentous bacteria that can be found in hot spring mats (such as *Chloroflexus*, *Heliothrix* and the purple sulfur bacterium, *Chromatium*) also show light-dependent motility but the molecular mechanisms for these movements are still not well understood. Some of these photo-movements may be UV-A irradiation dependent suggesting that complex photo-sensory mechanisms may have evolved to protect the phototrophs in mats (Richardson and Castenholz, 1987; Bebout and Garcia-Pichel, 1995; Castenholz and Garcia-Pichel, 2000). Indeed, some cyanobacteria produce UV-protective pigments such as scytonemin, for which part of the biosynthesis pathway has been elucidated (Castenholz and Garcia-Pichel, 2000; Soule et al., 2007).

D. Molecular Techniques to Study Community Structure and Composition

Analysis and examination of the microorganisms in the mat were initially based on classic microbiological tools such as attempts to cultivate microbes under specific enrichment conditions and/or microscopy (Brock, 1978; Ward et al., 1998). These methodologies are used by many laboratories and are still considered to be important but not necessarily comprehensive for identification. Thus, early on the use of these techniques led to the identification of cyanobacterial species in the mat that were categorized as unicellular *Synechococcus* sp. Similarly, other major microbial components of the mat were identified to be filamentous *Chloroflexus* sp. and the related *Roseiflexus* sp. by Castenholz and others (Pierson and Castenholz, 1974). However, with the development of simple but powerful molecular tools, namely 16S RNA clone sequencing and denaturing gradient gel electrophoresis (or DGGE) to identify bacterial species, the picture that emerged of the microbial populations in the mat got considerably more complex. Both of these techniques are still widely used for the identification and classification of bacterial populations in the environment because of the ease of use (in the case of DGGE) and the ability to use 16SRNA and 16S–23S internal transcribed spacer (ITS) sequence to build phylogenetic relationships (Stahl et al., 1985; Ward et al., 1998). Use of these techniques in the mat environment has shown that the mat contains a very large number and diversity of unique 16S RNA sequences (Ward et al., 1990; Ferris and Ward, 1997). Furthermore, these 16S RNA sequences did not match those of the most readily cultivated isolates from the mat. DGGE studies showed that the distribution of 16S RNA gene variants from cyanobacteria varied along the temperature gradients of the mat. An example of this type of distribution was the characterization of closely related 16S RNA sequences, designated A″, A′, A, B″ and B, that were detected along a temperature gradient ranging from ~70°C to ~50°C (Ferris and Ward, 1997; Ward et al., 1998). Interestingly, 16S RNA sequences of the *Synechococcus* populations were also found to vary predictably along the vertical transect of the mat and appeared to be correlated with the presence of differentially fluorescent *Synechococcus* populations (Ramsing et al., 2000). These detailed molecular studies coupled with other measurements (such as light availability within the depth of the mat) substantiate the view that cyanobacterial (*Synechococcus* sp.) communities within alkaline siliceous mats have well-defined distributions of 16S rRNA and this diversity appears to correlate with established environmental gradients (Ward and Cohan, 2005; Ward 2006a). Questions, some of which will be addressed in subsequent sections (also see Chapter 1), that follow from these results include:

1. Do these *Synechococcus* populations with different 16S sequences have measurable functional differences that can be probed by molecular approaches, and if so, how?
2. Have certain populations "adapted" to a particular niche, and what functional differences have developed? How does this correlate with the regulation of photosynthesis and related processes in the mat?
3. Does this microbial mat system provide a good model system to understand how populations evolve and adapt to environmental fluctuations?
4. Can one begin to understand how different members of a community are interacting metabolically and are sharing resources at a molecular and metabolic level?
5. Can the use of functional genomics and metagenomics clarify notions on the still evolving concept of a "bacterial species" (Konstantinidis and Tiedje, 2004; Gevers et al., 2005; Bhaya et al., 2007; Achtman and Wagner, 2008; Doolittle and Zhaxybayeva, 2009)?

III. Comparative Genomics and Transcriptomics to Probe Function of Closely Related *Synechococcus* sp.

A. Genomic Content and Architecture of Two Related Synechococcus Isolates

To address some of the questions posed above, we took advantage of two available *Synechococcus* isolates (see Ward et al., Chapter 1 in this volume). *Synechococcus* OS-A (*Syn* OS-A) was isolated by dilution culturing (filter cultivation approach) from samples derived from a high-temperature region of the mat (58–65°C), while *Synechococcus* OS-B′ (*Syn* OS-B′) was isolated from low-temperature mat samples (51–55°C) (Allewalt et al., 2006). Allewalt et al. (2006) demonstrated that *Syn* OS-A also showed the highest temperature optimum and upper limit for photosynthesis. This provided partial support to the concept that isolates from different temperature ranges of the mat had growth and photosynthetic consistent with their location and suggested that these gross measures of "adaptation" may also be reflected at the gene/genomic level.

We acquired complete genome sequences of *Syn* OS-A and *Syn* OS-B′ by means of shotgun sequencing (Bhaya et al., 2007). Both isolates came from the Octopus Spring mats and contained circular genomes of approximately the same size (~3.0 Mb), exhibited a relatively high G+C% content of 60.3 and 58.5, and included 2,892 and 2,933 predicted coding sequences, respectively (Fig. 2.3). Comparative analysis of these two closely related and mat-dominant cyanobacteria revealed that *Syn* OS-A and *Syn* OS-B′ each have two identical copies of the rRNA genes in their respective genomes. A comparison of the *Syn* OS-A and *Syn* OS-B′ 16S rRNA sequences showed 96.4% identity (i.e. 3.6% differences in sequence). Previous research by Ward and colleagues had demonstrated that (i) *Synechococcus* sp. dominating the mat were substantially different from the readily cultivated *Synechococcus lividus* strains with as much as 8–10% difference in 16S rRNA sequence (Ferris et al., 1996; Ward et al., 1998 and references therein), whereas (ii) five of the predominant *Synechococcus* genotypes (A, A′, A″, B and B′) were more closely related with < 3% difference at the 16S rRNA sequence level). The sequences also matched previously identified 16S rRNA sequences from the mat and confirmed that we were working with isolates that were dominant in the mat. Similar comparative analyses regarding phylogenetic relationships based on ribosomal sequences of related marine cyanobacterial ecotypes (*Prochlorococcus* and *Synechococcus* sp.) that have been sequenced have also been carried out (West et al., 2001; Rocap et al., 2002; Ernst et al., 2003; Ahlgren and Rocap, 2006).

A comparison of the coding sequences of the two *Synechococcus* genomes showed that they share a large fraction (~83% based on bidirectional best BLAST scores) of their coding sequences with a high identity between putative

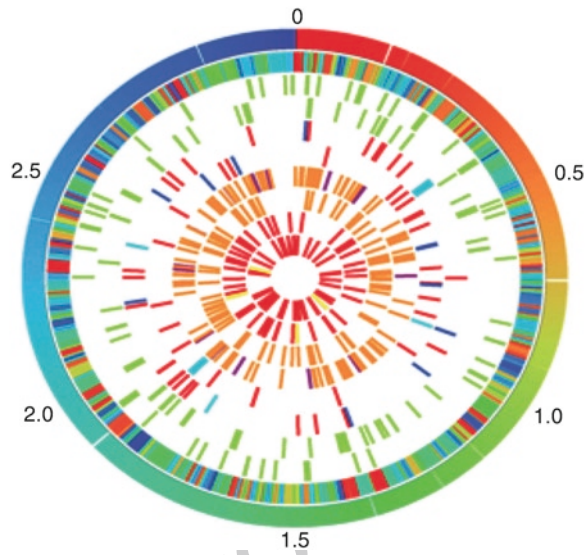

Fig. 2.3. Genomes of *Synechococcus* OS-A and OS-B′. Circle 1 (outermost); *Syn* OS-B′ genes assigned pseudo-spectrum colors based on relative position along the length of the genome, with genes nearest to the putative origin of replication colored red and genes most distal from this origin colored blue. Circle 2; putative homologs in *Syn* OS-A are assigned the same color as is presented in circle 1 for the *Syn* OS-B0 genes; the scrambling of colors in circle 2 is a reflection of the striking lack of relative conserved large-scale gene order between the genomes of Syn OS-A and OSB′. Circles 3 and 4; photosynthesis genes (*green*). Circles 5 and 6; nitrogen fixation genes (*light blue*), urease genes (*dark blue*) and fermentation genes (*red*). Circles 7 and 8; insertion elements (*orange*) and ISSoc13 transposases (*purple*). *Circles 9 and 10*; tRNAs (*red*) and rRNAs (*yellow*). Circles 3, 5, 7 and 9 show the relative gene positions on the genome of Synechococcus OS-B0. Circles 4, 6, 8 and 10 show the relative gene positions on the genome of Synechococcus OS-A. (Modified from Bhaya et al., 2007)

[AU7]

orthologs (~87% amino acid identity on average) (Bhaya et al., 2007). This is consistent with the information that these cyanobacterial isolates are morphologically identical and closely related at the 16S RNA sequence level (which is a well-accepted robust marker for phylogenetic relationships). Despite this, at the level of whole-genome architecture a comparison of *Syn* OS-A and OS-B′ genomes shows a marked lack of synteny or conserved, large-scale gene order, indicating an extensive history of rearrangement events (Fig.2.3). Comparison of complete genomes of closely related bacteria usually reveals extensive synteny, so this is a surprising deviation that requires further analysis. Regions of co-linearity between the *Syn* OS-A and OS-B′ genomes were short with the largest region of conserved gene order between the two isolates being ~32 Kbp (containing the genes required for nitrogen fixation).

Genome rearrangements and recombination events are often mediated by transposons or phage and could play an important role in evolution (Rocha, 2004, 2008). Both *Synechococcus* OS-A and OS-B′ genomes contain many transposon-like or insertion sequence (IS) elements (~100 intact IS genes on each genome as well as many IS gene fragments (Nelson, W., Heidelberg, J., and Bhaya, D, unpublished). *Synechococcus* OS-B′ contains 17 identical copies of an IS4 family of transposase genes (ISSoc13 or Interpro ID 002559) that are completely absent in the *Synechococcus* OS-A genome and that may still be active in *Synechococcus* OS-B′ (Fig. 2.3, Circle 7). We found that the IS elements are not always located at the borders of re-arranged regions, so their role in the large-scale gene rearrangements cannot be easily assigned (Parkhill et al., 2003). It has been estimated that these thermophilic *Synechococcus* isolates contain a higher percentage of transposable elements than expected based on genome size (difference at the 16S rRNA sequence level *Prochlorococcus* which has a small or "minimal" genome does not contain any transposable elements (Zhou et al., 2008). The role of IS elements in cyanobacteria is largely unknown but now that a large number of cyanobacterial genomes have been sequenced such a study may be quite fruitful, particularly since IS elements are known to be responsible for significant and rapid changes in genome architecture.

B. Functional Categories and Unique Genes in Genomes and Their Roles in Adaptation

The *Synechococcus* OS-A and OS-B′ genomes contain genes encoding complete sets of proteins required for, among others, photosynthesis, the biosynthesis of glycolate (Bateson and Ward, 1988), glycogen (Bateson and Ward, 1988; Konopka, 1992), and sulfolipids (Ward et al., 1994), and fermentative and respiratory metabolisms (Nold and Ward, 1996). We also identified genes required for the biosynthesis of Type IV pili and photoreceptors associated with phototaxis, which fits with earlier reports of motility of *Synechococcus* cells in the mats (Ramsing et al., 1997; Bhaya, 2004). However, the presence and activity of a functional pathway for nitrogen fixation in *Synechococcus* OS-A and OS-B′ were unexpected, as previous attempts had failed to measure nitrogen fixation in the mats at higher temperatures (see Section III.D).

To identify functional differences between *Synechococcus* OS-A and OS-B′, we examined subsets of genes unique to each of these isolates. There are ~400 and ~500 isolate-specific genes in *Synechococcus* OS-A and OS-B′, respectively, of which about half are annotated as either "hypothetical" or "conserved hypothetical". Within this set we identified examples of genes encoding proteins with known functions that are present on one but not on the other genome. Only the *Syn* OS-B′ genome contains genes for the synthesis and metabolism of cyanophycin, a N storage compound. Cyanophycin synthetase is the enzyme that synthesizes cyanophycin non-ribosomally from aspartate and arginine; and cyanophycinase can degrade the polymer to provide the cell with a source of N when needed (Simon, 1987). Cyanophycin levels vary with growth conditions, but can be high in stationary-phase cultures or under conditions in which the growth potential of the cell declines because of a limitation for other nutrients such as sulfate or phosphate. Cyanophycin has also been implicated in the integration of carbon and nitrogen metabolism in unicellular and filamentous cyanobacteria (Mackerras et al., 1990). The ability to store N in the form of cyanophycin granules by *Syn* OS-B′ suggests that it may experience fluctuating nitrogen levels. Since *Synechococcus* OS-A does not appear to have either of these genes it would unable to store

nitrogen as effectively as *Synechococcus* OS-B′ but the significance of this in the context of its environment or 'niche' is not clear.

Another interesting example of genome-specific functionality is the presence of a large 8 Kbp region on the *Synechococcus* OS-B′ genome, which contains ten genes (*phn* genes) responsible for the transport and metabolism of phosphonates. This could enable the organism to utilize phosphonate (compounds in which a carbon-oxygen-phosphorus bond is replaced by a direct carbon-phosphorus linkage) as a source of phosphorus in addition to phosphate. Phosphonates are relatively inert, stable compounds and may have preceded phosphates in the early atmosphere when oxygen levels were low. Although the importance of biogenic phosphonates in the terrestrial biosphere has not been established, phosphonate levels are high in the marine environment (Quinn et al., 2007). The entire *phn* gene cluster is missing in *Synechococcus* OS-A, but the region flanking the *phn* cluster is syntenic between the *Synechococcus* OS-A and OS-B′ genomes, indicating that the *phn* gene cluster was either recently acquired by *Synechococcus* OS-B′ or lost in the *Synechococcus* OS-A lineage. There is evidence suggesting that operons required for phosphonate uptake and utilization may be acquired through lateral gene transfer events in prokaryotes (Huang et al., 2005). Recently, genes for phosphonate utilization have been identified in metagenomic studies of marine, oxygenic photosynthetic prokaryotes, but the *phn* operon is not universally found in cyanobacteria, perhaps reflecting the different availability of phosphonates in various environments (Palenik et al., 2003; Dyhrman et al., 2006).

One approach to explore these differential abilities between two closely related cyanobacteria is to use axenic cultures of both isolates under defined laboratory conditions (e.g., low and high N conditions) to explore the benefits of storing cyanophycin or the advantages of being able to use phosphonate as a P source. We have pioneered such an approach with *Synechococcus* OS-B′ with the rationale that some questions are more powerfully addressed with axenic isolates while other questions are much better explored with *in situ* techniques (both will be addressed later in this section).

C. Axenic Cultures to Study Ecologically Important Questions

With the advent of high-throughput genome sequencing and metagenomics there is a flood of information about the genetic repertoire of various bacteria. Although this is a powerful information database that has been exploited in many ways (see Section IV), it still is a big leap to advance from a dictionary of genes in the genome (or the environment in the case of a metagenomics approach) to an understanding of the biology of dominant players in any particular environment. One way to achieve this deeper insight is to be able to work with axenic isolates from the environment of interest. This is not always feasible since only a very small percentage of bacteria can be axenically grown in the laboratory. However, for strains where axenic growth is possible, it opens the door for a number of exciting new areas for research since one can combine *in situ* approaches with more detailed experiments under controlled conditions.

The initial experiments to check if isolates derived from different temperature regions of the mat would show different physiological characteristics consistent with their location were carried out with isolates that were uni-cyanobacterial but not axenic (Allewalt et al., 2006). Although this may not have significantly impacted the interpretation of results, axenic cultures are preferable. Towards that end we repeatedly streaked the enriched cultures on plates at low agarose concentration and placed these plates in directional light (Fig. 2.2). Since *Synechococcus* OS-B′ cells are phototactic, we were able to separate them away from non-motile heterotrophs. We used 16S ribosomal sequencing, growth on nutrient-rich plates (to test for slow-growing contaminants), and phase-contrast microscopy to ensure that the culture was axenic. Two examples of approaches with axenic cultures are described to demonstrate how it has provided insight into the physiology and acclimation of phototrophs to light and nutrients.

1. Acclimation to High and Fluctuating Light Levels

To understand how thermophilic cyanobacteria in microbial mats can respond to fluctuating environmental parameters such as light, we used axenic isolates of *Synechococcus* OS-B′. The

ability to monitor the growth and other key parameters of this isolate under environmentally relevant temperature conditions, as well as prior knowledge about how cyanobacterial cells respond to high-light conditions at the biochemical and gene regulation levels, allowed us to assess the physiological state of the cells. Surprisingly, even though the microbial mats may contend with very high irradiances during the day, *Synechococcus* OS-B′ did not appear to cope well with continuous high-light conditions. Axenic cultures of *Synechococcus* OS-B′ grew optimally at relatively low light fluence rates of between 75 and 130 µmol photons m^{-2} s^{-1} as shown by their blue-green color and characteristic absorption spectrum, while cells grown at higher irradiances were chlorotic and lost phycobiliproteins (Fig. 2.4). Cells grown in continuous light at an irradiance of 400 µmol photons m^{-2} s^{-1} stopped growing after 3 days and died, but it is important to note that these culture conditions do not replicate the mat conditions. Within the mat, cells are extensively packed, there may be protective pigments present, and light is strongly attenuated, particularly in the blue and red regions of the spectrum (Kuhl et al., 1997), so cells may be experiencing a very different light regime under these conditions.

Photosynthetic organisms acclimate to the damaging consequences of the absorption of excess light energy in a number of ways, including by a marked decline in light-harvesting pigments, changes in the level and composition of photosynthetic reaction centers, the development of sinks to efficiently remove electrons from the electron transport chain, the establishment of mechanisms to eliminate reactive oxygen species (ROS) that might accumulate, and the ability to repair damaged cellular components. We attempted to measure some of these parameters in the axenic isolates. At 200 µmol photons m^{-2} s^{-1}, we noted several responses that had previously been associated with acclimation of cyanobacteria to high light levels such as strong bleaching of cells, reduced levels of phycobilisomes and chlorophyll, and elevated levels of carotenoids. Quantification of the abundance of transcripts encoding the polypeptides that make up the PBS was consistent with this observation. These results suggest that at higher light irradiances there is a reduction in the absorbance cross section of the light-harvesting antenna. Other parameters tested also suggested that the cells are acclimating to high light in a number of ways. Interestingly, 77 K fluorescence emission spectra suggest that *Synechococcus* OS-B′ accumulates very small amounts of photosystem II relative to that of photosystem I. This ratio was further decreased at higher growth irradiances, which may reflect potential photo-damage following exposure to high light intensity. High light intensity also reduced levels of transcripts encoding phycobilisome components, particularly for CpcH, which is a 20.5-kDa rod linker polypeptide. There was enhanced transcript abundance of genes encoding terminal oxidases, superoxide dismutase, tocopherol

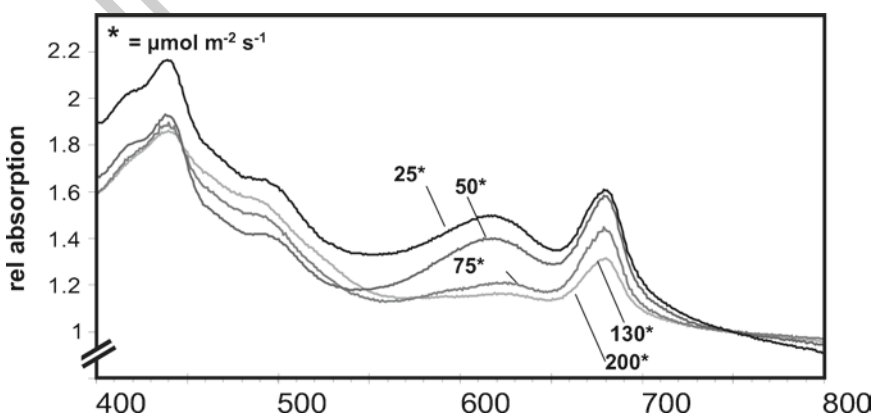

Fig. 2.4. Whole-cell absorption spectra of *Synechococcus* OS-B′ absorption at different light irradiances. Wavelength is on the X axis and absorbance on the Y axis. (Modified from Kilian et al., 2007)

cyclase, and phytoene desaturase. Genes encoding the photosystem II D1:1 and D1:2 isoforms (*psbAI* and *psbAII*/*psbAIII*, respectively) were also regulated according to the light regimen (Kilian et al., 2007).

2. Acclimation to Nutrient Limitation

Since the *Synechococcus* OS-B′ but not the OS-A genome harbors the *phn* gene cluster, which might allow OS-B′ to grow on phosphonate (Phn) as a sole phosphorus source, we grew axenic cultures in medium lacking Pi as well as on different Phn sources (Adams et al., 2008). Cells continued to progress through 3–4 cell divisions after Pi was removed from the growth medium suggesting that the storage and metabolism of intracellular poly P may be one mechanism that enables the organism to cope with low exogenous Pi. Consistent with this possibility, we found that there are large poly-P pools present in *Syn* OS-B′ (M. R. Gomez-Garcia, A. Grossman, and D. Bhaya, unpublished). Pi limitation of *Syn* OS-B′ was found to elicit the accumulation of extracellular alkaline phosphatase activity and increased levels of transcripts encoding several putative phosphatases. The gene encoding PhoX of *Syn* OS-B′ was most highly induced (based on Q-PCR measurements) and PhoX may be responsible for most of the extracellular phosphatase activity assayed during P deprivation. In addition, transcripts encoding the high-affinity ABC-type Pst transport systems as well as the genes in the *phn* operon were induced. Many Pho regulon genes that are present in the hot spring cyanobacteria, including *phoX* and the *phn* gene cluster, are only present on a few other cyanobacterial genomes, which suggest environmental or niche-specific adaptation of P metabolism (Adams et al., 2008).

Even though *phn* transcripts of *Syn* OS-B′ accumulated rapidly in response to P starvation, *Syn* OS-B′ was unable to effectively use the methyl Phn (MePhn) as a sole P-source until the cells had acclimated for approximately 3 weeks. Although this is a somewhat unexpected result, it is possible that the long and variable acclimation phase during which cells grow very slowly (Phn is supplied as a sole source of P) may be the consequence of steps that are limiting in Phn degradation. This long acclimation period has been noted in various other bacteria, such as *E. coli*, as they acclimate to different phosphonate sources (Wanner, 1994). For instance, induction of transport systems may take varying amounts of time to enable the transport of different Phn compounds, or Phn may be toxic to certain cellular processes. After ~20 days, the cells started to grow more rapidly and once the cells had acclimated, they initiated growth immediately upon transfer into fresh medium containing MePhn as a sole source of P and attained a doubling time similar to that of cells using Pi as their sole P-source. Currently we do not know if the MePhn acclimation phenomenon represents a genetically-based mechanism or whether a small subpopulation becomes responsive during the lag phase and ultimately outgrows the cells that were unable to acclimate. The *phn* gene cluster has been found in many microorganisms isolated from a variety of ecosystems, including marine ecosystems, in which Phns constitute a substantial fraction of dissolved organic P of the total P pool (Clark et al., 1998). The capacity to utilize Phns when other sources of P are limiting could confer an adaptive advantage to the *Syn* OS-B′ cells in an environment where Pi is scarce. However, it is not known whether the available Pi fluctuates on a temporal or spatial scale. Pi starvation on a daily or seasonal basis may allow *Syn* OS-B′ to acclimate to low P conditions, which includes an increased capability for utilizing phosphonates to satisfy the P demand. If fluctuations in the Pi concentration are frequent, the cells may remain in the acclimated state even when availability is elevated over a short time interval.

D. In Situ Transcriptomics to Probe Diel Cycles

In the course of our comparative analysis of the *Syn* OS-A and OS-B′ genomes, we identified a 30 Kbp region that harbored genes required for nitrogen fixation (Rubio and Ludden, 2005; Bhaya et al., 2007). This was surprising since most reports of nitrogen fixation (N_2-fixation) in hot springs suggested that it occurred only at low-temperature regions of the mat, possibly catalyzed by filamentous heterocystous cyanobacteria (Stewart, 1970; Belay et al., 1984). To provide strong experimental evidence that these genes were indeed functional and that N fixation was occurring in the mats, we developed an *in situ* transcriptomics approach: *nif* gene-specific primers were used for quantitative RT-PCR (qPCR) on

RNA samples isolated from the mats at different times of the diel cycle (Steunou et al., 2006); the *nif* genes were expressed *in situ* in the mat (Steunou et al. 2006, 2008) but were only detected at night and into the early morning (i.e., the period of time when the mat was anoxic) (Fig. 2.5). During the diel cycle collections, light and oxygen levels were also measured. Nitrogenase activity (monitored by acetylene reduction assay) and nitrogenase subunits (monitored by Western blot analyses) in mat samples were also detected in the evening and early morning. This suggests that nitrogenase activity is restricted to certain parts of the diel cycle. Since nitrogenase activity is irreversibly inactivated by oxygen and N fixation is energetically expensive (requiring at a minimum 16 ATP molecules per N fixed), it is important to determine how nitrogenase activity is regulated and how energy is made available for this process. This required us to accurately monitor other genes of interest, e.g., genes involved in photosynthesis, fermentation and respiration. Thus, our approach was to monitor mat metabolism and gene regulation, and to correlate these data with environmental parameters such as light, pH and nutrient availability to build a model of how various energetic processes vary over a diel cycle (Fig. 2.5).

A conceptual model showing the factors that influence nitrogenase activity over the diel cycle in hot spring mats has now been developed (Steunou et al., 2008). During the day, the upper few millimetres of the mat are supersaturated with O_2 because of cyanobacterial oxygenic photosynthesis. Under these conditions, the *nif* genes are not expressed. As irradiance falls towards the end of the day, the O_2 concentration in the mat also drops because of (i) a decline in photosynthetic O_2 evolution and (ii) a sustained or increased respiratory consumption of O_2 by cyanobacteria and other microbes in the community. At the same time, both *nif* and specific fermentation transcripts increase, corresponding polypeptides are synthesized and assembled into active complexes, and N_2 fixation can be initiated. By the time the level of nitrogenase becomes maximal and its activity is fully established, photosynthetic energy production has decreased substantially due to the absence of light. Oxygen is largely depleted in the upper 0.1–0.2 mm of the mat due to respiration and re-oxidation of reduced compounds. Thus, the only source of energy for cyanobacterial

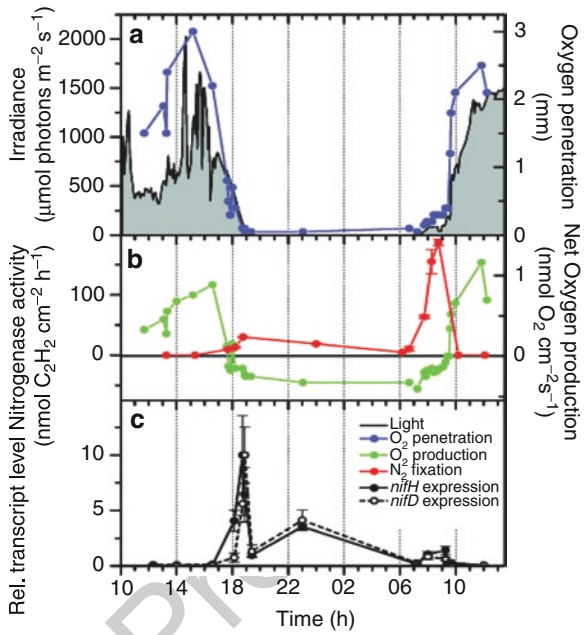

Fig. 2.5. In situ nitrogenase activity, levels of NifH subunit and transcripts encoding NifH and NifD, and oxygen penetration and net production over the diel cycle in the microbial mat of Mushroom Spring in September 2005. (**a**) Incident down welling irradiance (micromol photons m^2 s^1) and O_2 penetration (mm) in the hot spring mat over the diel cycle. (**b**) In situ nitrogenase activity (nmol ethylene per cm^2 per h) and net O_2 production (nmol O_2cm^2 s^1) over the diel cycle. (**c**) Relative abundance of transcripts encoding the nitrogenase subunits NifH and NifD over the diel cycle. Curves and colors are defined in the inset of panel C. Nitrogenase activity and gene expression data points represent means ± sd (N = 3). (Modified from Steunou et. al., 2008) [AU3]

N_2-fixation in the anoxic part of the mat is derived from fermentation of organic carbon accumulated during the preceding day. A similar scenario has been proposed for hypersaline mats (Bebout et al., 1993).

In the morning, as light levels increase, nitrogenase activity increases in parallel with photosynthetic activity. The increased nitrogenase activity is not accompanied by increased nitrogenase transcript or protein levels, but reflects elevated production of ATP and reductant. The mat remains largely anoxic with net O_2 consumption during the early hours of the day period, until increasing irradiance drives the rate of photosynthetic O_2 evolution above the rate of respiratory O_2 consumption. As O_2 begins to accumulate in the mat, the nitrogenase activity is strongly inhibited. Since these processes are interlinked, other

factors such as day length, light intensity and accumulation of stored carbon compounds such as polysaccharides (Steunou et al., 2008) may affect the dynamics of the metabolism over a diel cycle.

Integrating eco-physiological methods to quantify the mat microenvironment and nitrogenase activity *in situ*, together with transcriptomics give new insights into the complex dynamics of N_2 fixation in hot spring cyanobacterial mats. N_2 fixation in thermophilic *Synechococcus* sp. is closely linked to their energy metabolism and photosynthesis, which also shows pronounced shifts during a diel cycle. The next challenge with this system will be to develop an understanding of the functional interactions with other microbes in the mat community, as well as the variants in the populations of different cyanobacteria in the mat.

The integrative approach presented here attempts to tie together three methodologies: eco-physiology, *in situ* microsensor measurements of activity, and transcriptomics. This approach requires, as a prerequisite, a detailed knowledge of the physiology of the microbial mat system and the genomic content of some of the dominant players in the microbial mat. This approach sheds light on the metabolic dynamics of the cyanobacterial populations, which are well-adapted to the diurnal fluctuations in light and oxygen.

E. Recent Acquisition or Loss of Nutrient Utilization Pathways

As the *phn* cluster and the cyanophycin synthesis/metabolism pathway are only found in one of the two *Synechococcus* isolates (OS-B′) we further explored the possibility that nutrient utilization functionality can be gained and lost in related lineages. In this context, we noted significant differences between the *Synechococcus* OS-A and OS-B′ isolates for genes required for the putative uptake and utilization of urea. Some cyanobacteria can utilize urea as a source of nitrogen, which requires the enzyme urease and a dedicated transport system (Luque et al., 1994; Collier et al., 1999). *Synechococcus* OS-A has one genomic region encoding urease (*ureA1B1C*) and accessory factors (*ureEFG1D1*) (Cluster 1 urease) (Fig. 2.6). The genes of Cluster 1 urease are flanked by transposons. This cluster is lacking in *Synechococcus* OS-B′ genome but the synteny between *Synechococcus* OS-A and OS-B′ in the region flanking Cluster 1 is maintained, suggestive of a relatively recent gain of genes by *Synechococcus* OS-A. Alternatively, it might represent a loss of genes from *Synechococcus* OS-B′ if the shared ancestor of *Synechococcus* OS-A and OS-B′ contained these genes. Without further analysis it is not possible to distinguish between these possibilities, but the presence of

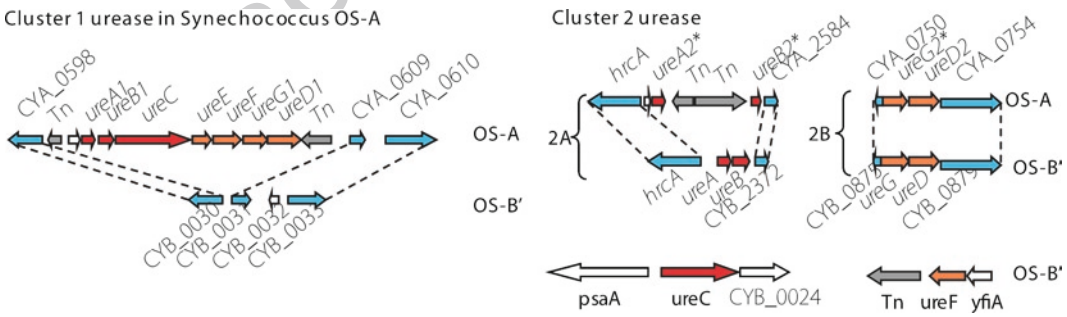

Fig. 2.6. Urease genes in *Synechococcus* OS-A and OS-B′. *Left Panel*: *top row* represents the Cluster 1 urease and flanking genes of *Synechococcus* OS-A; the bottom row shows homologs of the flanking genes in *Synechococcus* OS-B′. *Right panel* shows: locations of the Cluster 2 urease genes in *Synechococcus* OS-A and OS-B″. In *Synechococcus* OS-A, ureB2 is inactivated by a transposon; ureA and ureG2 contain frameshift mutations (inactive genes are designated by * symbols). Genes encoding urease (ureA, –B and –C) are shown in *red*; accessory factor genes are shown in *orange*; transporter genes in *green*; putative transposons are shown in *gray*; flanking gene homologs in *blue* and other genes in *white*. Syntenic regions are indicated by *dashed lines*

flanking transposons suggests a gain of function in OS-A. *Synechococcus* OS-B′ contains a second set of urease genes, which are scattered at five different regions of the genome, and remnants or transposon-mediated interruptions of these genes are found in *Synechococcus* OS-A. In *Synechococcus* OS-B′ these Cluster 2 urease genes are intact. These obvious differences in gene content underscore the possibility that both gene gain and loss are ongoing events in the *Synechococcus* populations. We envisage a possible scenario, in which there was a relatively recent acquisition of Cluster 1 urease by *Synechococcus* OS-A as suggested by the transposons flanking this region. Subsequently there may have been a progressive loss of functionality of the original urease cluster (i.e., the Cluster 2 urease), which is still maintained by and likely to be functional in *Synechococcus* OS-B′. These observations raise obvious questions to be addressed in the future:

(a) If *Synechococcus* OS-A acquired this cluster by "recent" lateral gene transfer, can we date this process and possibly identify the donor organism?
(b) Could further metagenomic analyses of the populations reveal other variants, with different gene arrangements, suggestive of fluid functionalities that are evolving in this environment?
(c) The more recently acquired Cluster 1 may provide selective advantage to *Synechococcus* OS-A, which the Cluster 2 urease did not provide. Can the use of axenic strains of these organisms allow us to test some of these hypotheses?
(d) Are there other examples of such loss/acquisition of nutrient utilization capacity? If so, is this a general theme of phototrophs that are evolving in fluctuating environments?

F. Genomic Rearrangements/Fusions in the Context of Photosynthesis

The high-light-inducible proteins (Hlips) of cyanobacteria are members of an extended CAB/ELIP/HLIP superfamily of distantly related polypeptides, which have between one and four transmembrane helices (TMHs) (Green, 1995; Jansson, 1999). The best-studied members of this superfamily are the membrane-integral light-harvesting chlorophyll *a/b* binding proteins (LHCs or also known as CABs) which are abundant in the chloroplast (see Chapter 11, this volume). They bind chlorophylls and carotenoids, associate with both photosystem I and II, and are regulated by various environmental factors including light levels and nutrients. The LHC proteins are well-conserved polypeptides with three TMHs and they may have evolved by gene duplication of an ancestral gene encoding a single TMH (Kuhlbrandt et al., 1994).

The Hlips are small membrane proteins consisting of a single TMH, similar to those in the LHC proteins that are induced by various stress conditions including illumination with UV-A and high light intensity, and are considered ancestral to the LHC proteins (Green and Pichersky, 1994; Dolganov et al., 1995; Funk and Vermaas, 1999; He et al., 2001). A *Synechocystis* sp. PCC 6803 mutant in which all four *hli* genes were inactivated died under high light conditions, suggesting that Hlips are crucial for acclimation to high light (He et al., 2001). It has also been suggested that certain members of this family may regulate an early step of tetrapyrrole biosynthesis and act as regulators that function based on chlorophyll availability, such that they activate chlorophyll biosynthesis steps when their pigment binding sites are unoccupied (Xu et al., 2002). One of the Hlips is associated with the periphery of photosystem II following exposure to high light intensity, suggesting that it may be required for photosystem II assembly or repair (Promnares et al., 2006; Yao et al., 2007).

In *Synechococcus* OS-A and *Synechococcus* OS-B′ there are five *hli* genes comprising a small gene family (Kilian et al., 2007); *hli4* and *hli5* are tandemly arranged on the genome in both organisms. Such an arrangement is often taken as evidence of a recent duplication event (Koonin, 2005). The high-light acclimated ecotype of marine *Prochlorococcus* MED4 has 22 *hli* gene copies including four tandemly arranged, almost identical ones (*hli6–hli10*); this entire region is also duplicated (*hli16–hli19*) elsewhere in the genome (Bhaya et al., 2002). Transcripts from all of these genes are significantly up-regulated at high light intensity, although other members of the *hli* gene family are constitutively expressed (Steglich et al., 2006). In *Synechococcus* OS-B′ Hlips1–4 are between 50 and 73 amino acids in length, as is the case in most other cyanobacteria (Dolganov et al., 1995; Kilian et al., 2007).

However, Hlip5 is considerably larger (102 amino acids) than the other HLIPs and appears to represent a novel fusion event between another small membrane protein of unknown function (Coh1) at the N-terminus (called TMH1) and the Hlip TMH domain (called TMH2) (Fig. 2.7).

The small Coh1 (cyanobacterial one helix) protein is highly conserved and its gene is found in most other cyanobacterial genomes with the exception of the marine lineage cyanobacteria *Synechococcus* sp. and *Prochlorococcus* sp. It is annotated as "conserved hypothetical". It has been suggested that new multi-domain proteins are formed during the course of evolution by the process of gene duplication followed by gene fusion. Such a fusion may also result in tight co-regulation of these genes and serve to expand the functional role of fusion proteins (Yanai et al., 2000; Bashton and Chothia, 2007). If *hli4* and *hli5* represent a recent duplication event, then the gene fusion between the *coh1* homolog (represented by TMH1 in Hlip5) and a *hli*-like gene (represented by TMH2 in Hlip5) may also have occurred recently. Hlip5, which has two TMH domains, may have increased stability in the membrane and/or additional partner proteins may alter its ability to bind chlorophyll or change its association with complexes of the photosynthetic apparatus (Kuhlbrandt et al., 1994; Funk and Vermaas, 1999; Standfuss et al., 2005; Kilian et al., 2007). Since *Synechococcus* OS-B′ is unable to survive when grown at photon fluence levels of 400 μmol photon $m^{-2}s^{-1}$ (Kilian et al., 2007), the role of Hlips, and in particular the novel two-TMH Hlip5, is of interest. Membrane topology algorithms predict that the loop between TMH1 and TMH2 is more likely to be located in the lumen; thus the conserved aspartate/glutamate residues that are positioned between TMH1 and TMH2 would be present in the lumen and could serve a regulatory function. Molecular analysis of non-photochemical quenching in *A. thaliana* suggests that PsbS senses hyper-excitation of photosynthetic electron transport via the development of a large ΔpH across the thylakoid membranes, which, in turn, leads to protonation of conserved, lumenal glutamate residues in PsbS (Li et al., 2004). PsbS may bind carotenoids that may quench chlorophyll excitation through the formation of a carotenoid radical (Holt et al., 2005; Standfuss et al., 2005).

Hlips may represent one of the single TMH progenitors of the CAB/ELIP/HLIP superfamily (Hoffman et al., 1987; Green and Pichersky, 1994; Kuhlbrandt et al., 1994). An Hlip sequence-like may have fused with another TMH to create a two-TMH structure. A four-helix protein (e.g. PsbS) could represent an event, in which a gene for a two-TMH protein experienced a tandem duplication, followed by fusion of the duplicated genes. The evolution of a three-TMH LHC from a PsbS-type, four-TMH protein would involve loss of the

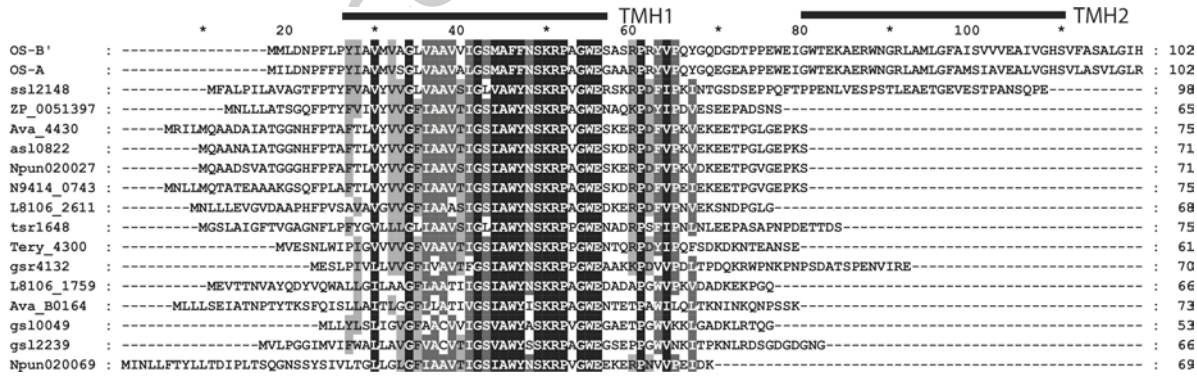

Fig. 2.7. ClustalW alignment of HLIPs from *Synechococcus* OS-B′ and OS-A with Coh1 found in other cyanobacteria. These are *Synechocystis* PCC 6803 (ssl2148), *Crocosphaera watsoni* (ZP_00051397), *Nodularia spumigena* (N9414_0743), *Trichodesmium erythraeum* (Tery_4300), *Nostoc* sp. PCC 7120 (asl0822), *Thermosynechococcus elongatus* (tsr1648), *Nostoc punctiforme* (Npun02002710 and Npun02006958), *Anabaena variabilis* (Ava_4430 and Ava_B0164), and *Lyngbya* sp. PCC8106 (L8106_2611 and L8106_1759) and *Gloeobacter violaceus* PCC7421(gsr4132, gsl0049,and gsl2239). *Black*, *dark gray* and *light gray boxes* indicate completely conserved (100%), highly conserved (.80%) or moderately conserved (.60%) residues, respectively. The positions of TMH1 and TMH2 (of Hlip5) are shown as *black bars* above (Modified from Kilian et al., 2008) [AU4]

last TMH (Green et al., 1991; Green and Pichersky, 1994; Montane and Kloppstech, 2000). The TMH1 and TMH3 of the LHCs are similar to each other and are also similar to TMHs of the distantly related CAB/ELIP/Hlip polypeptides, but the TMH2 of the LHCs is not well conserved (Heddad and Adamska, 2000; Montane and Kloppstech, 2000). The discovery of Hlip5, which represents a novel protein containing an Hlip domain and a domain representing a conserved protein in cyanobacteria, suggests that scrutiny of the extensive genomic databases might reveal evidence of other novel fusion events. The extensive and ongoing rearrangements experienced by the genomes of *Synechococcus* OS-A and OS-B′ (Bhaya et al., 2007) could accelerate the rate of gene fusion events. The large majority of these events would be lethal or lead to a non-competitive phenotype, but if a gene fusion event conferred a selective advantage it may be selected for during evolution.

IV. Metagenomic Analysis

The term metagenomics (also known as "community/environmental genomics") was first used in 1998 (Handelsman et al., 1998), but its impact on research areas such as ecology, evolution and bacterial diversity has really begun to grow exponentially over the last 4 years. Metagenomics can be broadly defined as a field, in which genomic approaches are taken to analyze microbial communities. Because this approach requires neither prior knowledge of the microbial members of the community nor purification of particular species, it has the potential to provide an unbiased view of the functionalities in a community (Streit and Schmitz, 2004; Konstantinidis et al., 2006; Handelsman, 2008; Vieites et al., 2009; Wilmes et al., 2009). Thus, metagenomic analyses are particularly attractive for environments where a large fraction of microbes remain uncultured or are not amenable to standard microbiological analyses (Woyke et al., 2006; Eisen, 2007; Kowalchuk et al., 2007; Raes and Bork, 2008). Sequencing projects have been initiated for a number of different purposes including (i) identification of dominant species (Yooseph et al., 2007), (ii) characterization of the community structure of complex consortia ('phylogenomics') (Strous et al., 2006; Woyke et al., 2006), (iii) identification of novel genes and operons in the population (Beja et al., 2001; Entcheva et al., 2001), and (iv) 'bioprospecting' or searching for desirable gene functions, such as cellulases or chitinases (Lorenz et al., 2002; Riesenfeld et al., 2004; Allen et al., 2007; Gabor et al., 2007; Ferrer et al., 2009). The focus of the first Sargasso Sea metagenome sequence analysis was an attempt to recreate complete genomes of surface-water organisms, notably *Prochlorococcus*. Subsequent metagenomic, genomic and comparative analyses of ocean microbes has led to insights on community genomics in stratified environments in the oceans (Coleman et al., 2006). Metagenome analysis of acid mine-drainage biofilms led to the complete or partial reconstruction of five genomes because of the domination of a small number of relatively homogenous species (*Leptospirillum* and *Ferroplasma*) (Tyson et al., 2004; Tyson and Banfield, 2005). Several groups have focused on the analysis of specific marker genes, such as the *nif* or 16S RNA genes, in metagenomic samples (Johnston et al., 2005; Sogin et al., 2006).

Methods are being developed to use comparative metagenomics and metaproteomic analysis to probe the diversity of bacterial populations and in various environments without relying on the assembly of genomes (Ram et al., 2005; Tringe et al., 2005; DeLong et al., 2006; Ward 2006b; Kowalchuk et al., 2007; Raes et al., 2007; von Mering et al., 2007). Tringe et al. (2005), applied an automated annotation process to establish "environmental gene tags" based on predicted genes on DNA fragments from the metagenome. To derive a measure of the functional profile of different microbial communities, they developed a method to "bin" similar sequences. This analysis allows one to compare the abundance of binned sequences belonging to specific metabolic pathways across communities. This approach rests on the assumption that a few defining habitats may determine genomic profiles in the microbes, since the habitats provide the context for the physiology and the pathways/processes critical for survival of the organism, and place less importance on relationships of specific genes to specific organisms. For instance, photosynthesis genes were found to be highly represented in the Sargasso Sea metagenomic study, but were not nearly as prevalent in other environments.

These results demonstrate that an aggregative approach to understand major functional metabolic pathways may be used without the need to link these functions to particular organisms. Likewise, DeLong et al., (2006) identified protein categories (based on KEGG and COG databases) and subsequently used cluster analysis to identify specific genes that were differentially distributed in the water column.

There are now well over a hundred metagenome projects underway and this number may rise further as sequencing costs fall (Edwards et al., 2006) and high throughput platforms become accessible (Havre et al., 2005; Huson et al., 2007). However, with the exception of studies that target low-complexity environments (Tyson et al., 2004), or where sequence coverage is very high, it is a major technical challenge to assemble complete or even partial microbial genomes from metagenomic data; the depth of sequencing required is usually far greater than most projects allow (Chen and Pachter, 2005; Rusch et al., 2007; Raes and Bork, 2008).

A. Microbial Mat Metagenomics

We have recently obtained an environmental genomic dataset from random shotgun sequencing of total DNA collected from the top green layer of microbial mats at two different temperature sites of Octopus Spring (the spring from which *Syn* OS-A and *Syn* OS-B′ were originally isolated) and from Mushroom Spring, a nearby spring with similar physicochemical characteristics (Fig. 2.8) (Bhaya et al., 2007). Recombinant libraries containing small inserts (2–6 Kbp) or large inserts (10–12 Kbp) from total DNA isolated from the top green layer were created using standard sequencing vectors. Sequencing was carried out on ~200,000 clones and "paired end" sequences were derived; this represented ~200 Mbp of sequence data from the four collection sites. By way of comparison, the read coverage from other metagenome sequencing projects range from 76.2 Mbp from acid-mine biofilms (Tyson et al., 2004) to 1.6 Gbp (surface water, Sargasso Sea) in the first global ocean sampling trip in 2004 (Venter et al., 2004). These datasets are a significant community asset and an exploitable addition to the toolbox available for this well-studied microbial ecosystem. The microbial-mat metagenomic data set is being used to better understand population structure in the mats as well as to gain an understanding of diversity within the *Synechococcus* populations at the different temperatures (Bhaya et al., 2007).

We were able to create large environmental genome scaffolds such that a "virtual" or "composite" genome spanned almost the entire reference genomes of *Synechococcus* OS-A and OS-B′. This suggests that the organisms from which the reference genomes were derived are abundant members of the community (Bhaya et al., 2007). This approach, in which metagenome sequences can be "pasted" back onto an "anchor" or reference genome, is somewhat different from many other metagenomic ongoing projects where no such reference genomes are available. In fact, lack of these reference genomes makes it difficult or impossible to create any large

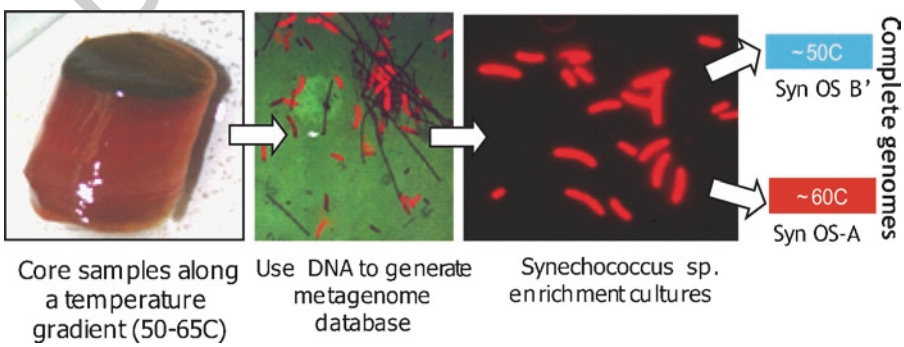

Fig. 2.8. A simplified schematic showing the major steps going from *Panel 1*, initial collection of core mat sample (which can be harvested from different regions), to *Panel 2*, in which total DNA can be extracted from the *top* green layer of the core. This layer which contains a diversity of microbes yields the metagenomic database of sequences. *Panel 3* shows the dilution enrichment culture in which specific *Synechococcus* isolates are identified and the entire genome of specific isolates can be sequenced

contiguous sequences or "contigs" and even more difficult to accurately identify the species of microorganism from which it is derived. The metagenome dataset provided further evidence that there is a rich population of *Synechococcus* spp. in the mats.

Interestingly, a significant fraction (~40%) of the metagenomic sequence reads were quite similar to the *Syn* OS-A and OS-B′ sequences, but could not be confidently assembled onto the anchor genome because they failed to meet the stringent criteria that we set for assembling the sequences. These clones, which we dubbed "illegal clones", might represent regions of recombination, transversion or gene gain/loss ('indels'). As significant gene rearrangements between populations of *Synechococcus* sp. along the temperature gradient (caused by transposon activity, recombination events etc.) have occurred and as there is significant diversity at the gene level, it is clear that this model system can be used for the analysis of genetic variation and of its origins and causes.

B. Functional Diversity in the Metagenome

The variations in nutrient acquisition and utilization capabilities suggested by differences in genome content between the two sequenced *Synechococcus* isolates prompted us to carry out a detailed examination of *Synechococcus*-like sequences in the metagenome dataset for further examples of functionally specialized populations. We investigated a category of "illegal clones", in which one end of the clone had high sequence identity to a specific region on the *Synechococcus* OS-B′ genome, while the paired end sequence did not match any sequences in the *Synechococcus* OS-B′ genome. These clones could represent a *Synechococcus* population closely related to *Synechococcus* OS-B′ that had acquired additional sequences that are absent in the 'anchor' genome. Indeed, we found an example in which a sequenced clone matched the *Synechococcus* OS-B′ genome at both ends but contained an extra 5.5 Kbp region relative to the *Synechococcus* OS-B′ anchor genome (Fig. 2.9). This clone contained seven additional genes, and is flanked by genes that are almost identical to genes in the *Synechococcus* OS-B′ genome (CYB_0562 (99.72% NAID) on the left and CYB_0565 (99.76% NAID) on the right). Of the genes present in the 5.5 Kbp region, two exhibited significant identity to the *feoA* and *feoB* genes of the unicellular thermophilic cyanobacterium, *Thermosynechococcus elongatus.*

The *feoA* and *feoB* genes encode proteins required for ferrous ion transport in several bacteria (Andrews et al., 2003). Iron in the hot-spring environment exists in both ferrous and ferric forms, with the ferrous form predominating under conditions of low oxygen. The *feoA*/*feoB*-like sequences are absent on the *Synechococcus* OS-A and OS-B′ anchor genomes, although they contain several genes required for ferric ion uptake and assimilation. The presence of *feo* genes in an organism that appears to be closely related to *Synechococcus* OS-B′ is interesting because it suggests the presence of functionally specialized populations capable of using Fe^{2+}. Based on Q-PCR results, we hypothesize that the ferrous transport system may accumulate as the mat becomes anoxic during the night, allowing some

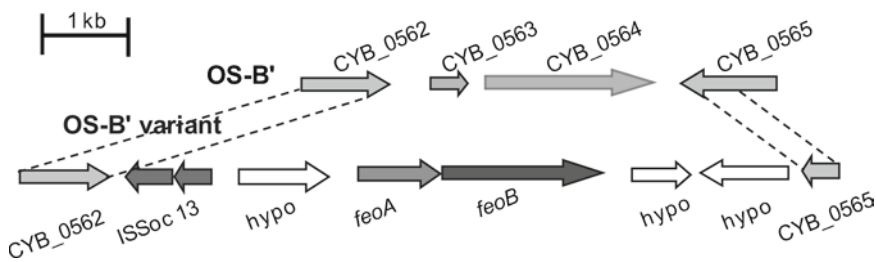

Fig. 2.9. Feo genes containing variant of *Synechococcus* OS-B′. Region containing CYB_0562, CYB_0563, CYB_0564 and CYB_0565 from *Synechococcus* OS-B′. Below is the variant clone containing the *feo* genes plus seven additional CDSs; including OrfB and OrfA of ISSoc13 transposase which appears to be Synechococcus OS-B′-specific, a *feoA* (158 amino acids), *feoB* (763 amino acids). The similarity of this variant to *Synechococcus* OS-B′ beyond this region is not known

Synechococcus OS-B′-like organisms to scavenge the reduced ferrous ions that increase during anoxic mat conditions. This is an example of the advantages of using metagenomic information to identify *Synechococcus* variants in the mat and raises questions relating to functional genomics and technology development in the context of the environment.

C. Lateral Gene Transfer, Transposons and Viruses in the Creation of Variant Populations in Microbial Mats

The dense microbial mats, which harbor a diversity of organisms, are likely to be fertile territory for the study of how genes are transferred between organisms. Although this aspect has not yet been studied in detail among the phototrophs of the mat, it has the potential of revealing the importance of gene transfer processes in communities. Transposons are an important part of the genetic repertoire of the cyanobacteria and in some cases they may be quite active (Bhaya et al., 2007; Zhou et al., 2008). A consequence of this activity would be an accelerated pace of recombination events, which may be responsible to a larger or lesser extent for the large number of "variants" that we see in the populations of cyanobacteria. Some transpositions by IS elements may have occurred very recently (i.e., after/during the process in which the isolates were brought into culture from the original mat environment) and lateral gene transfer may be occurring in the mat cyanobacterial populations (Bhaya et al., unpublished). This suggests that these genomes may be quite fluid and that genetic change in natural populations is an ongoing occurrence. Lateral gene transfer has been extensively documented in numerous bacterial lineages and is considered to play a significant role in genome evolution (Boucher et al., 2003; Gogarten and Townsend, 2005; Lerat et al., 2005). Hot spring mats have been examined for the presence of viruses using both a viral metagenomic approach (Schoenfeld et al., 2008) as well as by other methods (Rice et al., 2001; Ortmann et al., 2006; Snyder et al., 2007) but cyanophages have not yet been carefully studied. A recent study concluded that the newly discovered "viral immunity" systems mediated by CRISPRS are active in the mat *Synechococcus* populations (Heidelberg et al., 2009). This would imply that a better understanding of virology in the context of cyanobacterial populations in the microbial mat is warranted. Marine cyanophages are capable of carrying specific cyanobacterial (or host genes, such as *psbA*) within their genomes and these genes, when functional, may confer some advantage to the viruses. Terrestrial cyanophages may also have similar strategies but this has not yet been shown for cyanobacteria in the hot-spring microbial mats.

V. Future Directions

I will briefly mention three directions that are likely to be directly relevant to many of the issues and concepts discussed here in the future.

A. From a 'Wild' Cyanobacterium to a Model Organism

One of the unique and powerful aspects of our approach was the ability to combine microbial ecology with sophisticated molecular tools to probe function and to correlate it with important and variable environmental parameters. Furthermore, the ability to grow axenic strains of environmentally relevant and dominant thermophilic cyanobacteria adds another important tool to our arsenal. However, one of the crucial requirements for any microorganism to be used as a "model organism" is for it to be effectively genetically manipulated. Some cyanobacteria are naturally transformable and are able to take up foreign DNA and integrate it into the genome by recombination. So it has been relatively easy to develop systems for gene inactivation and the creation of targeted mutants (Koksharova and Wolk, 2002). We have recently demonstrated that we can transform *Synechococcus* OS-B′ and that it appears to be stably transformed (Bhaya et al., unpublished). This opens the door for the next generation of experiments to be performed and to develop a powerful model platform that extends from the *in situ* experiments and environmental measurements in the mat to the testing and refinement of hypotheses under controlled laboratory conditions. This development combined with microarray analysis makes the system much more powerful for examining questions about how environmental parameters affect gene expression.

B. Single Cells and Their Applications

Single-cell analysis is a new technology that has provided new insights into the concept that there is marked phenotypic heterogeneity even in genetically homogenous microbial cultures (Avery, 2006). Single-cell analysis allows one to observe cells as individuals, and to manipulate them in many ways (Breslauer et al., 2006; El-Ali et al., 2006). Microfluidic platforms, which allow for the capture of single cells, can be coupled with numerous downstream analytical manipulations including capillary electrophoresis to examine protein content of a single cell as well as counting individual fluorescent molecules (Huang et al., 2007; Kim et al., 2007). It can also be powerfully combined with new techniques that include whole genome amplification (Lasken, 2007). Particularly in the context of the environment the ability to capture and sequence the entire genome of a single cell opens up untold possibilities (Ottesen et al., 2006; Ishoey et al., 2008; Woyke et al., 2009), since most bacteria cannot be cultivated and the genetic diversity of the microbial world is still to be detailed (Warnecke and Hugenholtz, 2007). This sort of analysis would provide a large "bank" of sequences relevant to a particular environment and therefore would likely improve the ability to build larger scaffolds from metagenomic sequence data.

C. Community Proteomics

Metagenomics and transcriptomics provide a strong basis for assessing the inherent encoded capabilities of a community. However, information about the "business end" of a cell, i.e., the proteins, is perhaps the most technically challenging (Ram et al., 2005). Proteins are made, modified and degraded over the diel cycle or during the life cycle of a cell and new technology, which uses mass spectrometry and 2-D gel electrophoresis, may help identify these proteins and monitor changes and modifications. In the context of the hot spring mats, responses of phototrophic microbes to dynamic environmental conditions, which include temperature, light, lack of oxygen and nutrient levels, would be particularly important to examine. A challenge offered by such an undertaking is to generate concepts and principles from large datasets and to reveal physiological and metabolic changes underlying adaptation to extreme and fluctuating environmental conditions. Modeling and computation analyses will be key components of such an endeavor and are likely to provide essential insights.

Acknowledgments

I would like to acknowledge generous research support from the Frontiers in Integrative Biology Program of the National Science Foundation (Grant EF-0328698), which allowed us to develop various aspects of this multifaceted project. I would particularly like to thank members of the FIBR group (Fred Cohan, Arthur Grossman, John Heidelberg, Michael Kuhl, and Dave Ward) for their individual intellectual contributions to the success of the project. This chapter represents a synthesis of several joint publications, although the emphasis on specific topics is my perspective. Many of the post-docs and students at the Carnegie Institution and elsewhere put in heroic efforts at various stages. Funding and partial support from the Carnegie Institution and from other NSF grants (including Grant 0758374 to Doron Levy, University of Maryland and D.B.) is also gratefully acknowledged. The extended cooperation of Christie Hendrix and members of the Parks Department of Yellowstone National Park is also greatly appreciated.

References

Achtman, M., and Wagner, M. (2008) Microbial diversity and the genetic nature of microbial species. *Nat Rev Micro* **6**: 431–440.

Adams, D.G. (2000) Symbiotic interactions. In *The Ecology of Cyanobacteria*. Whitton, B.A., and Potts, M. (eds). Norwell, MA: Kluwer Academic Publishers, pp. 37–59.

Adams, D.G., and Duggan, P.S. (2008) Cyanobacteria-bryophyte symbioses. *J Exp Bot* **59**: 1047–1058.

Adams, M.M., Gomez-Garcia, M.R., Grossman, A.R., and Bhaya, D. (2008) Phosphorus deprivation responses and phosphonate utilization in a thermophilic Synechococcus sp. from microbial mats. *J Bacteriol* **190**: 8171–8184.

Ahlgren, N.A., and Rocap, G. (2006) Culture Isolation and Culture-Independent Clone Libraries Reveal New Marine Synechococcus Ecotypes with Distinctive Light and N Physiologies. *Appl. Environ. Microbiol.* **72**: 7193–7204.

Allen, R.T., and Day, A.L. (1935) *Hot Springs of the Yellowstone National Park*: Carnegie Institute of Washington.

Allen, E.E., Tyson, G.W., Whitaker, R.J., Detter, J.C., Richardson, P.M., and Banfield, J.F. (2007) Genome dynamics in a natural archaeal population. *PNAS* **104**: 1883–1888.

Allewalt, J.P., Bateson, M.M., Revsbech, N.P., Slack, K., and Ward, D.M. (2006) Effect of temperature and light on growth of and photosynthesis by *Synechococcus* isolates typical of those predominating in the octopus spring microbial mat community of Yellowstone National Park. *Appl Environ Microbiol* **72**: 544–550.

Andrews, S.C., Robinson, A.K., and Rodriguez-Quinones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**: 215–237.

Arakawa, S., Sato, T., Sato, R., Zhang, J., Gamo, T., Tsunogai, U. et al. (2006) Molecular phylogenetic and chemical analyses of the microbial mats in deep-sea cold seep sediments at the northeastern Japan Sea. *Extremophiles* **10**: 311–319.

Archibald, J.M. (2009) The puzzle of plastid evolution. *Curr Biol* **19**: R81-88.

Avery, S.V. (2006) Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Micro* **4**: 577.

Bashton, M., and Chothia, C. (2007) The generation of new protein functions by the combination of domains. *Structure* **15**: 85–99.

Bateson, M.M., and Ward, D.M. (1988) Photoexcretion and consumption of glycolate in a hot spring cyanobacterial mat. *Appl Environ Microbiol* **54**: 1738–1743.

Bebout, B.M., and Garcia-Pichel, F. (1995) UV B-Induced Vertical Migrations of Cyanobacteria in a Microbial Mat. *Appl Environ Microbiol* **61**: 4215–4222.

Bebout, B.M., Fitzpatrick, M.W., and Paerl, H.W. (1993) Identification of the Sources of Energy for Nitrogen Fixation and Physiological Characterization of Nitrogen-Fixing Members of a Marine Microbial Mat Community. *Appl Environ Microbiol* **59**: 1495–1503.

Beja, O., Spudich, E.N., Spudich, J.L., Leclerc, M., and DeLong, E.F. (2001) Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786–789.

Belay, N., Sparling, R., and Daniels, L. (1984) Dinitrogen fixation by a thermophilic methanogenic bacterium. *Nature* **312**: 286–288.

Bhaya, D. (2004) Light matters: phototaxis and signal transduction in unicellular cyanobacteria. *Mol Microbiol* **53**: 745–754.

Bhaya, D., Dufresne, A., Vaulot, D., and Grossman, A. (2002) Analysis of the *hli* gene family in marine and freshwater cyanobacteria. *FEMS Microbiol Lett* **215**: 209–219.

Bhaya, D., Arthur R. Grossman, Anne-Soisig Steunou, Natalia Khuri, Frederick M. Cohan, Natsuko Hamamura et al. (2007) Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME Journal* **1**: 703–713.

Botero, L.M., Brown, K.B., Brumefield, S., Burr, M., Castenholz, R.W., Young, M., and McDermott, T.R. (2004) Thermobaculum terrenum gen. nov., sp. nov.: a non-phototrophic gram-positive thermophile representing an environmental clone group related to the Chloroflexi (green non-sulfur bacteria) and Thermomicrobia. *Arch Microbiol* **181**: 269–277.

Boucher, Y., Douady, C.J., Papke, R.T., Walsh, D.A., Boudreau, M.E., Nesbo, C.L. et al. (2003) Lateral gene transfer and the origins of prokaryotic groups. *Annu Rev Genet* **37**: 283–328.

Breslauer, D.N., Lee, P.J., and Lee, L.P. (2006) Microfluidics-based systems biology. *Mol Biosyst* **2**: 97–112.

Brock, T.D. (1978) *Thermophilic microorganisms and life at high temperatures*. Berlin: Springer Verlag.

Brock, T.D. (1997) The value of basic research: discovery of Thermus aquaticus and other extreme thermophiles. *Genetics* **146**: 1207–1210.

Cardenas, E., and Tiedje, J.M. (2008) New tools for discovering and characterizing microbial diversity. *Curr Opin Biotechnol* **19**: 544–549.

Castenholz, R.W., and Garcia-Pichel, F. (2000) Cyanobacterial responses to UV-radiation. In *The Ecology of Cyanobacteria*. Whitton, B.A. and Potts, M. (eds.). Norwell, MA: Kluwer Academic Publishers, pp. 591–611.

Chen, K., and Pachter, L. (2005) Bioinformatics for whole-genome shotgun sequencing of microbial communities. *PLoS Comput Biol* **1**: 106–112.

Christner, B.C., Mosley-Thompson, E., Thompson, L.G., and Reeve, J.N. (2003) Bacterial recovery from ancient glacial ice. *Environ Microbiol* **5**: 433–436.

Clark, L.L., Ingall, E.D., and Benner, R. (1998) Marine phosphorus is selectively mineralized. *Nature* **393**: 426.

Cohen, Y., Padan, E., and Shilo, M. (1975) Facultative anoxygenic photosynthesis in the cyanobacterium Oscillatoria limnetica. *J Bacteriol* **123**: 855–861.

Coleman, M.L., Sullivan, M.B., Martiny, A.C., Steglich, C., Barry, K., DeLong, E.F., and Chisholm, S.W. (2006) Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* **311**: 1768–1770.

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**: 447–459.

Decker, K.L.M., Potter, C.S., Bebout, B.M., Marais, D.J.D., Carpenter, S., Discipulo, M. et al. (2005) Mathematical simulation of the diel O, S, and C biogeochemistry of a hypersaline microbial mat. *FEMS Microbiol Ecol* **52**: 377–395.

DeLong, E.F., and Karl, D.M. (2005) Genomic perspectives in microbial oceanography. *Nature* **437**: 336–342.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.-U. et al. (2006) Community genomics

among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.

Dolganov, N.A., Bhaya, D., and Grossman, A.R. (1995) Cyanobacterial protein with similarity to the chlorophyll a/b binding proteins of higher plants: evolution and regulation. *Proc Natl Acad Sci U S A* **92**: 636–640.

Doolittle, W.F., and Zhaxybayeva, O. (2009) On the origin of prokaryotic species. *Genome Res* **19**: 744–756.

Dupraz, C., and Visscher, P.T. (2005) Microbial lithification in marine stromatolites and hypersaline mats. *Trends in Microbiology* **13**: 429.

Dyhrman, S.T., Chappell, P.D., Haley, S.T., Moffett, J.W., Orchard, E.D., Waterbury, J.B., and Webb, E.A. (2006) Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* **439**: 68–71.

Edwards, R.A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D.M. et al. (2006) Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* **7**: 57.

Eisen, J.A. (2007) Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. *PLoS Biol* **5**: e82.

El-Ali, J., Sorger, P.K., and Jensen, K.F. (2006) Cells on chips. *Nature* **442**: 403–411.

Entcheva, P., Liebl, W., Johann, A., Hartsch, T., and Streit, W.R. (2001) Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Appl Environ Microbiol* **67**: 89–99.

Ernst, A., Becker, S., Wollenzien, U.I., and Postius, C. (2003) Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. *Microbiology* **149**: 217–228.

Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034–1039.

Ferrer, M., Beloqui, A., Timmis, K.N., and Golyshin, P.N. (2009) Metagenomics for mining new genetic resources of microbial communities. *J Mol Microbiol Biotechnol* **16**: 109–123.

Ferris, M.J., and Ward, D.M. (1997) Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **63**: 1375–1381.

Ferris, M.J., Ruff-Roberts, A.L., Kopczynski, E.D., Bateson, M.M., and Ward, D.M. (1996) Enrichment culture and microscopy conceal diverse thermophilic Synechococcus populations in a single hot spring microbial mat habitat. *Appl Environ Microbiol* **62**: 1045–1050.

Ferris, M.J., Nold, S.C., Santegoeds, C.M., and Ward, D.M. (2001) Examining bacterial population diversity within Octopus Springs microbial mat community. In *Thermophiles*. Reysenbach, A.L., Voytek, M. and Mancinelli R (ed). NY: Kluwer, pp. 51–64.

Ferris, M.J., Kuhl, M., Wieland, A., and Ward, D.M. (2003) Cyanobacterial ecotypes in different optical microenvironments of a 68 degrees C hot spring mat community revealed by 16S-23S rRNA internal transcribed spacer region variation. *Appl Environ Microbiol* **69**: 2893–2898.

Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., and DeLong, E.F. (2008) Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* **105**: 3805–3810.

Funk, C., and Vermaas, W. (1999) A cyanobacterial gene family coding for single-helix proteins resembling part of the light-harvesting proteins from higher plants. *Biochemistry* **38**: 9397–9404.

Gabor, E., Liebeton, K., Niehaus, F., Eck, J., and Lorenz, P. (2007) Updating the metagenomics toolbox. *Biotechnol J* **2**: 201–206.

Gentry, T.J., Wickham, G.S., Schadt, C.W., He, Z., and Zhou, J. (2006) Microarray applications in microbial ecology research. *Microbial Ecol* **52**: 159–175.

Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J. et al. (2005) Opinion: Re-evaluating prokaryotic species. *Nat Rev Microbiol* **3**: 733–739.

Gogarten, J.P., and Townsend, J.P. (2005) Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol* **3**: 679–687.

Green, B.R. (1995) Sequence conservation of light-harvesting and stress response proteins in relation to the molecular structure of LHCII. *Photosynth Res* **44**: 139–148.

Green, B.R., and Pichersky, E. (1994) Hypothesis for the evolution of three helix Chl*a/b* and Chl*a/c* light harvesting antenna proteins from two-helix and four helix ancestors. *Photosynth Res* **39**: 149–162.

Green, B.R., Pichersky, E., and Kloppstech, K. (1991) Chlorophyll *a/b* binding proteins: an extended family. *Trends Biochem Sci* **16**: 181–186.

Guerrero, R., Piqueras, M., and Berlanga, M. (2002) Microbial mats and the search for minimal ecosystems. *Int Microbiol* **5**: 177–188.

Handelsman, J. (2008) Metagenomics is not enough. *DNA Cell Biol* **27**: 219–221.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* **5**: R245-249.

Haverkamp, T., Acinas, S.G., Doeleman, M., Stomp, M., Huisman, J., and Stal, L.J. (2008) Diversity and phylogeny of Baltic Sea picocyanobacteria inferred from their ITS and phycobiliprotein operons. *Environ Microbiol* **10**: 174–188.

Havre, S.L., Webb-Robertson, B.J., Shah, A., Posse, C., Gopalan, B., and Brockman, F.J. (2005) Bioinformatic Insights from metagenomics through visualization. *Proc IEEE Comput Syst Bioinform Conf*: 341–350.

He, Q., Dolganov, N., Bjorkman, O., and Grossman, A.R. (2001) The high light-inducible polypeptides in

Synechocystis PCC6803. Expression and function in high light. *J Biol Chem* **276**: 306–314.

Heddad, M., and Adamska, I. (2000) Light stress-regulated two-helix proteins in Arabidopsis thaliana related to the chlorophyll *a*/*b*-binding gene family. *Proc Natl Acad Sci U S A* **97**: 3741–3746.

Heidelberg, J.F., Nelson, W.C., Schoenfeld, T., and Bhaya, D. (2009) Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS ONE* **4**: e4169.

Herter, S., Farfsing, J., Gad'On, N., Rieder, C., Eisenreich, W., Bacher, A., and Fuchs, G. (2001) Autotrophic CO_2 fixation by *Chloroflexus aurantiacus*: study of glyoxylate formation and assimilation via the 3-hydroxypropionate cycle. *J Bacteriol* **183**: 4305–4316.

Hoehler, T.M., Bebout, B.M., and Des Marais, D.J. (2001) The role of microbial mats in the production of reduced gases on the early Earth. *Nature* **412**: 324.

Hoffman, N.E., Pichersky, E., Malik, V.S., Castresano, C., Ko, K., Darr, S.C., and Cashmore, A.R. (1987) A cDNA encoding a photosystem I protein with homology to photosystem II chlorophyll a/b-binding proteins. *Proc Natl Acad Sci USA* **84**: 8844–8848.

Holo, H. (1989) *Chloroflexus aurantiacus* secretes 3-hydroxypropionate, a possible intermediate in the assimiliation of CO_2 and acetate. *Arch Microbiol* **151**: 252–256.

Holt, N.E., Zigmantas, D., Valkunas, L., Li, X.P., Niyogi, K.K., and Fleming, G.R. (2005) Carotenoid cation formation and the regulation of photosynthetic light harvesting. *Science* **307**: 433–436.

Howe, C.J., Barbrook, A.C., Nisbet, R.E., Lockhart, P.J., and Larkum, A.W. (2008) The origin of plastids. *Philos Trans R Soc Lond B Biol Sci* **363**: 2675–2685.

Huang, J., Su, Z., and Xu, Y. (2005) The evolution of microbial phosphonate degradative pathways. *J Mol Evol* **61**: 682–690.

Huang, B., Wu, H., Bhaya, D., Grossman, A., Granier, S., Kobilka, B.K., and Zare, R.N. (2007) Counting low-copy number proteins in a single cell. *Science* **315**: 81–84.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Res* **17**: 377–386.

Inskeep, W.P., and McDermott, T.R. (2005) *Geothermal Biology and Geochemistry in Yellowstone National Park*: Thermal Biology Institute, Montana State University Publications, Bozeman MT.

Ishii, M., Chuakrut, S., Arai, H., and Igarishi, Y. (2004) Occurence, biochemistry and possible biotechnological application of the 3-hydroxypropionate cycle *Appl Microbiol Biotechnol* **64**: 605–610.

Ishoey, T., Woyke, T., Stepanauskas, R., Novotny, M., and Lasken, R.S. (2008) Genomic sequencing of single microbial cells from environmental samples. *Curr Opin Microbiol* **11**: 198–204.

Jansson, S. (1999) A guide to the Lhc genes and their relatives in Arabidopsis. *Trends Plant Sci* **4**: 236–240.

Johnston, A.W., Li, Y., and Ogilvie, L. (2005) Metagenomic marine nitrogen fixation-feast or famine? *Trends Microbiol* **13**: 416–420.

Jungblut, A.-D., Hoeger, S.J., Mountfort, D., Hitzfeld, B.C., Dietrich, D.R., and Neilan, B.A. (2006) Characterization of microcystin production in an Antarctic cyanobacterial mat community. *Toxicon* **47**: 271.

Keiter, R.B., and Boyce, M.S. (1991) *The Greater Yellowstone Ecosystem*: Yale University Press.

Kettler, G.C., Martiny, A.C., Huang, K., Zucker, J., Coleman, M.L., Rodrigue, S. et al. (2007) Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet* **3**: e231.

Kilian, O., Steunou, A.S., Fazeli, F., Bailey, S., Bhaya, D., and Grossman, A.R. (2007) Responses of a thermophilic *Synechococcus* isolate from the microbial mat of Octopus Spring to light. *Appl Environ Microbiol* **73**: 4268–4278.

Kim, S., Huang, B., and Zare, R.N. (2007) Microfluidic separation and capture of analytes for single-molecule spectroscopy. *Lab Chip* **7**: 1663–1665.

Koksharova, O.A., and Wolk, C.P. (2002) Genetic tools for cyanobacteria. *Appl Microbiol Biotechnol* **58**: 123–137.

Konopka, A. (1992) Accumulation and utilization of polysaccharide by hot-spring phototrophs during a light-dark transition. *FEMS Microbiol Lett* **102**: 27–32.

Konstantinidis, K.T., and Tiedje, J.M. (2004) Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc Natl Acad Sci U S A* **101**: 3160–3165.

Konstantinidis, K.T., Ramette, A., and Tiedje, J.M. (2006) The bacterial species definition in the genomic era. *Philos Trans R Soc Lond B Biol Sci* **361**: 1929–1940.

Koonin, E.V. (2005) Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet* **39**: 309–338.

Kowalchuk, G.A., Speksnijder, A.G.C.L., Zhang, K., Goodman, R.M., and van Veen, J.A. (2007) Finding the needles in the metagenome haystack. *Microb Ecol* **53**: 475–485**.**

Krumbein, W.E., Paterson, D.M., and Zavarin, G.A. (2003) *Fossil and Recent Biofilms: A Natural History of Life on Earth*: Kluwer Academic Publishers.

Kuhl, M. (2005) Optical microsensors for analysis of microbial communities. *Methods Enzymol* **397**: 166–199.

Kuhl, M., Lassen, C., and Revsbech, N.P. (1997) A simple light meter for measurements of PAR (400–700 nm) with fiber-optic microprobes. *Aquatic Microbial Ecol.* **13**: 197–207.

Kuhl, M., Steukart, C., Eikert, G., and Jeroschewski, P. (1998) A H_2S sensor for profiling biofilms and sediments: application in an acidic lake sediment. *Aquatic Microbial Ecol* **15**: 201–209.

Kuhlbrandt, W., Wang, D.N., and Fujiyoshi, Y. (1994) Atomic model of plant light harvesting complex by electron crystallography. *Nature* **367**: 614–621.

Lasken, R.S. (2007) Single-cell genomic sequencing using Multiple Displacement Amplification. *Current Opin Microbiol* **10**: 510–516.

Lau, E., Nash, C.Z., Vogler, D.R., and Cullings, K.W. (2005) Molecular diversity of cyanobacteria inhabiting coniform structures and surrounding mat in a Yellowstone hot spring. *Astrobiol* **5**: 83–92.

Lerat, E., Daubin, V., Ochman, H., and Moran, N.A. (2005) Evolutionary origins of genomic repertoires in bacteria. *PLoS Biol* **3**: e130.

Ley, R.E., Harris, J.K., Wilcox, J., Spear, J.R., Miller, S.R., Bebout, B.M. et al. (2006) Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl Environ Microbiol* **72**: 3685–3695.

Li, X.P., Gilmore, A.M., Caffarri, S., Bassi, R., Golan, T., Kramer, D., and Niyogi, K.K. (2004) Regulation of photosynthetic light harvesting involves intrathylakoid lumen pH sensing by the PsbS protein. *J Biol Chem* **279**: 22866–22874.

Lorenz, P., Liebeton, K., Niehaus, F., and Eck, J. (2002) Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Curr Opin Biotechnol* **13**: 572–577.

Ludwig, R., Pringault, O., de Wit, R., de Beer, D., and Jonkers, H.M. (2006) Limitation of oxygenic photosynthesis and oxygen consumption by phosphate and organic nitrogen in a hypersaline microbial mat: a microsensor study. *FEMS Microbiol Ecol* **57**: 9–17.

Luque, I., Flores, E., and Herrero, A. (1994) Nitrate and nitrite transport in the cyanobacterium *Synechococcus* sp. PCC 7942 are mediated by the same permease. *Biochim Biophys Acta* **1184**: 296–298.

Mackerras, A.H., Youens, B.N., Wier, R.C., and Smith, G.D. (1990) Is cyanophycin involved in the integration of nitrogen and carbon metabolism in the cyanobacteria *Anabaena cylindrica* and *Gleothece* grown on light/dark cycles? *J Gen Microbiol* **136**: 2049–2056.

Montane, M.-H., and Kloppstech, K. (2000) The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? *Gene* **258**: 1–8.

Nagy, M.L., Perez, A., and Garcia-Pichel, F. (2005) The prokaryotic diversity of biological soil crusts in the Sonoran Desert (Organ Pipe Cactus National Monument, AZ). *FEMS Microbiol Ecol* **54**: 233–245.

Namsaraev, Z.B., Gorlenko, V.M., Namsaraev, B.B., Buriukhaev, S.P., and Iurkov, V.V. (2003) [The structure and biogeochemical activity of the phototrophic communities from the Bol'sherechenskii alkaline hot spring]. *Mikrobiologiia* **72**: 228–238.

Nold, S.C., and Ward, D.M. (1996) Photosynthate partitioning and fermentation in hot spring microbial mat communities. *Appl. Environ. Microbiol.* **62**: 4598–4607.

Nubel, U., Garcia-Pichel, F., Kuhl, M., and Muyzer, G. (1999) Quantifying microbial diversity: morphotypes, 16S rRNA genes, and carotenoids of oxygenic phototrophs in microbial mats. *Appl Environ Microbiol* **65**: 422–430.

Nubel, U., Bateson, M.M., Vandieken, V., Wieland, A., Kuhl, M., and Ward, D.M. (2002) Microscopic examination of distribution and phenotypic properties of phylogenetically diverse Chloroflexaceae-related bacteria in hot spring microbial mats. *Appl Environ Microbiol* **68**: 4593–4603.

Ortmann, A.C., Wiedenheft, B., Douglas, T., and Young, M. (2006) Hot crenarchaeal viruses reveal deep evolutionary connections. *Nat Rev Microbiol* **4**: 520–528.

Ottesen, E.A., Hong, J.W., Quake, S.R., and Leadbetter, J.R. (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* **314**: 1464–1467.

Paerl, H.W. (2000) Marine plankton. In *The Ecology of Cyanobacteria*. Whitton, B.A. and Potts, M. (eds). Norwell, MA: Kluwer Academic Publishers, pp. 121–148.

Paerl, H.W., Pinckney, J.L., and Steppe, T.F. (2000) Cyanobacterial-bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. *Environ Microbiol* **2**: 11–26.

Palenik, B., Brahamsha, B., Larimer, F.W., Land, M., Hauser, L., Chain, P. et al. (2003) The genome of a motile marine *Synechococcus*. *Nature* **424**: 1037–1042.

Papke, R.T., Ramsing, N.B., Bateson, M.M., and Ward, D.M. (2003) Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* **5**: 650–659.

Parkhill, J., Sebaihia, M., Preston, A., Murphy, L.D., Thomson, N., Harris, D.E. et al. (2003) Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. *Nat Genet* **35**: 32–40.

Partensky, F., Hess, W.R., and Vaulot, D. (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Reviews* **63**: 106–127.

Pierson, B.K., and Castenholz, R.W. (1974) A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Arch Microbiol* **100**: 5–24.

Podar, M., and Reysenbach, A.L. (2006) New opportunities revealed by biotechnological explorations of extremophiles. *Curr Opin Biotechnol* **17**: 250–255.

Promnares, K., Komenda, J., Bumba, L., Nebesarova, J., Vacha, F., and Tichy, M. (2006) Cyanobacterial small chlorophyll-binding protein ScpD (HliB) is located on the periphery of photosystem II in the vicinity of PsbH and CP47 subunits. *J. Biol. Chem.* **281**: 32705–32713.

Quinn, J.P., Kulakova, A.N., Cooley, N.A., and McGrath, J.W. (2007) New ways to break an old bond: the bacterial carbon-phosphorus hydrolases and their role in biogeochemical phosphorus cycling. *Environ Microbiol* **9**: 2392–2400.

Raes, J., and Bork, P. (2008) Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol* **6**: 693–699.

Raes, J., Foerstner, K.U., and Bork, P. (2007) Get the most out of your metagenome: computational analysis of

environmental sequence data. *Curr Opin Microbiol* **10**: 490–498.

Ram, R.J., Verberkmoes, N.C., Thelen, M.P., Tyson, G.W., Baker, B.J., Blake, R.C., 2nd et al. (2005) Community proteomics of a natural microbial biofilm. *Science* **308**: 1915–1920.

Ramsing, N.B., Ferris, M.J., and Ward, D.M. (1997) Light-induced motility of thermophilic *Synechococcus* isolates from Octopus Spring, Yellowstone National Park. *Applied & Environmental Microbiology* **63**: 2347–2354.

Ramsing, N.B., Ferris, M.J., and Ward, D.M. (2000) Highly ordered vertical structure of *Synechococcus* populations within the one-millimeter-thick photic zone of a hot spring cyanobacterial mat. *Appl Environ Microbiol* **66**: 1038–1049.

Reid, R.P., Visscher, P.T., Decho, A.W., Stolz, J.F., Bebout, B.M., Dupraz, C. et al. (2000) The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. *Nature* **406**: 989–992.

Revsbech, N.P., and Ward, D.M. (1984) Microelectrode studies of interstitial water chemistry and photosynthetic activity in a hot spring microbial mat. *Appl Environ Microbiol* **48**: 270–275.

Reysenbach, A.L., and Cady, S.L. (2001) Microbiology of ancient and modern hydrothermal systems. *Trends Microbiol* **9**: 79–86.

Reysenbach, A.L., and Shock, E. (2002) Merging genomes with geochemistry in hydrothermal ecosystems. *Science* **296**: 1077–1082.

Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S. et al. (2001) Viruses from extreme thermal environments. *Proc Natl Acad Sci U S A* **98**: 13341–13345.

Richardson, L.L., and Castenholz, R.W. (1987) Diel vertical movements of the cyanobacterium *Oscillatoria terebriformis* in a sulfide-rich hot spring microbial mat. *Appl Environ Microbiol* **53**: 2142–2150.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004) Metagenomics: genomic analysis of microbial communities. *Annu Rev Genet* **38**: 525–552.

Rocap, G., Distel, D.L., Waterbury, J.B., and Chisholm, S.W. (2002) Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. *Appl Environ Microbiol* **68**: 1180–1191.

Rocha, E.P. (2004) Order and disorder in bacterial genomes. *Curr Opin Microbiol* **7**: 519–527.

Rocha, E.P. (2008) The organization of the bacterial genome. *Annu Rev Genet* **42**: 211–233.

Rubio, L.M., and Ludden, P.W. (2005) Maturation of nitrogenase: a biochemical puzzle. *J Bacteriol* **187**: 405–414.

Ruff-Roberts, A.L., Kuenen, J.G., and Ward, D.M. (1994) Distribution of cultivated and uncultivated cyanobacteria and *Chloroflexus*-like bacteria in hot spring microbial mats. *Appl Environ Microbiol* **60**: 697–704.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S. et al. (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: e77.

Sato, M., Masuda, Y., Kirimura, K., and Kino, K. (2007) Thermostable ATP regeneration system using polyphosphate kinase from *Thermosynechococcus elongatus* BP-1 for D-amino acid dipeptide synthesis. *J Biosci Bioeng* **103**: 179–184.

Schloss, P.D., and Handelsman, J. (2007) The last word: books as a statistical metaphor for microbial communities. *Annu Rev Microbiol* **61**: 23–34.

Schoenfeld, T., Patterson, M., Richardson, P.M., Wommack, K.E., Young, M., and Mead, D. (2008) Assembly of viral metagenomes from Yellowstone hot springs. *Appl Environ Microbiol* **74**: 4164–4174.

Schopf, J.W. (2000) The fossil record: tracing the roots of the cyanobacterial lineage. In *The Ecology of Cyanobacteria*. Whitton, B.A., and Potts, M. (eds): Kluwer, pp. 13–35.

Sheehan, K.B., Patterson, D.J., Dicks, B.L., and Henson, J.M. (2005) *Seen and Unseen, Discovering the Microbes of Yellowstone*: Globe Press and Monatana State University.

Simon, R.D. (1987) Inclusion bodies in the cyanobacteria: cyanophycin, polyphosphate, polyhedral bodies. In *The Cyanobacteria*. Fay, P., and Van Baalen, C. (eds.). Amsterdam: Elsevier Science Publishers B. V., pp. 199–225.

Snyder, J.C., Wiedenheft, B., Lavin, M., Roberto, F.F., Spuhler, J., Ortmann, A.C. et al. (2007) Virus movement maintains local virus population diversity. *Proc Natl Acad Sci U S A* **104**: 19102–19107.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R. et al. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci U S A* **103**: 12115–12120.

Soule, T., Stout, V., Swingley, W.D., Meeks, J.C., and Garcia-Pichel, F. (2007) Molecular genetics and genomic analysis of scytonemin biosynthesis in *Nostoc punctiforme* ATCC 29133. *J Bacteriol* **189**: 4465–4472.

Stahl, D.A., Lane, D.J., Olsen, G.J., and Pace, N.R. (1985) Characterization of a Yellowstone hot spring microbial community by 5 S rRNA sequences. *Appl Environ Microbiol* **49**: 1379–1384.

Stal, L.J. (2000) Cyanobacterial mats and stromatolites. In *The Ecology of Cyanobacteria*. Potts, M., and B.A.Whitton (eds). Norwell, MA: Kluwer Academic Publishers, pp. 61–120.

Stal, L.J., and Caumette, P. (1994) *Microbial Mats: Structure, Development and Environmental Significance*: Springer Verlag.

Staley, J.T. (1997) Biodiversity: are microbial species threatened? *Curr Opin Biotechnol* **8**: 340–345.

Standfuss, J., Terwisscha van Scheltinga, A.C., Lamborghini, M., and Kuhlbrandt, W. (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 A resolution. *Embo J* **24**: 919–928.

Steglich, C., Futschik, M., Rector, T., Steen, R., and Chisholm, S.W. (2006) Genome-wide analysis of light sensing in *Prochlorococcus*. *J Bacteriol* **188**: 7796–7806.

Steunou, A.S., Bhaya, D., Bateson, M.M., Melendrez, M.C., Ward, D.M., Brecht, E. et al. (2006) In situ analysis of nitrogen fixation and metabolic switching in unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats. *Proc Natl Acad Sci U S A* **103**: 2398–2403.

Steunou, A.S., Jensen, S.I., Brecht, E., Becraft, E.D., Bateson, M.M., Kilian, O. et al. (2008) Regulation of *nif* gene expression and the energetics of N_2 fixation over the diel cycle in a hot spring microbial mat. *ISME J* **2**: 364–378.

Stewart, W.D.P. (1970) Nitrogen fixation by blue-green algae in Yellowstone thermal areas. *Phycologia* **9**: 261–268.

Strauss, G., and Fuchs, G. (1993) Enzymes of a novel autotrophic CO_2 fixation pathway in the phototrophic bacterium *Chloroflexus aurantiacus*, the 3-hydroxypropionate cycle. *Eur J Biochem* **215**: 633–643.

Streit, W.R., and Schmitz, R.A. (2004) Metagenomics - the key to the uncultured microbes. *Curr Opin Microbiol* **7**: 492–498.

Strous, M., Pelletier, E., Mangenot, S., Rattei, T., Lehner, A., Taylor, M.W. et al. (2006) Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**: 790–794.

Suzuki, M.T., and DeLong, E.F. (2002) Marine prokaryotic diversity. In *Biodiversity of Microbial Life*. Staley, J.T., and Reysenbach, A.L. (eds). New York: Wiley Liss, pp. 209–234.

Teske, A., and Stahl, D. (2002) Microbial mats and biofilms:evolution, structure, and function of fixed microbial communities. In *Biodiversity of Microbial Life*. Staley, J.T., and Reysenbach, A.L. (eds). New York: Wiley-Liss, pp. 49–100.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W. et al. (2005) Comparative metagenomics of microbial communities. *Science* **308**: 554–557.

Tyson, G.W., and Banfield, J.F. (2005) Cultivating the uncultivated: a community genomics perspective. *Trends Microbiol* **13**: 411–415.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M. et al. (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**: 37–43.

van der Meer, M.T., Schouten, S., Bateson, M.M., Nubel, U., Wieland, A., Kuhl, M. et al. (2005) Diel variations in carbon metabolism by green nonsulfur-like bacteria in alkaline siliceous hot spring microbial mats from Yellowstone National Park. *Appl Environ Microbiol* **71**: 3978–3986.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A. et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Vieites, J.M., Guazzaroni, M.E., Beloqui, A., Golyshin, P.N., and Ferrer, M. (2009) Metagenomics approaches in systems microbiology. *FEMS Microbiol Rev* **33**: 236–255.

Vincent, W.F. (2000) Cyanobacterial dominance in the polar regions. In *The Ecology of Cyanobacteria*. Whitton, B.A., and Potts, M. (eds.). Norwell, MA: Kluwer Academic Publishers, pp. 37–59.

Vincent, W.F., Mueller, D.R., and Bonilla, S. (2004) Ecosystems on ice: the microbial ecology of Markham Ice Shelf in the high Arctic. *Cryobiology* **48**: 103–112.

von Mering, C., Hugenholtz, P., Raes, J., Tringe, S.G., Doerks, T., Jensen, L.J. et al. (2007) Quantitative phylogenetic assessment of microbial communities in diverse environments. *Science* **315**: 1126–1130.

Wanner, B.L. (1994) Phosphate-regulated genes for the utilization of phosphonates in members of the family *Enterobacteriaceae*. In *Phosphate in Microorganisms. Cellular and Molecular Biology*. Torriani-Gorini, A., Yagil, E., and Silver, S. (eds.). Washington, DC: ASM Press, pp. 215–222.

Ward, D.M. (2006a) Microbial diversity in natural environments: focusing on fundamental questions. *Antonie Van Leeuwenhoek* **90**: 309–324.

Ward, N. (2006b) New directions and interactions in metagenomics research. *FEMS Microbiol Ecol* **55**: 331–338.

Ward, D.M., and Castenholz, R.W. (2000) Cyanobacteria in geothermal habitats. In *The Ecology of Cyanobacteria*. Whitton, B.A., and Potts, M. (eds). Norwell, MA: Kluwer Academic Publishers, pp. 37–59.

Ward, D.M., and Cohan, F.M. (2005) Microbial diversity in hot spring cyanobacterial mats: pattern and prediction. In *Geothermal Biology and Geochemistry in Yellowstone National Park*, Inskeep, W.P., and McDermott, T. (eds.): Thermal Biology Institute (Montana State University), pp. 185–201.

Ward, D.M., Weller, R., and Bateson, M.M. (1990) 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**: 63–65.

Ward, D.M., Panke, S., Kloppel, K.D., Christ, R., and Fredrickson, H. (1994) Complex polar lipids of a hot spring cyanobacterial mat and its cultivated inhabitants. *Appl Environ Microbiol* **60**: 3358–3367.

Ward, D.M., Ferris, M.J., Nold, S.C., and Bateson, M.M. (1998) A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol Mol Biol Rev* **62**: 1353–1370.

Ward, D.W., Papke, R.T., Nubel, U., and McKitrick M.C. (2002) Natural History of Microorganisms Inhabiting Hot spring Microbial Mat Communities: Clues to the Origin of Microbial Diversity and Implications for Microbiology and Macrobiology. In *Biodiversity of Microbial Life*. Staley, J.T., and Reysenbach, A.L. (eds), pp. 25–48.

Warnecke, F., and Hugenholtz, P. (2007) Building on basic metagenomics with complementary technologies. *Genome Biol* **8**: 231.

West, N.J., Schonhuber, W.A., Fuller, N.J., Amann, R.I., Rippka, R., Post, A.F., and Scanlan, D.J. (2001) Closely related *Prochlorococcus* genotypes show remarkably different depth distributions in two oceanic regions as revealed by in situ hybridization using 16S rRNA-targeted oligonucleotides. *Microbiology* **147**: 1731–1744.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V. et al. (2008) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **36**: D13-21.

Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* **95**: 6578–6583.

Whitton, B.A., and Potts, M. (2000) Introduction to the cyanobacteria. In *The Ecology of Cyanobacteria*. Whitton, B.A., and Potts, M. (eds). Norwell, MA: Kluwer Academic Publishers, pp. 1–11.

Wilmes, P., Simmons, S.L., Denef, V.J., and Banfield, J.F. (2009) The dynamic genetic repertoire of microbial communities. *FEMS Microbiol Rev* **33**: 109–132.

Woyke, T., Teeling, H., Ivanova, N.N., Huntemann, M., Richter, M., Gloeckner, F.O. et al. (2006) Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* **443**: 950–955.

Woyke, T., Xie, G., Copeland, A., Gonzalez, J.M., Han, C., Kiss, H. et al. (2009) Assembling the marine metagenome, one cell at a time. *PLoS ONE* **4**: e5299.

Wynn-Williams, D.D. (2000) Cyanobacteria in deserts- Life at the limit? In *The Ecology of Cyanobacteria*. Whitton, B.A., and Potts, M. (eds). Norwell, MA: Kluwer Academic Publishers, pp. 37–59.

Xu, H., Vavilin, D., Funk, C., and Vermaas, W. (2002) Small Cab-like proteins regulating tetrapyrrole biosynthesis in the cyanobacterium *Synechocystis* sp. PCC6803. *Plant Molecular Biology.* **49**: 149–160.

Yanai, I., Camacho, C.J., and DeLisi, C. (2000) Predictions of gene family distributions in microbial genomes: evolution by gene duplication and modification. *Phys Rev Lett* **85**: 2641–2644.

Yao, D., Kieselbach, T., Komenda, J., Promnares, K., Prieto, M.A., Tichy, M. et al. (2007) Localization of the small CAB-like proteins in photosystem II. *J Biol Chem* **282**: 267–276.

Yooseph, S., Sutton, G., Rusch, D.B., Halpern, A.L., Williamson, S.J., Remington, K. et al. (2007) The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biol* **5**: e16.

Yoshihara, S., and Ikeuchi, M. (2004) Phototactic motility in the unicellular cyanobacterium Synechocystis sp. PCC 6803. *Photochem Photobiol Sci* **3**: 512–518.

Zhou, F., Olman, V., and Xu, Y. (2008) Insertion sequences show diverse recent activities in Cyanobacteria and Archaea. *BMC Genomics* **9**: 36.

Author Queries

Chapter No.: 2 0001279820

Queries	Details Required	Author's Response
AU1	Please provide the opening parenthesis in the sentence "difference at the 16S rRNA sequence level."	
AU2	Please provide the closing parenthesis in the sentence "difference at the 16S rRNA sequence level…"	
AU3	Please confirm the unit symbols (negative sign) are mismatch with figure and figure caption.	
AU4	Please provide details for reference Kilian et al. (2008) in the list.	
AU5	Please check the inserted closing quote is appropriate in sentence "The term metagenomics…"	