Microbiological Targets for Ocean Observing Laboratories (Micro-TOOLs) 2010
Workshop I

Gordon and Betty Moore Foundation
January 13-15, 2010
Palo Alto, California

Workshop I Final Report

I. Introduction

The oceans play critical roles in maintaining the habitability of the planet. There are many ecosystem services provided by the oceans, including being a major source of food [1]. Although virtually invisible to the eye, ocean food webs depend on the microalgae and bacteria in addition to the macroscopic invertebrates, fish and mammals most commonly associated with life in the sea [2]. Marine microorganisms are abundant and diverse but not well-characterized, because they are small, and look similar even under the most powerful microscopes. Bacteria are present at concentrations of nearly a million cells per drop of seawater, but most of these cells have never been brought into stable culture in the laboratory. These millions and billions of cells range in size, shape, and more importantly metabolic diversity, and together these cells create a complex biological milieu called "microbial communities". Microscopic microbial communities are composed of microbial (Bacteria, Archaea and Eukarya) food chains and chemical interactions that underlie biogeochemical cycles, including those necessary for maintaining habitability of our planet and that fuel the food chains of larger organisms used for human consumption [3-4].

Biogeochemical cycles are the chemical transformations of elements through biological, chemical and physical processes. The transformation of elements is significant since it affects their biological availabilities, solubilities, and solid-liquid-chemical states. Key biogeochemical cycles include the cycling of carbon through photosynthetic microorganisms and macroorganisms, animals and the atmosphere, the cycling of nutrients including nitrogen and phosphorus, and the cycling of trace elements and gases. Microorganisms control the nutritional status of ocean waters, and the fluxes of greenhouse gases between the atmosphere and oceans. Marine microorganisms are diverse metabolically and taxonomically. Even within taxa, there appears to be a high degree of genetic and genomic diversity, with closely related strains varying in gene content, genome size, composition and gene organization. It is believed that a large fraction of marine microbial diversity still eludes cultivation and are known only from nucleic

acid sequences [5-6]. However, major advances have been made in cultivating the most abundant open ocean marine microorganisms that are useful model systems [7].

The biological complexity of microbial assemblages at microscopic scales is paralleled by the complexity of the ocean habitat ranging from microscopic to global scales [2]. The world's oceans are themselves physically and chemically diverse, varying in circulation patterns, salinity, chemical composition (including nutrient concentrations), and temperature. The marine microbial habitats not only range from boiling to below freezing temperature extremes, but also differ widely in hydrostatic pressure and salinity. The rotation of the Earth, winds, and physical-chemical properties result in global scale circulation patterns (the thermohaline circulation) and basin-scale and regional currents. The microbial habitat itself is dynamic and evolving over both spatial and temporal scales, which adds to the difficulty of determining the function of diverse microorganisms in the vast ocean basins.

The climate of Earth is in a state of change [8], dramatically affected by anthropogenic activities, including but not limited to carbon dioxide emissions [9]. Future climate scenarios are very uncertain, but the oceans, and life within them will be affected by [10-11] and play a major role in dampening or amplifying environmental change. Although we have recently learned much about the diversity, genomics and physiology of microbial community components [7], we know little about the relationship of microbial community structure to the physical and chemical dynamics of the oceans, including long-term environmental change. Time-series of key biogeochemical parameters show that the ocean ecosystems are in a state of change [9,12-15], but currently we only monitor very basic aspects of microbial community structure-for example, total microbial genetic diversity, pigment concentrations, or total cell numbers. These measurements provide little information on the key components of microbial communities that control the ocean habitat.

The key chemical reactions of biogeochemical cycles [16] are associated with specific microorganisms, enzymes and genes. It is possible to measure, monitor and study the key biogeochemical processes that control climate-relevant features and ecosystem function, by assaying microorganisms, genes, gene expression or enzyme activity. The rapid development of technologies for assaying genes and gene expression has paralleled development of ocean laboratories, from moored submersed instrumentation to autonomous underwater vehicles that can traverse great distances and depths [17-18]. By coupling these technological advances, a revolution in approach to studying the marine microbiota is at hand. Acquisition of physical samples from the ocean's interior needs not rely exclusively on manned ship-based expeditions and limited mooring installations. New technologies for observing the ocean autonomously will provide fundamentally new information at spatial and temporal scales that have not been possible previously. This approach will provide not only an understanding of the function of microorganisms in the environment, but will also provide data on the factors that control microbial distributions and activities. In turn, predictive power for understanding the fate of ocean ecosystems during environmental change will improve.

One key challenge in approaching this new observing framework is enumerating specific microorganisms and linking them to key biogeochemical processes. Techniques that reveal cellular markers, such as RNA or protein, have the potential to provide taxonomic identification as well as information on function [19-20]. Existing techniques can now be implemented in relatively high throughput assays such as microarrays or "chips" as well as a number of other approaches such as high throughput quantitative PCR and RT-PCR. Gene targets, chosen carefully, have the potential to provide much more informative data on the physiological status, nutrient limitation, biogeochemical processes and stress in microbial communities. Furthermore, progress in biological ocean instrumentation now makes it feasible to begin to deploy the most highly prioritized probes now, and begin to develop a time-series of biological function in addition to diversity information. Both approaches will help to gain microbial biogeochemical information with higher temporal and spatial resolution, and begin to close the gap between physical and chemical data time scales and those of biogeochemistry. In that light, the goal of this workshop was to identify ecologically-important characteristics of community metabolism that reflect microbial ecosystem properties important for understanding change in ecosystem function or health. Furthermore, the workshop focused on how to develop and implement strategies for studying microbiological processes at time and space scales relevant for understanding the links between ocean physics and microbial processes, and to provide the foundation for examining the change in roles of microbially-driven processes through time. Although this field is evolving rapidly, workshop participants were challenged to ground their recommendations in what is possible in the near term so as to provide a clear set of priorities for guiding research and development of opportunities that exist now.

II. Goals for implementing biogeochemical probe strategies

Ultimately, remote instrumentation deployed in strategically placed arrays will provide high resolution information on multiple components of biogeochemical cycles in concert with chemical and physical data. Sensors and instrumentation for chemical and physical data are well-developed. Sensors for microbes require substantial development, partially because the microbes of the oceans are not well-characterized, and partially because the targets for microbial or biogeochemical properties need to be chosen, prioritized and evaluated [21]. The focus of this workshop was on identifying important microbial targets that would provide critical information for understanding the biogeochemical function of ocean ecosystems. The following specific goals reflect the needs for developing a comprehensive suite of microbial probes for studying and monitoring the ocean.

1. Identification of key characteristics of microbial communities for monitoring and comparing microbial community function.

Microbial communities are diverse and are responsible for a wide variety of biogeochemical transformations that are critical for ecosystem function. Nitrogen uptake and assimilation, nitrogen fixation, phosphorus stress, and Fe metabolism are examples for which there are already existing probe strategies [22-26]. There are many other similar biogeochemical

properties of microbial communities that could serve as targets for assays, including biogeochemically relevant genes or specific microorganisms that are keystone in microbial food webs. The goal of this workshop was to begin to identify a larger probe set that could be developed and applied by the scientific community, rather than in individual efforts studying a specific process.

2. Create a cooperative, community level effort to develop comprehensive microbial targets and analytical tools

The design and implementation of a comprehensive suite of probes for molecular targets requires extensive knowledge and expertise. In order to develop a comprehensive toolset, the scientific community needs to work in a coordinated manner, drawing on the vast expertise of experts on organisms, biogeochemical transformations and microbial food webs. Such a coordinated effort could result in probe suites and remote instrumentation that comprehensively sampled the ocean habitat(s) for a variety of biogeochemical parameters.

3. Prioritize targets

There are two molecular probe strategies that will benefit the study and monitoring of microbial biogeochemical function. First, high throughput hybridization and qPCR arrays could facilitate the routine use of a comprehensive probe suite by many investigators, even those without expertise in the specific genes. Second, remote instrumentation can be used now to obtain near real-time information on a limited number, but key biological variables. It is possible to envision high-throughput lab- or ship-based analytical systems that can analyze up to tens of thousands of genetic features simultaneously. In parallel it is important to begin to implement probes using the more limited throughput but higher resolution remote instrumentation approaches. For both approaches, a critical step in implementing is the choice of targets that provide the most appropriate, meaningful and interpretable data, relevant to marine microbial population growth and its relationship to biogeochemical cycles. Prioritization of the many potential microorganism or gene targets is an important but complex and challenging issue that was the subject of this workshop.

4. Contextual biogeochemical measurements

The implementation of probe strategies for microbial gene targets requires appropriate contextual information for interpretation. For example, perhaps nitrate concentrations are necessary to provide the context for ammonia oxidation genes or nitrate assimilation genes, or phosphate concentrations in order to interpret phosphorus limitation (stress) genes. As part of the selection and design of targets for probe strategies, the necessary contextual information needs to be considered, as collecting and analyzing chemistry or isotope tracer experiments puts additional constraints on the strategy for implementation.

5. Compare/contrast needs for different ocean habitats

Many components of microbial communities and biogeochemical cycles are common across ocean habitats, but some are not. Probe strategies, particularly in the comprehensive approach considered here, require development for specific ocean habitats, although some targets may be similar or identical across habitats (for example, *Prochlorococcus* or *Pelagibacter*

ubique gene targets). Unique habitats in the open ocean include for example, oxygen minimum zones. The workshop addressed the issues of specific habitats and the relevant probe strategies that are unique to these habitats.

6. Strategy for future

The implementation of comprehensive probe arrays, lab and ship-based, and deployed on remote instrumentation or autonomous systems, promises to yield important information on the critical roles of microorganisms in the environment, the factors that control their growth and activity, and how they are changing over long time scales. Implementing this vision requires substantial investment in effort and resources to devise informative tools that can be deployed in different marine habitats and by different investigators or research programs. Strategies for obtaining scientific community input and performing critical steps in development and implementation are needed in order to expedite the process, and lead to implementation on time scales needed for informed ocean ecosystem management decisions.

7. Timeline for implementation

The MicroTOOLs goals require short-, mid- and long-term objectives. Short-term objectives include testing of probe strategies on remote platforms to determine major issues on *in situ* equipment, but also to evaluate how molecular target information can be interpreted in the context of high spatial and temporal complexity of the ocean. Short-term objectives also include testing probes as part of time series and on spatial transects, in order to evaluate their utility for studying microbial communities and aid in prioritization of different probe strategies. Mid- and long-term objectives include implementation of larger numbers of probes in concerted efforts to obtain genetic information that is informative for a comprehensive biogeochemical analysis, rather than focus on individual processes. In parallel, development of new sensors and samplers will be required to meet needs of a broad research and resource management community.

8. Future roles and opportunities for funding

The successful deployment of the MicroTOOLs strategy requires coordination of various institutions, laboratories, and multidisciplinary teams in order to implement a cohesive approach applicable to different habitats. Although much progress can be made with the collaborative work of individuals with extant funding, ultimately additional funding may help to expedite implementation of comprehensive gene arrays for assessing microbial function in multiple oceanic regimes.

III. Current technologies

A. Molecular/genomic approaches

Technical advances in the biological sciences over the past twenty years have provided the means to interrogate microorganisms for specific genes, composition and organization of genomes, gene expression and regulation, composition and activity of proteins and enzymes, and the composition of metabolites within the cell. In the post-genomics era, many of these

approaches have become high throughput, providing the ability to assay thousands of gene products or gene transcripts simultaneously. A variety of molecular approaches have been developed to assess microbial diversity, such as TRFLP [27], ARISA [28], and more recently the high throughput sequencing of variable regions of rRNA [29]. High density microarrays can provide information on gene presence or whole genome expression [30]. High-density arrays have been used to obtain diversity information on biogeochemical function in terrestrial microbial communities (the "GEOCHIP" [31], Huang et al. 2007}. High throughput sequencing has provided the means to obtain genomic and gene expression information from complex microbial communities [32-35]. Methods for assaying for the presence of individual genes and gene transcription products have been used to probe microbial communities in the marine environment [36-38]. It is now possible to design high throughput molecular or proteomic approaches to obtain new information on how the ocean microbial systems function, and how they interact with their environment. The goal now is to take this information on diversity, and more specifically, biogeochemistry, and use this information to monitor and study the microbial communities from a biogeochemical system function perspective.

B. Remote instrumentation

Achievements in molecular analyses have fueled the idea of applying "molecular sensors" in remote settings [17]. This vision draws from many, including concepts outlined by Doney et al. [39], Devereux et al. [40], Bowler et al. [41] and the "sensor robot" as schematized in Figure 1. Scholin et al. [21] advanced the use of the phrase "ecogenomic sensor" as being that field-deployable instrument which combines autonomous sample collection with molecular analytical capability and data telemetry. These systems were conceived of as providing assessments of the presence and abundance of a specified set of organisms, their genes and metabolites in near real-time, in time-series fashion. The Autonomous Microbial Genosensor (AMG, [17]) and Environmental Sample Processor (ESP, [18]) both provide a glimpse into the work that has gone on to realize the ecogenomic sensor vision [21]. Autonomous *in situ* application of low density DNA and protein arrays, Nucleic Acid Sequence Based Amplification (NASBA), and quantitative PCR has been demonstrated. Ecogenomic sensors can now be deployed in a variety of settings and embedded within an ocean observatory infrastructure. For example, the ESP has been deployed on coastal moorings, piers, and the deep-sea (Scholin et al. 2009 and unpublished).

With the capacity to analyze material in real-time and archive samples for later analysis in a laboratory, ecogenomic sensors will create opportunities for remotely selecting when and where discrete measurements and samples are acquired in response to a hierarchical set of environmental variables. When this capability is combined with modeling, it should be possible to generate and test hypotheses without a strict requirement for being physically present at-sea to acquire necessary samples. These same systems can also enhance ship-based expeditions by providing additional sampling and analytical power within a region around a vessel.

C. Models, and links between models and microbiology

Models provide a platform in which to synthesize and quantify conceptual understanding. Ocean models can be also used to interpolate and extrapolate sparse observations. Laboratory and field studies of marine microbes are expensive and labor intensive, even with the aid of automated underwater vehicle (AUV) and similar platforms. Mathematical and numerical models provide a tool with which to understand and quantify the large-scale ecological and biogeochemical implications of processes and characteristics as well as detailed interactions at the cellular and metabolic pathway level. Large-scale ocean models may include explicit representations of the physical environment (including dispersal by the circulation), resource environment (light and nutrients), and predatory context. Shaped by, and feeding back upon, these environmental factors, modeled microbial communities must be represented by parameterizations of the physiological trade-offs which govern their response and adaptability. The physiological trade-offs which relate resource (energy and nutrient) allocation at the individual level may ultimately be captured in models of metabolic pathways, rooted in the genetic encoding of particular organisms and ecotypes. Ultimately, such detailed models, linked with laboratory studies, will provide the backbone for parameterizations of traits and trade-offs used in coarse-grained descriptions which are economical for ecological and biogeochemical studies.

In the near future, models of marine microbes and microbial communities must proceed along these two intersecting trajectories:

(i) Ecological and biogeochemical studies, focused on simulating, interpreting and rooted in highly parameterized, computationally efficient descriptions of microbial physiology.

(ii) Detailed physiological models which resolve single cells or small, laboratory populations in significant detail and from which parameters can be determined to apply to large scale environmental studies.

Mathematical models provide the means to extend observations over basin-scale, fore- and hind-cast and develop hypotheses. Even advanced technologies (like the ESP which will provide unparalleled biological time-series at a few oceanographic locations) will not provide useful information far from the sampling location. Mathematical models of the ocean are designed to represent or simulate ocean physics, biology and chemistry to a sufficient degree to provide estimates of key quantities of interest (e.g., standing stock or cell count of organisms or rate processes that control biogeochemical fluxes). Many ocean models do so in 3-space dimensions as well as time. Traditional biogeochemical models were primarily based on nutrients and grazing (N-P-Z) with parameterized variables, and did not include microbial components or microbial diversity. In the past decade the need for inclusion of microbial components, including genomics, has been recognized [39], and a number of different modeling approaches have addressed microbial distributions and activities [42-47] .

Mathematical modeling development is a critical component of the development of the MicroTOOLs approach. Models make several important contributions in assessing and interpreting environmental molecular and biological information. Models that incorporate the dynamics of physics and chemistry can provide important information for informed placement and sampling strategies for microbial measurements, be they from ships or remote instrumentation. Further, models provide the means to take the information sampled from ships or remote instrumentation, and extend the results over time- and space-scales so that the regional or basin-scale implications can be determined. Model output, for example the recent report on distribution of N_2-fixing microorganisms (Goebel et al. [43]), can provide testable hypotheses that can be examined with contextual information or targeted sampling programs.

IV. Applications and promise of molecular approaches for biogeochemistry

A. Definition of molecular "targets"

Environmental molecular targets are biosignatures (gene, RNA, protein) that provide information on the ecological status (stress, nutrient limitation, growth, biogeochemical cycling) of natural microbial communities. Gene targets can be used for proxies of organism abundance, but also for the presence of key genes whose transcription or protein products are involved in key biogeochemical transformations. Some genes (for example ammonia oxidation or nitrogen fixation genes, are not present in all organisms) so the detection of these genes can indicate a selection for that biogeochemical process. Better targets are RNA or protein (or enzyme activity), since these molecules mean the genes are being transcribed and thus the presence of the target is more likely to be correlated with the presence of the biogeochemical activity (although not necessarily the rate). RNA or protein targets, or a suite of RNA and proteins, can be indicators of key ecosystem parameters such as species growth rates, cell stress or death or nutrient limitation.

There are numerous examples of molecular probes that have already been successfully implemented to access microbial composition and activities in the marine environment (see example of available and applicable cyanobacterial probes in Appendix I, Table 3). These probes were developed for PCR to identify diversity and abundance of microorganisms (such as 16S-rRNA [48], *rbcL* [49], *rpoC* [50]), and/or for RT-PCR to determine the activity and physiology of microorganisms (such as *ftsZ* for cell cycle protein [51], *nifH* for nitrogen fixation [36,52-53], *ntcA* for nitrogen limitation [54], *phoX* for phosphorus stress [55]). Depending on the question, molecular probes for a target can be specific to a phylum, class, genera, and/or species. For example, *nifH* degenerate primers identify the presence and expression of the *nifH* gene from a majority of microorganisms that fix nitrogen [36], or group-specific *nifH* primers can differentiate between groups of organisms that fix nitrogen (e.g. Church et al. [37]). This concept can be extended to all of the major components of marine biogeochemical cycles, including ammonia oxidation, iron

(Fe) metabolism and limitation, utilization of different phosphorus sources (e.g. phosphonate vs. inorganic phosphate), and many others.

B. Implementation of molecular approaches for targets.

Water masses are defined by disparate physical and chemical parameters, and each water body has a distinct microbial component. Determining those environmental factors that define the presence and absence of microbial taxa in distinct water masses is the first step in evaluating their role in marine biogeochemistry over space and time. In traditional microbial oceanographic studies, samples are only taken on a monthly basis and at few sites/depths per region. There is therefore a great limitation of data for the vast suite of environmental parameters determining microbial abundances, in addition to the limitation of data on the abundances themselves (Karl et al., 1998), precluding our understanding of spatial and temporal dynamics. In addition, much can be gained by high resolution analyses of expression profiles of microbial communities *via* examination of RNA or protein dynamics.

Once targets are identified, probes can be developed for qPCR/RT-PCR, protein assays or microarrays. Microarrays allow robust and sensitive identification of species diversity and gene expression, but are suitable for analysis of microorganisms ranging in abundance from average to high. A PCR/RT-PCR approach is preferred for the low abundant microorganisms, and a plate approach can be implemented for high-throughput. Since PCR can be performed both on ship and in the lab, a set of PCR/RT-PCR targets will be defined as a universal set (although it may differ for various marine environments) that can be used on a ship as part of the core sample analyses.

In situ instrumentation potential applications

As noted earlier, existing instrumentation that allows for interrogation of microbial communities *in situ* employs rRNA hybridization arrays, ELISA protein assays, NASBA, and qPCR. Samples can also be archived for more in-depth analyses upon instrument retrieval. Deployments are coupled with physical and chemical sensors to provide an appropriate environmental framework against which molecular analyses and archives can be referenced. Real-time capability gives the investigator the option of event response, for example by altering the function of the remote instrument to perform specific analyses or by changing sampling regimes. Access to real-time data also allows for directing collection of additional samples for more comprehensive analyses and experimentation not possible using remote systems alone.

Application of hybridization arrays provide microbial fingerprinting of water masses and are discussed below. Quantitative PCR on the ESP delivers high-sensitivity (down to 10s of organisms), high-resolution (hourly/daily) abundances of groups of interest and can be used with both 16S rRNA and functional gene primer/probe sets in order to quantify the subclades of

organisms performing a function of interest (Preston *et al*., unpublished). Molecular probes that are currently in use on ESP or are being developed are shown in Appendix I (Tables 1 and 2). Both types of data, from hybridization arrays and from qPCR, can be coupled to physical and chemical data to gain knowledge of the environmental limitations of microbial groups. This data can be fed into models in order to achieve more reliable forecasts in microbial oceanography (Follows et al., 2007).

As a tool to measure microbial activities, ELISA protein assays have been used successfully to detect the toxin domoic acid from eukaryotic phytoplankton [56]. This detection capability can be used for quantification of abundant proteins in the marine environment, though its potential for rare proteins from the marine environment has not yet been proven. ELISA assays are a rapid means of determining microbial activities by detecting the proteins that are directly responsible for carrying out processes, and coupled to hybridization arrays, can provide data on the abundance and activity of ecologically-important microorganisms in a single sample. NASBA has also been used as a means of detecting viral RNA *in situ* (Patterson et al., 2006) and its applications to assess activities of prokaryotic phytoplankton *via* changes in transcription are currently being evaluated.

In situ biological instrumentation can gain from developments in contextual sensors. The suite of deployable physical and chemical sensors is limited and real-time data on the concentrations of various inputs and outputs of microbial metabolism would allow investigators to more clearly discern the role of chemistry in shaping biology and vice-versa.

High throughput, multi-target ship- or lab-based approaches

DNA microarrays are currently applied in two main formats, both of which are amenable to high throughput analysis on archived samples. Such samples could be collected on a "ships-of-opportunity" basis or from moorings that can collect large numbers of samples for later analysis on-shore.

i) Genome, genome proxy or partial genome arrays: These arrays contain probes derived from single organisms, or from genomic fragments representing unknown organisms detected in field samples. By hybridizing them with DNA or RNA extracted from field samples, the distribution and relative abundance of the target organism or close relatives can be determined to yield temporal and spatial patterns (biogeography).

ii) Functional gene arrays: The probes in this case represent specific functional genes and can be synthesized or derived by PCR amplification of known genes from natural samples and cultures. To the extent that specific genes relate to particular biogeochemical processes, information on the genetic capability and gene expression for these processes can be analyzed in the context of

physical and chemical factors to investigate ecosystem function. They can be used in manipulative experiments to investigate short term response or in natural samples to monitor ecosystem response to perturbations and anthropogenic change. Functional gene arrays can go beyond biogeography to investigate transient responses to environmental variables by quantifying gene expression among different phylotypes.

Phylogenetic and functional DNA microarray have been increasingly used in marine ecology over the past 10 years (for example, for N cycle -Taroncher-Oldenburg et al.[57]; nitrogen fixation - Moisander et al. [58]; ammonium oxidation - Ward et al. [38]; genome-proxi - Rich et al. [59]). However, high density arrays such as PhyloChip (Brodie et al. [60]) and GeoChip (He et al. [31]), that have been developed and applied for soil environment, have not yet been utilized for marine environments. These arrays feature tens to hundreds of thousands of probes, and allow highly specific identification of genes that differ only in a few nucleotides. Development of high density phylogenetic and functional microarrays will increase the robustness and sensitivity of analyses in marine environments.

A few obvious steps could take these approaches to the next level. 1) Sequence the culture collection for key functional genes of interest. At present we can detect many more different functional genes in the environment than we can identify beyond generally phylogenetic affinity or align with known organisms. Without information about the organism they represent, it is difficult to make ecological inferences. It is much faster and cheaper to sequence a few functional genes than a whole eukaryotic genome, so this investment would go far to improve interpretation of field samples without incurring major sequencing costs. 2) Decide on which functional genes or proxies are of highest priority for array development and commission both the sequencing effort required to build a sufficient database, and the bioinformatics analysis to design arrays based on them.

V. Identification of targets and applications

The development of a MicroTOOLs strategy involves consideration of organisms and habitats. Targets such as *Synechococcus* or *Prochlorococcus* may provide good sentinels of ecosystem state. Processes that deserve focus and attention, in addition to the key N, P and Fe cycles, are organic matter decomposition and cell mortality. Some habitats provide good sites for implementation of a MicroTOOLS approach because of the nature of the habitat or because existing research or monitoring programs are in place. Some examples of key organism types that are good targets are the picocyanobacteria and the eukaryotic phytoplankton. The following sections provide examples of MicroTOOLs project foci.

A. Organisms

Picocyanobacteria

Cyanobacteria such as *Prochlorococcus* and *Synechococcus* are major primary producers across diverse ecosystems [61-63]. These are an ideal group of organisms to begin to implement molecular tools for investigating the distribution of specific taxa as a function of geography and time. Marker genes such a 16S rRNA, RNA polymerase (*rpoC*), 16S-23S rRNA intergenic spacer, *ntcA* gene, etc. have provided a glimpse into the community structure of these organisms, and probe development using these genes is being implemented using high throughput techniques and *in situ* monitoring instruments. In addition, due to the funding from NSF, JGI, Genescope, and the Moore Foundation, the genomes of multiple picocyanobacterial species are available. This facilitates the investigation of genome-wide gene expression under environmental stresses such as nitrogen limitation, phosphate limitation [64-66], etc. In turn, this work promotes the development of robust molecular probes for marker genes of specific environmental conditions and these could in turn be implemented on *in situ* monitoring instruments. This will allow, in the relatively near future, the measurement of the biological conditions of specific taxa in relation to the physiochemical conditions in which they find themselves. In other words, this will allow a real time measurement of the state and health of marine ecosystems.

Eukaryotic phytoplankton

Eukaryotic phytoplankton support roughly one-half of global primary production, greatly impact the biogeochemistry of different macronutrients and trace elements, and act as a critical link in marine food webs. In coastal systems, for example, diatoms are often the primary drivers of carbon fixation and export. Phytoplankton activities and their resultant biogeochemical impacts are hard to constrain, and one of the ongoing challenges in biological oceanography is to understand the links between phytoplankton community structure, function, and biogeochemical processes well enough to predict how the system will respond in the future. Looking towards the future, these assays should be linked with 'key genes' to serve as expression targets and act as 'canaries' for how cells experience their environment – providing feedback on the system changes that are affecting population dynamics. Molecular tools already show great promise in this regard, and tracking of specific proteins (e.g. alkaline phosphatase, flavodoxin) has been used with success to monitor the physiology and activities of eukaryotic phytoplankton *in situ*. One challenge with the marine eukaryotic phytoplankton, is that whole genome sequences, and metagenome/transcriptome datasets are much more limited relative to marine cyanobacteria. With these limitations, much of the previous gene/protein probe work has focused on a few targets in a few test species, where the sequences, regulation patterns, activities etc. could be ascertained in detail. Thus presently, a concerted effort to render the current state of knowledge regarding eukaryotic phytoplankton genomes, to tools for assessing physiological status of the cell or population is needed. Studies to identify basic expression patterns of functional genes predicted to respond to biogeochemical and other forcing is an important initial step towards this

goal. This synthesis will provide a platform for comparing and developing more potential targets, and this will improve as the sequence databases for the eukaryotic phytoplankton expands. A large number of genes in sequenced genomes, and metagenomic datasets, are still of unknown function. These genes may hold novel information about currently overlooked biological drivers, but are difficult to express and study in laboratory conditions. Therefore a structured approach that incorporates genes of unknown function, and those of known function, may serve as research tools, enhancing future efforts on ocean sensing. Taken together, these integrated data types on taxonomy and expressional responses should greatly benefit modeling efforts that still lack relevant information on biological drivers of change.

Due to their size, intrinsic fluorescence, and unique ultrastructure, there is a rich literature on the presence and population dynamics of larger cell-sized eukaryotic groups. Resolution of eukaryotic populations in general, is increasingly aided by advances in flow cytometry (e.g. image flow cytobot) and molecular detection and tracking (e.g. qPCR, sandwich hybridization). In fact, the presence and dynamics of toxic phytoplankton species that have particular relevance to the oceans and human health are increasingly assayed using molecular tools. The high throughput processing of the image flow cytobot has enhanced our understanding of the fine scale dynamics of the eukaryotic phytoplankton in their natural environment. Similarly, increased molecular screening (e.g. via tools like the ESP, once it is more widely available with a well-validated, commonly shared set of microarrays or qPCR methods) could provide molecular information to complement the image flow cytobot, or stand alone to track key groups or species. Historically, most of the probes target rDNA for these applications. However other genes such as RuBisCO and various functional genes can provide insights to both population structure and serve as first indicators of functional aspects. Though not as comprehensive as prokaryotic genome databases, eukaryotic sequence databases are sufficiently strong that it should be possible to develop a comprehensive screen, whether via a traditional microarray, or newer technologies, for evaluating relative change in different eukaryotic phytoplankton populations with higher temporal, spatial and taxonomic resolution than previously possible.

B. Mortality/decomposition

Two major, opposing fluxes in the global carbon cycle are primary production and remineralization of autochthonous and allochthonous organic matter in the oceans. Primary production can be represented in biogeochemical models by equations that do a reasonably good job of expressing its dependence on major controlling factors (e.g., light, nutrients, temperature) given an estimate of phytoplankton biomass. While the accuracy of primary production measurements and estimates is still being improved, our mechanistic understanding of photosynthesis has made the relatively simple measurement of bulk chlorophyll a very useful oceanographic tool. In contrast, the decomposition and remineralization of organic matter remains less well understood. There are many reasons for the relatively advanced state of our understanding of primary production, not least of which are our good understanding of the

biochemistry and physiology of photosynthesis, that photosynthesis works in fundamentally the same way in all oxygenic phototrophs, and that oxygenic phototrophs generally use the same substrates (light, N, P, inorganic C). In contrast to the similarity of basic growth processes among phototrophs, there are multiple distinct major decomposition and loss processes that reflect the diverse and poorly characterized suite of dissolved and particulate organic matter serving as potential heterotrophic substrates, predation by a variety of consumers and lysis by a variety of viruses. Combined, these processes involve a complex array of organic compounds and particles that are transformed via many distinct pathways, reflecting the great phylogenetic and physiological diversity of heterotrophs in the oceans. To improve our understanding of C and other elemental cycles in the oceans, we need to develop more simple, routine, and high throughput ways to measure these loss processes or proxies for them. The ultimate goal would be the development of process, community structure and function assays that can be automated and deployed on buoys, moorings, or in the underway seawater monitoring systems of oceanographic research vessels.

C. Habitats

A number of habitats provide important sites for implementation of a MicroTOOLs probe approach. A selection of habitats include the open ocean (Cape Verde, BATS, HOT), Monterey Bay, Martha's Vineyard, the California Current, Saanich Inlet, the Gulf of Mexico, the Costa Rica Dome, and Chesapeake Bay. These specific sites are mentioned since they have characteristics that are representative of important types of habitats (e.g. estuary, suboxic waters, open ocean, etc), and also may have existing programs in place that provide the backbone for deployment and provision of ancillary data.

Monterey Bay. Monterey Bay is an ideal site for tracking microbial dynamics due to its accessibility, the wealth of time series data available and the various physical processes that drive the biology in this region. Coastal regions are influenced by rainfall and agricultural runoff, which fertilizes the marine environment in this nitrogen-limited ecosystem (Kudela and Dugdale, 2000; Los Huertos *et al*., 2001). Sites in the bay are affected by coastal upwelling, which happens on a seasonal basis and typically corresponds with blooms of eukaryotic phytoplankton. Prokaryotic phytoplankton (*Synechococcus* sp.) are abundant in the outer bay (Paerl *et al*., in review) but also bloom periodically inshore, and the factors determining when these blooms occur are currently unknown. Time series observations have been made at stations within the bay by MBARI (Pennington *et al., in prep* [67]), and along 'CalCOFI Line 67,' which transects the bay, moving from high- to low-nutrient waters and thereby allowing comparisons of coastal ecosystems with more typical open ocean systems (http://www.calcofi.org/). The Monterey Bay Aquarium Research Institute houses ships and Autonomous Underwater Vehicles locally, which have been used to create four-dimensional reconstructions of the water column's physical and chemical properties at study sites. Detailed physical models of the California Current have been incorporated into emergent ecosystem models in order to predict the dominant phytoplankton ecotypes over time and space (Goebel *et al., in review* [68]), and the results of the model agree well

with observations, which are few relative to the models' capacities. We are at a point in Monterey Bay where the predictive capacity of the science is limited by the resolution of biological observations and gaining more information on the ecotypes present in distinct water masses will allow us to define those factors determining the relative success of groups of organisms, *in situ*. We have just begun to obtain daily resolution on functional genes involved in nitrate assimilation and carbon fixation by *Synechococcus*, and in ammonia oxidation by the *Crenarcheaota*, using the ESP (Robidart *et al*., in prep). These are just a few pieces of the nitrogen cycle puzzle and now we must transition to gain information about additional key players and processes (e.g. P, C cycling) in order to gain a more comprehensive view of factors driving Monterey Bay biogeochemistry over space and time.

The California Current. The California Cooperative Oceanic Fisheries Investigations (CalCOFI) are a partnership of the California Department of Fish and Game, the NOAA Fisheries Service and the Scripps Institution of Oceanography, whose focus is the study of the marine environment off the coast of California and the management of its living resources (http://www.calcofi.org/). Since 1949, CalCOFI has undertaken quarterly cruises sampling a grid of 75 stations to measure the physical and chemical properties of the California Current System and census populations of organisms from phytoplankton to avifauna. This is the foremost observational oceanography program in the United States. The NSF funded California Current (CCE) LTER shares some PIs with this program and is involved in characterizing changes in biogeochemical processes in this region over time. Initial deep sequencing of ribosomal rRNA amplicons as carried out in collaboration with the MIRADA project is providing an overview of the major bacterial, archaeal, and eukaryotic taxa present as well as the potentially important "rarer" taxa with important biogeochemical functions. The California Current would thus be an ideal ecosystem in which to introduce routine monitoring with molecular probes.

D. Equipment and standardization

In order to implement a scientific community-wide comprehensive probe strategy, laboratory procedures need to be selected that can be implemented across laboratories or developed as a centralized facility.

Currently there are only two remote instruments that can be deployed, the AMG and the ESP. The ESP is higher throughput and can be deployed for longer periods. Equipment selection is therefore limited, and testing and standardization needs to be done on the ESP or AMG modules, prior to deployment.

High throughput approaches, in contrast, provide many different methods (e.g. Taqman vs. molecular beacon quantitative polymerase chain reaction; different types of microarrays). In order to develop a community resource, choices of compatible instruments, sample handling and standardization methodologies will need to be made. Issues include sample size (which involves

limit of detection considerations) extraction efficiencies for different organisms and sample types, RNA or DNA amplification (necessary for microarray approaches), expected target concentrations, and how to develop standards for determining extraction and amplification efficiencies, and quantification of assay results. Development of standards needs to go hand-in-hand with assay development.

E. Databases

The MicroTOOLs initiative in ecogenomics has the potential to generate a large volume of data for target sequences under study. Procedures need to be developed to obtain and store data, along with quality control information and metadata (standardized location coordinates, depth, time, etc) in order to facilitate analysis of results by the scientific community. A stand-alone sequence database will need to be established for primer design. After that, an OLTP (on-line transaction processing) relational database will be needed for accumulating real-time data from sample sites. The "atomic" or "lowest level" data in the OLTP database will need to identify the sample location and time, and the data being collected. These transactions will need to be related to a file of sample locations and times containing physical (temperature, density, etc.) and chemical (salinity, pH, O_2 level, etc.) information. An annotation for the transaction will also need to be generated, identifying the enzyme, ribozyme, or structural RNA it represents, and the taxon from which it likely came. "Logical views" of the data in the relational OLTP database, consisting of linked data elements from samples and associated metadata, can then be constructed for input into analytical packages for gene expression analysis, phylogenomics, or any other open source sequence analysis tool. If there is sufficient interest and funding, an "Extract, Load, and Transform" tool could be used to transform the OLTP data into an OLAP (on-line analytical processing) "data cube" for interactive analysis of abundance of target sequences along taxonomic and biochemical hierarchies at any desired level in the hierarchy.

VI. Prioritization

Identification of probes for biogeochemistry involves selection from a wide range of targets for enzymes, transporters, signal proteins across a wide array of Bacteria, Archaea and Eukaryotes. Thousands of potential targets can be identified, but emphasis needs to be placed on selecting the probe targets that will provide the most meaningful ecological information on key biogeochemical and ecological characteristics of microbial communities. Potential probes range from those that are currently well-developed to those that may provide useful approaches for the future. Thus, prioritization for selection must evaluate both current status of information and stage of development vs. the potential information that probe strategies may provide.

VII. Needs for implementation

Currently a number of probes have, or are being developed for implementation on remote instrumentation, at the UCSC MEGAMER facility and a number of other laboratories (please see Appendices). In order to develop a comprehensive MicroTOOLs approach:

1. Key biogeochemical processes need to be identified for targets

2. High throughput equipment and procedures need to be selected by the scientific community.

3. Probes need to be identified and probe strategies need to be devised to target the relevant organism or organisms, to evaluate habitat specificity issues, and to determine the molecule (DNA, RNA, protein or other) that provides useful information.

4. Probe approaches need to be tested in culture and the environment. Probe testing includes determining that signals from the environment provide useful ecological information (for example in transect or time-series samples), and that the probe has the appropriate level of specificity.

5. The molecular microbial community needs to organize, develop a strategy, pipeline and timeline for implementing probe approaches in a concerted manner. This will involve short-term goals for implementing existing probes in conjunction with existing study sites, and developing high throughput approaches. As a result of this meeting, working groups have been developed and are working on these goals. Longer-term goals will prioritize and prepare for larger scale implementation on remote instrumentation.

VIII. Summary

Groups within the workshop discussed specifics on environments and organisms of interest, regarding how to best study systems in order to gain the most comprehensive picture of marine biogeochemistry, and this document only highlights a few specific cases. It was agreed among those present that where it is available, metagenomic and metatranscriptomic sequences provide a wealth of data from which molecular tools can be developed, and that environments where such sequence information is available allow investigators to 'hit the ground running' when it comes to biogeochemical investigations. As a result of this workshop, working groups have assembled to develop suites of targets for specific processes and for individual habitats. These groups are organized to move forward on executing the MicroTOOLs approach in a variety of habitats and targeting organisms and processes.

It is timely to accelerate development of molecular probe strategies for characterizing the biogeochemistry of the seas. Molecular (in the sense of DNA, RNA or protein) indicators can provide information on the ecosystem status, the activity levels of microorganisms, biogeochemical process potential and activity, and the microbial species involved. Much work is needed to identify and test individual molecular probes, and it requires the collective expertise of the broader marine microbiological community. However, the coordinated and collaborative development of a comprehensive microbial "toolbox" will facilitate both cross-system comparative studies, as well as provide meaningful tools for time-series analysis at individual ocean sites. This comprehensive approach would be the first global extension of molecular

biology techniques to microbial community analysis across biogeochemical cycles and across habitats, and promises to provide the information and tools needed to deploy remote instrumentation at key ocean sites. This information will be key in understanding the dynamics of ocean chemistry and biology, but more importantly provide the baseline for detecting how the ocean changes on annual, decadal and longer time scales. The time is now to develop this community wide effort to take microbial oceanography from the genetic discovery phase to biogeochemical observation, analysis and prediction.

REFERENCES

1 *Halpern, B., Walbridge, S., Selkoe, K. & Kappel, C. A global map of human impact on marine ecosystems. Science (2008).*

2 *Azam, F. & Malfatti, F. Microbial structuring of marine ecosystems. Nature Reviews Microbiology (2007).*

3 *Sommer, U., Stibor, H., Katechakis, A. & Sommer, F. Pelagic food web configurations at different levels of nutrient richness and their …. Hydrobiologia (2002).*

4 *Calbet, A. & Saiz, E. The ciliate-copepod link in marine ecosystems. Aquatic Microbial Ecology (2005).*

5 *Rappe, M. & Giovannoni, S. The uncultured microbial majority. Annual Reviews (2003).*

6 *Pedrós-Alió, C. Marine microbial diversity: can it be determined? Trends in Microbiology **14**, 257-263 (2006).*

7 *DeLong, E. The microbial ocean from genomes to biomes. Nature (2009).*

8 *Karl, T. & Trenberth, K. Modern global climate change. Science (2003).*

9 *Dore, J. E., Lukas, R., Sadler, D. W., Church, M. J. & Karl, D. M. Physical and biogeochemical modulation of ocean acidification in the central North Pacific. Proc Natl Acad Sci USA **106**, 12235-12240, doi:10.1073/pnas.0906044106 (2009).*

10 *Brander, K. Global fish production and climate change. Proceedings of the National Academy of … (2007).*

11 *Behrenfeld, M. J. et al. Controls on tropical Pacific Ocean productivity revealed through nutrient stress diagnostics. Nature **442**, 1025 (2006).*

12 *Karl, D. Minireviews: A Sea of Change: Biogeochemical Variability in the North Pacific …. Ecosystems (1999).*

13 *Karl, D. Nutrient dynamics in the deep blue sea. TRENDS in Microbiology **10**, 410-418 (2002).*

14 *Church, M., Ducklow, H. & Karl, D. Multiyear increases in dissolved organic matter inventories at Station …. Limnology and Oceanography (2002).*

15 *Sabine, C. L. et al. The oceanic sink for anthropogenic CO2. Science **305**, 367-371 (2004).*

16 *Falkowski, P. G., Fenchel, T. & Delong, E. F. The microbial engines that drive Earth's biogeochemical cycles. Science **320**, 1034-1039, doi:10.1126/science.1153213 (2008).*

17 *Paul, J., Scholin, C., Van den Engh, G. & Perry, M. J. In Situ Instrumentation. Oceanography **20**, 70-78 (2007).*

18 *Scholin, C. et al. Remote detection of marine microbes, small invertebrates, harmful algae, and biotoxins using the Environmental Sample Processor (ESP). Oceanography **22**, 158-167 (2009).*

19 *Zehr, J. P. & Hiorns, W. D. in Molecular Approaches to the Study of the Ocean (ed K E Cooksey) 91-112 (Chapman and Hall, 1998).*

20 *Zehr, J. P., Hewson, I. & Moisander, P. H. Molecular biology techniques and applications for ocean sensing. Ocean Science **5**, 101-113 (2009).*

21 *Scholin, C. A. What are "ecogenomic sensors?" A review and thoughts for the future. Ocean Science **6**, 51-60 (2010).*

22 *Allen, A. E. et al. Diversity and detection of nitrate assimilation genes in marine bacteria. Appl Environ Microbiol **67**, 5343-5348 (2001).*

23 *Webb, E. A., Jakuba, R. W., Moffett, J. W. & Dyhrman, S. T. Molecular assessment of phosphorus and iron physiology in Trichodesmium populations from the western Central and western South Atlantic. Limnol Oceanogr **52**, 2221-2232 (2007).*

24 *Paerl, R. W., Foster, R. A., Jenkins, B. D., Montoya, J. P. & Zehr, J. P. Phylogenetic diversity of cyanbacterial narB genes from various marine habitats. Environ. Microbiol. **10**, 3377-3387, doi:10.1111/j.1462-2920.2008.01741.x (2008).*

25 *Zehr, J. P., Crumbliss, L. L., Church, M. J., Omoregie, E. O. & Jenkins, B. D. Nitrogenase genes in PCR and RT-PCR reagents: implications for studies of diversity of functional genes. Biotechniques **35**, 996-1005 (2003).*

26 *Chapelle, F. H., Bradley, P. M., Thomas, M. A. & McMahon, P. B. Distinguishing Iron-Reducing from Sulfate-Reducing Conditions. Ground Water **47**, 300-305, doi:10.1111/j.1745-6584.2008.00536.x (2009).*

27 *Carlson, C., Morris, R., Parsons, R. & Treusch, A. Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones …. The ISME … (2008).*

28 *Fuhrman, J. A. Measuring diversity. Accessing Uncultivated Microorganisms:FROM THE ENVIRONMENT TO ORGANISMS AND GENOMES AND BACK, 131-151 (2008).*

29 *Sogin, M. L. et al. Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc Natl Acad Sci USA **103**, 12115-12120 (2006).*

30 *Shi, T., Ilikchyan, I. N., Rabouille, S. & Zehr, J. P. Genome-wide analysis of diel gene expression in the unicellular N_2-fixing cyanobacterium Crocosphaera watsonii WH 8501. ISME J, doi:10.1038/ismej.2009.148 (2009).*

31 *He, Z. et al. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J **1**, 67-77, doi:doi:10.1038/ismej.2007.2 (2007).*

32 *Rusch, D. B. et al. The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biology **5**, e77, doi:doi:10.1371/journal.pbio.0050077 (2007).*

33 *Poretsky, R. S. et al. Analysis of microbial gene transcripts in environmental samples. Appl Environ Microbiol **71**, 4121-4126 (2005).*

34 *DeLong, E. F. et al. Community genomics among stratified microbial assemblages in the ocean's interior. Science **311**, 496-503 (2006).*

35 *Frias-Lopez, J. et al. Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci USA **105**, 3805-3810, doi:10.1073/pnas.0708897105 (2008).*

36 *Zehr, J. P. & McReynolds, L. A. Use of degenerate oligonucleotides for amplification of the nifH gene from the marine cyanobacterium Trichodesmium spp. Appl Environ Microbiol **55**, 2522-2526 (1989).*

37 *Church, M. J., Jenkins, B. D., Karl, D. M. & Zehr, J. P. Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the oligotrophic North Pacific Ocean. Aquat Microb Ecol **38**, 3-14 (2005).*

38 *Ward, B. B. et al. Ammonia-oxidizing bacterial community composition in estuarine and oceanic environments assessed using a functional gene microarray. Environ. Microbiol. **9**, 2522-2538 (2007).*

39 *Doney, S. C., Abbott, M. R., Cullen, J. J., Karl, D. M. & Rothstein, L. From genes to ecosystems: the ocean's new frontier. Frontiers In Ecology and the Environment **2**, 457-466 (2004).*

40 *Devereux, R., Rublee, P., Paul, J. H., Field, K. G. & Santo Domingo, J. W. Development and Applications of Microbial Ecogenomic Indicators for Monitoring Water Quality: Report of a Workshop Assessing the State of the Science, Research Needs and Future Directions. Environ Monit Assess **116**, 459-479 (2006).*

41 *Bowler, C., Karl, D. M. & Colwell, R. R. Microbial oceanography in a sea of opportunity. Nature **459**, 180-184, doi:10.1038/nature08056 (2009).*

42 *Rabouille, S., Edwards, C. A. & Zehr, J. P. Modelling the vertical distribution of Prochlorococcus and Synechococcus in the North Pacific Subtropical Ocean. Environ Microbiol **9**, 2588-2602 (2007).*

43 *Goebel, N. L., Edwards, C. A., Church, M. J. & Zehr, J. P. Modeled contributions of three types of diazotrophs to nitrogen fixation at Station ALOHA. ISME J **1**, 606-619 (2007).*

44 *Goebel, N. L. et al. Growth and carbon content of three different sized diazotrophic cyanobacteria observed in the subtropical North Pacific. J Phycol **44**, 1212-1220 (2008).*

45 *Follows, M. J., Dutkiewicz, S., Grant, S. & Chisholm, S. W. Emergent biogeography of microbial communities in a model ocean. Science **315**, 1843-1846, doi:10.1126/science.1138544 (2007).*

46 *Monteiro, F. & Follows, M. On the interannual variability of nitrogen fixation in the subtropical gyres. Journal of Marine Research (2009).*

47 *Barton, A. D., Dutkiewicz, S., Bragg, J., Flierl, G. & Follows, M. J. Patterns of marine phytoplankton diversity. Science (2010).*

48 *Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol **173**, 697-703 (1991).*

49 *Pichard, S. L., Campbell, L. & Paul, J. H. Diversity of the ribulose bisphosphate carboxylase/oxygenase Form I gene (rbcL) in natural phytoplankton communities. Appl Environ Microbiol **63**, 3600-3606 (1997).*

50 *Palenik, B. & Haselkorn, R. Multiple evolutionary origins of prochlorophytes, the chlorophyll B-containing prokaryotes. Nature **355**, 265-267 (1992).*

51 *Holtzendorff, J. et al. Synchronized expression of ftsZ in natural Prochlorococcus populations of the Red Sea. Environ. Microbiol. **4**, 644-653 (2002).*

52 *Zani, S., Mellon, M. T., Collier, J. L. & Zehr, J. P. Expression of nifH genes in natural microbial assemblages in Lake George, NY detected with RT-PCR. Appl Environ Microbiol **66**, 3119-3124 (2000).*

53 *Church, M. J., Short, C. M., Jenkins, B. D., Karl, D. M. & Zehr, J. P. Temporal patterns of nitrogenase gene (nifH) expression in the oligotrophic North Pacific Ocean. Appl Environ Microbiol **71**, 5362-5370 (2005).*

54 *Lindell, D. & Post, A. F. Ecological aspects of ntcA gene expression and its use as an indicator of the nitrogen status of marine Synechococcus spp. Appl Environ Microbiol **67**, 3340-3349 (2001).*

55 *Sebastian, M. & Ammerman, J. W. The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. Isme Journal **3**, 563-572, doi:10.1038/ismej.2009.10 (2009).*

56 *Doucette, G. J. et al. Remote, subsurface detection of the algal toxin domoic acid onboard the Environmental Sample Processor: Assay development and field trials. Harmful Algae **8**, 880-888, doi:10.1016/j.hal.2009.04.006 (2009).*

57 *Taroncher-Oldenburg, G., Griner, E. M., Francis, C. A. & Ward, B. B. Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. Appl Environ Microbiol **69**, 1159-1171 (2003).*

58 *Moisander, P. H. et al. Application of a nifH oligonucleotide microarray for profiling diversity of N_2-fixing microorganisms in marine microbial mats. Environ. Microbiol. **8**, 1721-1735 (2006).*

59 *Rich, V. I., Konstantinidis, K. & DeLong, E. F. Design and testing of 'genome-proxy' microarrays to profile marine microbial communities. Environ. Microbiol. **10**, 506-521 (2008).*

60 *Brodie, E. L. et al. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl Environ Microbiol **72**, 6288-6298, doi:10.1128/aem.00246-06 (2006).*

61 *Olson, R. J., Chisholm, S. W., Zettler, E. R. & Armbrust, E. V. Pigments, Size, and Distribution of Synechococcus in the North-Atlantic and Pacific Oceans. Limnol Oceanogr **35**, 45-58 (1990).*

62 *Olson, R. J., Chisholm, S. W., Zettler, E. R., Altabet, M. A. & Dusenberry, J. A. Spatial and temporal distributions of prochlorophyte picoplankton in the North-Atlantic Ocean. Deep-Sea Research Part a-Oceanographic Research Papers **37**, 1033-1051 (1990).*

*63 Li, W. K. W. Primary production of prochlorophytes, cyanobacteria, andeucaryotic ultraphytoplankton: Measurements fromflow cytometric sorting. Limnol. Oceanogr **39**, 169-175 (1994).*

*64 Martiny, A. C., Coleman, M. L. & Chisholm, S. W. Phosphate acquisition genes in Prochlorococcus ecotypes: Evidence for genome-wide adaptation. Proc Natl Acad Sci USA **103**, 12552-12557 (2006).*

*65 Tetu, S. G. et al. Microarray analysis of phosphate regulation in the marine cyanobacterium Synechococcus sp WH8102. Isme Journal **3**, 835-849, doi:10.1038/ismej.2009.31 (2009).*

*66 Tolonen, A. C. et al. Global gene expression of Prochlorococcus ecotypes in response to changes in nitrogen availability. Molecular Systems Biology **2**, doi:10.1038/msb4100087 (2006).*

67 Pennington, J. T., Collins, C. A., Michisaki, R. M. & Chavez, F. P. A twenty year oceanographic time series in Monterey Bay, California: Spatial pattern between stations. (in preparation).

68 Goebel, N. L., Edwards, C. A., Zehr, J. P. & Follows, M. J. An emergent community ecosystem model applied to the California Current System Journal of Marine Systems (submitted).

MicroTOOLs I workshop participants

Participants	Institute
Alexandra Worden	Monterey Bay Aquarium Research Institute
Annika Mosier	Stanford University
Anton Post	Woods Hole Oceanographic Institution
Arthur Grossman	Stanford University
Bess Ward	Princeton University
Bethany Jenkins	University of Rhode Island
Brian Palenik	Scripps Institution of Oceanography
Chris Gobler	Stony Brook University
Chris Francis	Stanford University
Chris Scholin	Monterey Bay Aquarium Research Institute
Dave Caron	University of Southern California
David Scanlan	University of Warwick
Eric Webb	University of Southern California
Hans Paerl	University of North Carolina at Chapel Hill
Irina Ilikchyan	UC Santa Cruz
Jackie Collier	Stanford University
Jim Tripp	UC Santa Cruz
Jim Ammerman	Stony Brook University
John Paul	University of South Florida
Jon Zehr	UC Santa Cruz
Julie LaRoche	Leibniz Institute of Marine Sciences
Julie Robidart	UC Santa Cruz
Mak Saito	Woods Hole Oceanographic Institution
Mark Hildebrand	Scripps Institution of Oceanography
Mary Ann Moran	University of Georgia
Matt Church	University of Hawai`i at Manoa
Matt Sullivan	University of Arizona
Mick Follows	Massachusetts Institute of Technology
Phyllis Lam	Max Planck Institute for Marine Microbiology
Pia Moisander	UMass Darfmouth
Ryan Paerl	UC Santa Cruz
Scott Doney	Woods Hole Oceanographic Institution
Sonya Dyhrman	Woods Hole Oceanographic Institution
Steve Hallam	University of British Columbia
Tim Hollibaugh	University of Georgia
Virginia Armbrust	University of Washington