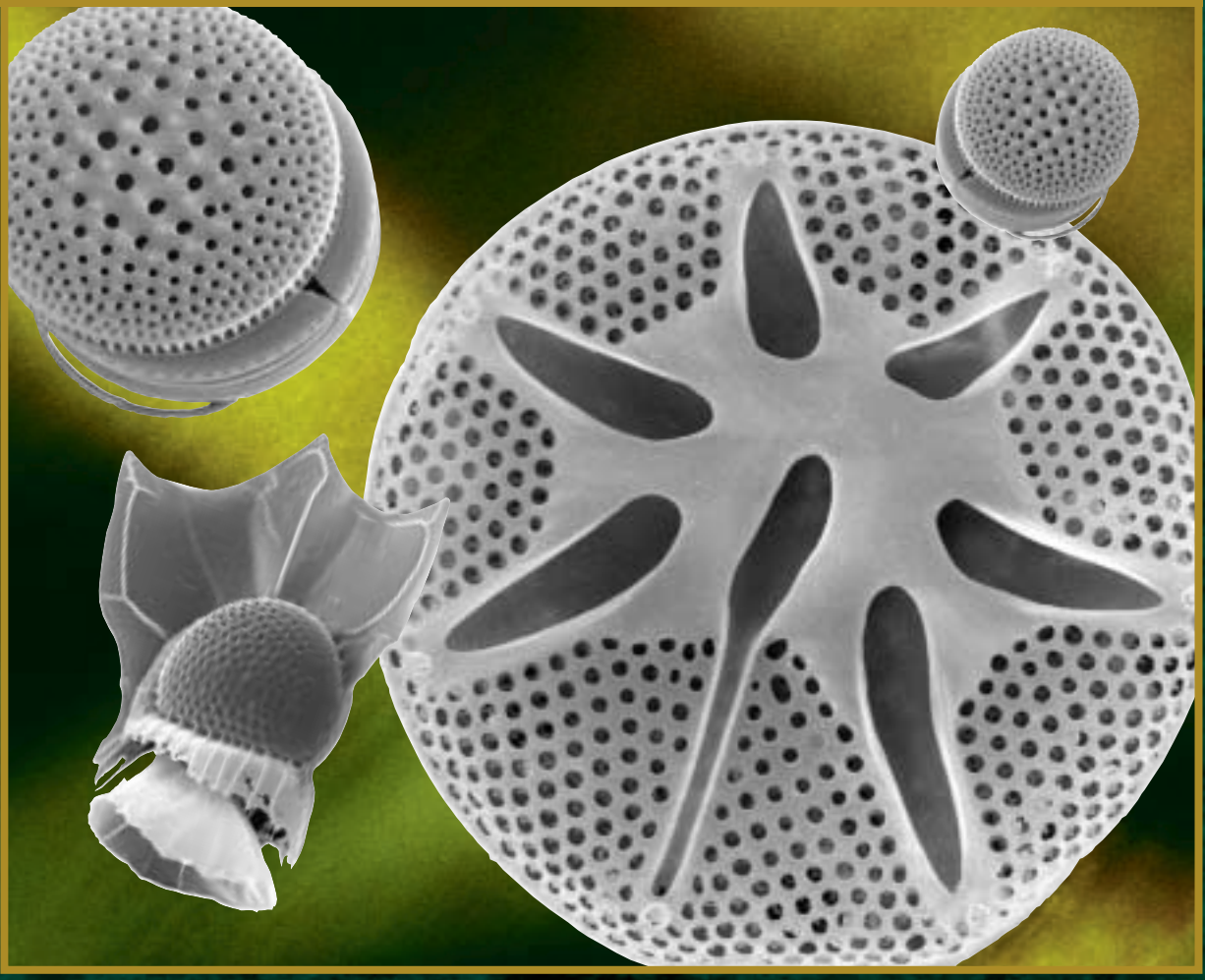
THE ECOLOGICAL SOCIETY OF AMERICA

Frontiers in Ecology and the Environment

Issue No 9 Volume 2 November 2004



From genes to ocean ecosystems

Using GPS effectively in field studies

Forum: managing the wild

esa

From genes to ecosystems: the ocean's new frontier

Scott C Doney[1], Mark R Abbott[2], John J Cullen[3], David M Karl[4], and Lewis Rothstein[5]

The application of new molecular and genomic techniques to the ocean is driving a scientific revolution in marine microbiology. Discoveries range from previously unknown groups of organisms and novel metabolic pathways to a deeper appreciation of the fundamental genetic and functional diversity of oceanic microbes. The "oceanic genotype" represents only the potential biological capacity and sets an upper constraint on possible pathways and ecosystem rates. The realized structure and functioning of marine ecosystems, the "oceanic phenotype", reflects the complex interactions of individuals and populations with their physical and chemical environment and with each other. A comprehensive exploitation of the wealth of new genomic data therefore requires a close synergy with interdisciplinary ocean research. Incorporating the information from environmental genomics, targeted process studies, and ocean observing systems into numerical models will improve predictions of the ocean's response to environmental perturbations. Integrating information from genes, populations, and ecosystems is the next great challenge for oceanography.

Front Ecol Environ 2004; 2(9): 457–466

Molecular and genomic techniques are fundamentally changing our view of life in the ocean. This is particularly true for marine microbes, organisms of less than ~100–150μm that include viruses and prokaryotes as well as single- and multi-celled eukaryotes (Figure 1; Stahl and Tiedje 2002). Photosynthetic microbes synthesize most of the organic matter in the sea, heterotrophic microbes degrade it, and microbial activity regulates the ocean redox state, nutrient cycling, and trace gases relevant to global climate, for example CO_2, dimethylsulfide (DMS), and nitrous oxide (N_2O). Thus, the ecology and biogeochemistry of the sea is governed at a basic level by the activity of microbes.

In a nutshell:

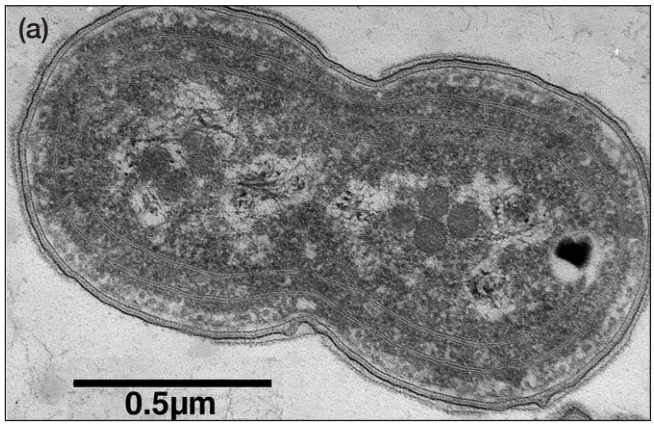
- Ocean microbes play critical roles in ecosystem dynamics, biogeochemical cycles, and climate systems
- Marine environmental genomics provides a wealth of new information on the functional diversity, physiology, and ecological potential of microbes
- Interpretation of the genomic data requires a rich oceanographic context from ocean observing systems, process studies, and numerical models
- The integration of marine genomics, population and community ecology, and biogeochemistry will lead to a new generation of more realistic oceanographic simulations, including improved climate change projections

[1]*Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, MA (sdoney@whoi.edu);* [2]*College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR;* [3]*Center for Marine Environmental Prediction, Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada;* [4]*School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, HI;* [5]*Graduate School of Oceanography, University of Rhode Island, Narragansett, RI*

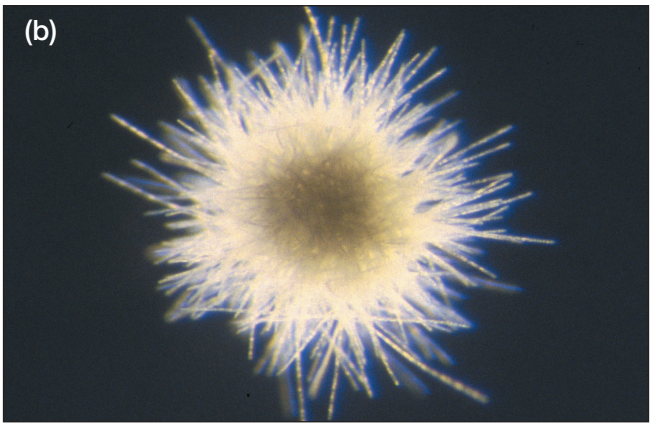
Microbial environmental genomics is a rapidly expanding field that uses nucleotide- and nucleic acid-based analyses of pure cultures and mixed natural microbial assemblages to investigate the diversity and function of microorganisms (Table 1). As is common for a new field in an exploratory phase, discoveries are occurring at a fast pace, challenging traditional paradigms and conceptual models. Examples from the ocean include a previously unknown group of unicellular prokaryotes that can fix nitrogen, a key biogeochemical process (Zehr *et al.* 2001), and the widespread presence of a novel and unexpected metabolic pathway, proteorhodopsin-based phototrophy (Béjà *et al.* 2000, 2001). Ocean scientists are also exploiting a vast and still growing gene sequencing capacity, a by-product of the Human Genome Project, and full genomes of key marine microbes are appearing at an unprecedented rate (Rocap *et al.* 2003; Palenik *et al.* 2003). Novel techniques are now available for sequencing the "meta-genome" of entire naturally occurring microbial communities (Venter *et al.* 2004).

Some general and intriguing trends are emerging. Individual species appear to have both unexpected physiological capabilities and limitations; in many cases, classically identified microbial "species" may be misnomers, lumping together genetically divergent organisms that occupy different oceanic niches and habitats. Furthermore, we now recognize that the smallest oceanic microbes – formerly thought to be heterotrophic bacteria reliant on pre-formed organic matter – include *Prochlorococcus* and *Synechococcus* (photoautotrophic picoplankton; Chisholm *et al.* 1988), the two most abundant "plants" on Earth. Members of the *Archaea* domain (Delong 1992; Fuhrman *et al.* 1992) make up a substantial fraction of mid- and deep-water microbes (Karner *et al.* 2001) and may be chemolithoautotrophic

© The Ecological Society of America



Courtesy of J Waterbury and E Webb, Woods Hole Oceanographic Institution



Courtesy of I Kaczmarska, Mount Allison University

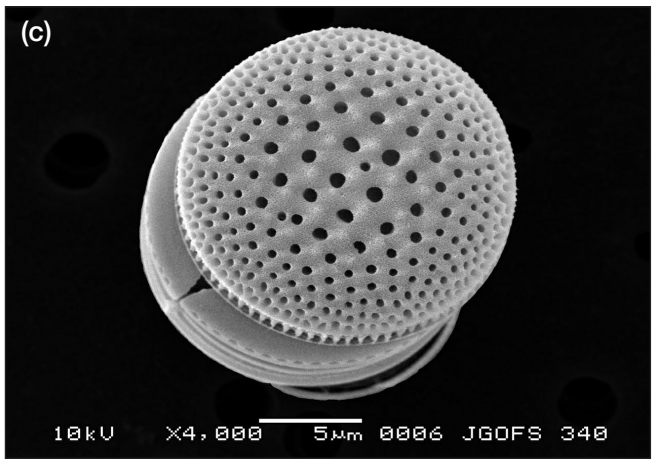


Figure 1. *Oceanic photosynthetic microorganisms: (a) prokaryotic* Synechococcus *cell, a key oceanic picoplankton species, especially in nutrient poor subtropical gyres, and one of the most abundant organisms on the planet; (b) colony of cyanobacteria* Trichodesmium *(scale of image ~4mm), a nitrogen fixing species common in warm, well-stratified subtropical environments; (c) eukaryotic, open-ocean centric diatom* Thalassiosira*, an organism that forms silica shells and a contributor to the vertical export of organic carbon from the surface ocean.*

(Table 2). The genomic data helping to elucidate key biogeochemical cycles also indicate that microorganisms are often able to conduct only a single specific step in a pathway. The overall transformations therefore require close coordination of microbial assemblages or consortia.

Many of the microbes in the ocean have not been cultured and their physiology and ecological roles remain somewhat ill defined. Yet the complexity of the ocean genome suggests that its ecosystems are even more diverse than previously thought, with a deeper genetic reservoir, or potential. And the expression of the ocean genome involves an intricate dynamic of biological responses to environmental forcing and interspecies interactions that is yet to be resolved. This raises troublesome issues about our ability to predict the future behavior of such a system under, for example, climate warming.

Rather than compromising existing approaches to describing ocean ecology, the new genomic data will serve as a basis for developing and testing specific hypotheses. Questions of particular relevance involve the interplay between microbial genomics, environmental variability, community structure, and biogeochemical

Table 1. Selected nucleotide- and nucleic acid-based methods, arranged by date of first use, present in the marine microbial ecologist's toolbox (see Cooksey 1998; Paul 2001; DeLong 2004)

Period	*Indices/parameters*	*Ecological information*
1960s	• ATP and DNA concentrations	• Total microbial biomass
1970s	• Nucleic acid staining – epifluorescence microscopy	• Total cell enumeration
	• Incorporation of radioisotopic labeled precursors into RNA and DNA	• Total nucleic acid production, growth rates
1980s	• rRNA sequence analysis	• Phylogeny and biodiversity
	• FISH with rRNA-targeted probes	• Enumeration of specific microbial groups in mixed assemblages
1990s	• PCR for cultivation-independent amplification of target genes	• Detection of specific microbial groups or metabolic function in mixed assemblages
	• q-PCR of target genes or mRNA copy number	• Quantitative assessment of specific organisms, gene content, or expression in mixed assemblages
	• BAC library construction	• Cultivation-independent detection of gene content and arrangement in mixed assemblages
	• Whole genome sequencing	• Determination of complete genotype
2000s	• DNA microarrays (chips) of multiple (up to 10^4) probes on solid surface for simultaneous hybridization of target genes	• Rapid and comprehensive characterization of community structure, gene expression, and ecosystem function

Abbreviations: ATP = adenosine-5'-triphosphate, DNA = deoxyribonucleic acid, RNA = ribonucleic acid, FISH = fluorescence in situ hybridization, mRNA = messenger RNA, PCR = polymerase chain reaction, rRNA = ribosomal RNA PCR, q-PCR = quantitative PCR, BAC = bacterial artificial chromosome.

© The Ecological Society of America

cycling. Addressing these issues will be both a critical challenge and an excellent opportunity for oceanography over the next decade. However, fully exploiting the wealth of new genomic data will require a close synergy with rapidly developing interdisciplinary ocean research approaches that provide an overall framework. Here we emphasize the melding of genomic, process study, and observing system data with a new generation of numerical ecological models.

Table 2. Microbial metabolic processes

Term	*Energy source*	*e⁻ Donor*	*C source*
Photolithoautotroph	Light	H_2O, H_2S, H_2	CO_2
Photolithoheterotroph	Light	H_2O, H_2S, H_2	Org-C
Photoorganoautotroph	Light	Org-C	CO_2
Photoorganoheterotroph	Light	Org-C	Org-C
Chemolithoautotroph	Chemical	H_2S, $S_2O_3^{-2}$, NH_4^+/NO_2^-, H_2, red-Fe/Mn	CO_2
Chemolithoheterotroph	Chemical	H_2S, $S_2O_2^{-2}$, NH_4^+/NO_2^-, H_2, red-Fe/Mn	Org-C
Chemoorganoautotroph	Chemical	Org-C	CO_2
Chemoorganoheterotroph	Chemical	Org-C	Org-C
Mixotroph[1]	Light/Chemical	Red inorganic/ Org-C	CO_2/Org-C

[1]Multiple possible use patterns of mixed energy sources, e⁻ donors and C-sources

Oceanic prokaryotic microorganisms use a diverse spectrum of metabolic processes to derive the energy and organic C required to support life, ranging from "pure" oxygenic photosynthesis (photolithoautotrophy) to classic heterotrophy (chemoorganoheterotrophy); recent work suggests that some of the intermediate pathways, such as photoorganoautotrophy and photoorganoheterotrophy, as well as mixotrophy, may play a much larger role in ocean ecology than previously thought

The rise of marine environmental genomics

High throughput sequencing of environmental DNA and RNA allows us to test directly many concepts about microbial ecology and the underlying mechanisms of biogeochemical processes. The use of molecular tools to study ribosomal RNA (rRNA) sequences of uncultured microbes in the late 1980s and early 1990s unveiled a huge diversity of prokaryotes in ocean ecosystems (DeLong *et al.* 1989; Giovannoni *et al.* 1990; Pace 1996, 1997). Using bacterial artificial chromosomes (BACs), DeLong and colleagues went on to create and sequence a library of relatively large DNA fragments that included the genes coding for the rRNA from γ-proteobacteria. In an example of the serendipity involved in environmental genomics, linked to one of the rRNA genes was a proteorhodopsin gene, later shown to be a light-driven proton pump, an energy system previously unknown within the bacterial domain (Béjà *et al.* 2000, 2001). If these organisms do indeed use proteorhodopsin for photoorganoheterotrophy (ie a metabolic process that gains energy from the sun and carbon from organic matter), their presence fundamentally alters our perception of how carbon is processed within the surface waters of the ocean (Karl 2002a; Table 2). Effectively, the subsidy of light energy supplements the more traditional heterotrophic lifestyle, resulting in a more efficient and possibly more competitive microbe (especially under conditions of organic C limitation).

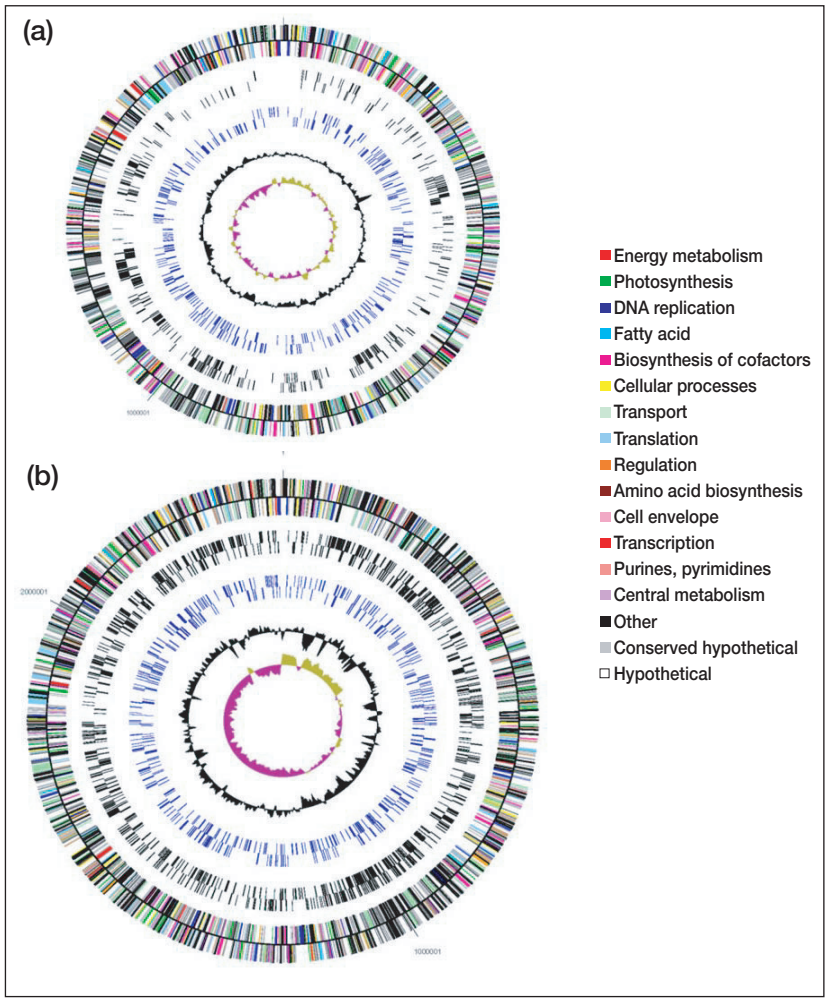
The proteorhodopsin discovery is a sobering reminder of how little we actually know about the fundamental oceanic genotype. This situation is improving. Several marine phytoplankton genomes have been sequenced by the US Department of Energy Joint Genome Institute http://www.jgi.doe.gov/) and the French GENOSCOPE project (http://www.genoscope.cns.fr/) and are available to the scientific community, including *Synechococcus* and *Trichodesmium*, several strains of *Prochlorococcus*, and a diatom, *Thalassiosira pseudonana*. Plans are underway to produce genomes for other important marine organisms, including the coccolithophore *Emiliania huxleyii* and the unicellular nitrogen fixer *Crocosphaera*. A recent initiative funded by the Gordon and Betty Moore Foundation may soon sequence another 100–150 selected marine microbial organisms (see Marine Microbiology Initiative, http://www.moore.org/).

Even this limited number of available genomic sequences is offering up surprises. Two distinct ecotypes of *Prochlorococcus*, the most abundant photosynthetic organism in the ocean, differ dramatically in their genetic material, having in common only about one half of the total number of genes (Figure 2; Rocap *et al.* 2003). Corresponding metabolic differences (eg in light–growth relationships and nitrogen-assimilation capacity) are striking. They support the niche differentiation of the two strains, one that grows in sunlit surface waters and the other deeper in the water-column, in a subsurface chlorophyll maximum often found near 100m depth in the subtropics. Equally remarkable, neither ecotype appears able to grow directly on nitrate, the primary fuel of new production (and carbon export) in the sea, according to the current paradigm (Dugdale and Goering 1967). These results contribute to a more complex view of oceanic new production that also includes substantial inputs from in situ nitrogen fixation (Karl *et al.* 2001).

Future genomic research directions

On a somewhat different tack, researchers from the Institute for Biological Energy Alternatives (http://www.bioenergyalts.org/) have embarked on the ambitious project of sequencing the entire marine microbial genome. Using a "shotgun" sequencing technique, this group conducted a pilot study of genomic diversity in the upper ocean microbial community from the Sargasso Sea (Venter *et al.* 2004). The results are stunning; sampling about 2000 liters of seawater, they report the discovery of more than one million new genes, about 70 000 of which

© The Ecological Society of America



Courtesy of the © Nature Publishing Group

Figure 2. *Circular representation of the* Prochlorococcus *genomes (Rocap* et al. *2003) for two ecotypes, (a) MED4 and (b) MIT 9313. While grouped into the same species, the two organisms contain strikingly different genetic material, which may reflect adaptation to different environmental habitats. For both genomes, the outermost circles (1 and 2) display the predicted protein coding regions on the plus and minus strands, respectively (color coding to the right shows putative metabolic/regulatory role). The next two circles indicate genes not present in the other* Prochlorococcus *genome on the plus (circle 3) and minus (circle 4) strands. (Adapted from Supp. Figures 1 and 2 from Rocap* et al. *2003.)*

are characterized as "novel" and include a wide range of potential biogeochemical pathways. They also report the discovery of more than a thousand new gene-based "species". Falkowski and de Vargas (2004) highlight the importance of this work for marine microbial ecology and evolutionary biology, emphasizing the merits of taking the Venter *et al.* (2004) strategy into a global oceanographic context.

A critical facet of environmental genomics is that sequencing DNA, either in a culture or field sample, is just the beginning. The next step involves identifying actual genes and determining the function of the coded protein. This is often done by comparison with other known genes, although in many cases there are no direct analogues in the databases, which are populated predominantly by terrestrial and human pathogenic microorganisms. Sorting through the tens to hundreds of millions of nucleotide arrangements from organism and environmental shotgun sequencing is a major computational task, as is the science – or art – of deciphering the resulting information. The next step is to assess the expression and regulation mechanisms of this genomic potential by sampling messenger RNA (mRNA) in cells across a range of environmental conditions. Finally, proteomic analyses (the study of expressed gene products) are needed to quantify the activity of the resulting proteins and enzymes *in situ*; only then can we directly connect the genomics to the biological rates that govern biogeochemistry. Progress is greatly hindered by our inability to grow many of the target microbe species in culture, and by other less tangible issues, including the lack of a larger number of scientists with the required cross-disciplinary training.

The rapidly growing database of marine microbial DNA sequences allows the application of cutting-edge technology from molecular biology, such as DNA micro- and macroarrays, to oceanography (Gibson 2002). The arrays consist of up to 10 000 unique DNA probes robotically spotted onto a small glass slide. The probes could correspond, for example, to all of the genes from a single organism or to different variants of a single enzyme across a wide range of microbes. Genetic material from an environmental sample will hybridize with specific probes on the array only if there is a close sequence match; this provides a tool for simultaneously determining the presence or absence of an entire suite of genes (using DNA) or the expression of those genes (using mRNA). Currently, several groups are developing whole or partial genome arrays for marine organisms and geochemically important functional genes such as those involved in nitrogen fixation, carbon uptake, metal uptake, or other metabolic processes.

Ocean physics, nutrients, and ecosystem dynamics

The past two decades have also seen a virtual explosion in open-ocean biological and biogeochemical research (Fasham *et al.* 2003). New insights provide a consistent mechanistic picture of the large-scale controls for surface ocean biomass, productivity, and community structure, as well as an essential framework for interpreting and understanding emerging genomic data. The basic picture is

 © The Ecological Society of America

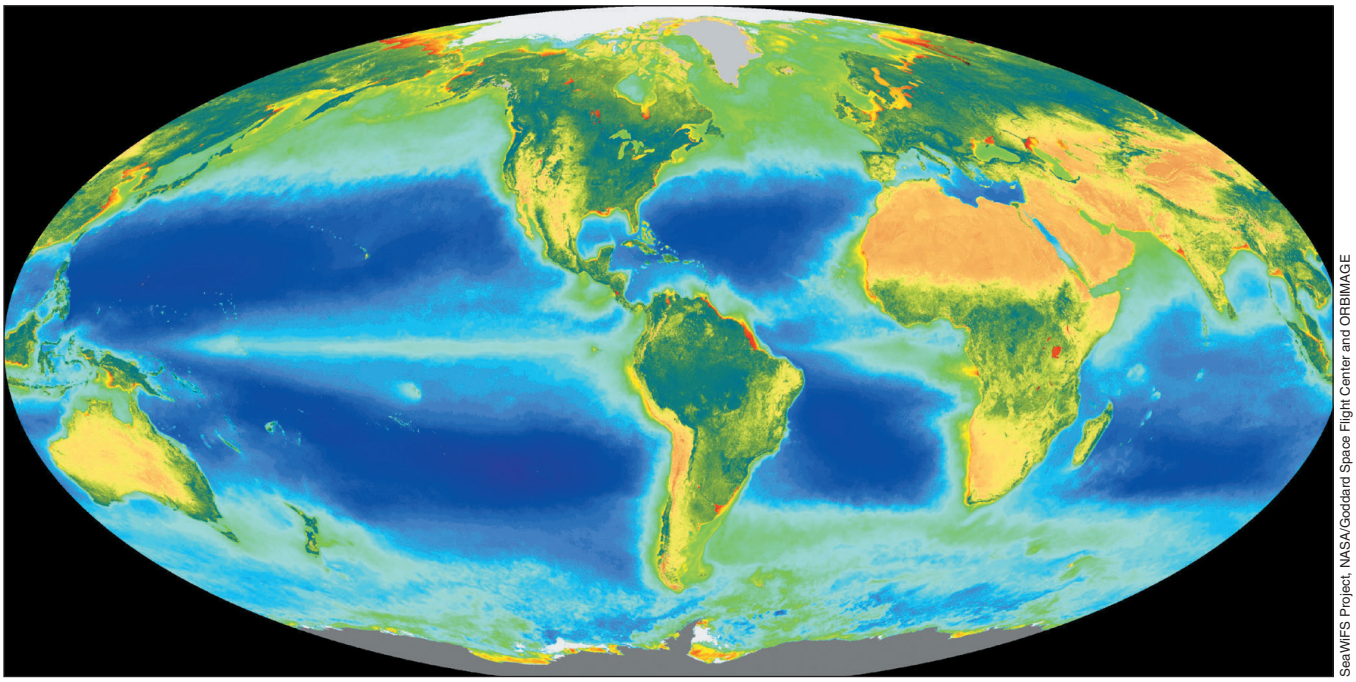


Figure 3. *Global map of annual average surface ocean chlorophyll, a measure of photosynthetic (autotrophic) biomass, derived from the SeaWiFS satellite ocean color sensor. The satellite data clearly illustrate the large-scale spatial patterns of ocean biomes driven by ocean mixing, light limitation, subsurface nutrient and iron fluxes, and atmospheric iron inputs. The ocean color scale is approximately logarithmic, with more than two orders of magnitude of change from the low biomass/low nutrient subtropical gyres (blue) to coastal upwelling regimes (yellow/orange).*

similar to that outlined more than half a century ago (eg Sverdrup *et al.* 1942), but with substantial refinement.

The regional and seasonal distribution of phytoplankton (Figure 3) is determined primarily by the supply of subsurface nutrients (eg nitrate, phosphate, silicate) and the depth of turbulent vertical mixing, which alters the average light field seen by cells. Supported by upwelling and deep seasonal convection, phytoplankton blooms occur in the temperate and subpolar latitudes and in eastern boundary coastal upwelling zones. These ecosystems are dominated by larger eukaryotic autotrophic cells such as diatoms (Margalef 1978), mesozooplankton grazers (eg copepods), and high organic matter export (30–50% of gross primary production) (Legendre and LeFevre 1989; Kiørboe 1993). Export is important because it governs the large-scale biogeochemistry of the ocean and the net sequestration of CO_2 in the ocean interior away from the atmosphere for decades to centuries.

The well-stratified subtropical oceans are characterized by downwelling, low biomass, and generally low nutrient concentrations, at most a few nanomoles (10^{-9} moles) of inorganic nitrogen and/or phosphorus per liter in surface waters. Despite this condition, primary productivity is relatively high, driven by a tight coupling between production by small prokaryotic autotrophs such as *Prochlorococcus* and *Synechococcus*, grazing by microzooplankton, and microbial consumption and respiration of detrital organic matter. These ecosystems tend to export only a small fraction (about 5%) of primary productivity as sinking particles or via the downward mixing of dissolved organic matter. Specialized organisms such as the cyanobacteria *Trichodesmium*, *Richelia intracellularis* (a symbiont of *Hemialus* and *Rhizosolenia* diatoms), and newly discovered unicellular prokaryotes contribute to high nitrogen fixation rates, elevated productivity, and export in some subtropical areas (Karl 2002b). Nitrogen fixation is an iron-intensive enzymatic process and thus may be tied to atmospheric inputs through dust deposition from desert source regions.

The Southern Ocean, equatorial Pacific, and eastern sub-polar North Pacific are anomalies; they have intermediate biomass and productivity levels and high levels of unused surface inorganic nitrogen and phosphorus. A critical discovery is that these areas are iron limited (Martin 1990; Boyd *et al.* 2000). Ocean productivity and community structure therefore reflect a variety of factors that vary from region to region – physical supply of subsurface macronutrients and iron, atmospheric iron inputs, and the regeneration and partitioning of iron between pools, which may be much more complicated than for nitrogen or phosphorus (Fung *et al.* 2000). The distinct major ocean biomes should be targeted for comparative microbial genomic analysis.

Towards better marine ecosystem models

Marine ecological modeling has a long history, dating back to the 1940s (Steele 1974), but has undergone an exciting renaissance over the past decade (Doney 1999). Numerical models are essentially tools for scaling the interactions of specific processes to the level of a whole ecosystem, or for predicting how an ecosystem will evolve

© The Ecological Society of America

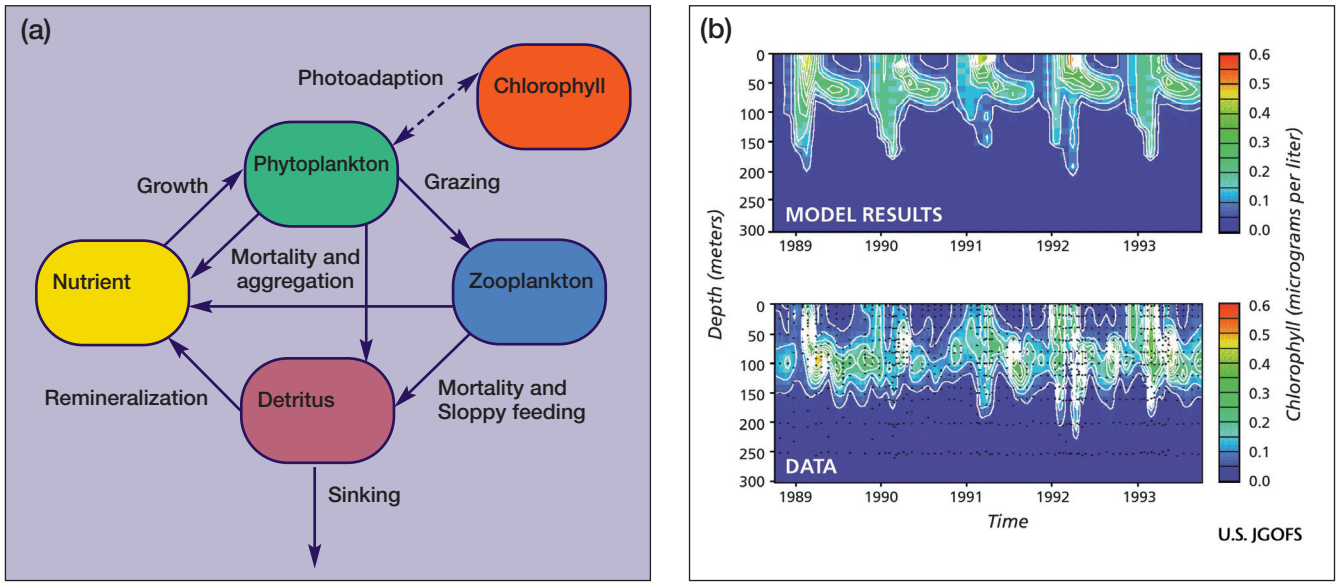


Figure 4. *(a) Schematic of a simple marine ecosystem model with separate compartments for nutrients, phytoplankton, zooplankton, detritus (collectively NPZD), and chlorophyll. (b) Depth-versus-time comparison of observed and simulated chlorophyll concentrations from a 1-D version of the model applied to a multi-year record from the Bermuda Atlantic Time-series Study (BATS) site in the Sargasso Sea. Deep convection in the winter brings nutrients to the surface, generating a winter/spring phytoplankton bloom. During the summer, biological export drives surface nutrients and chlorophyll to very low levels. A subsurface, deep chlorophyll maximum forms where light from above and nutrients from below are both available.*

in the presence of natural and human perturbations. The basic structure of many of the models consists of a simple conceptual food web with separate compartments for nutrients, phytoplankton, zooplankton, and detritus (collectively termed NPZD) (Figure 4; Fasham *et al.* 1990). Models in this class track the water-column concentrations and flows of nitrogen, phosphorus, or other elements through the trophic level compartments, but typically do not resolve individual organisms. Model functions for grazing and the limitation of primary productivity by nutrients and light are based on results from laboratory experiments and field bottle incubations.

Many efforts are underway to enhance the realism and sophistication of the classical NPZD framework. Modifications include the specific treatment of heterotrophic bacteria, the addition of multiple limiting nutrients (eg N, P, Si) and trace metals (eg iron), and the incorporation of multiple subgroups of different size phytoplankton, zooplankton, and detritus. Particle size can vary by several orders of magnitude in the ocean and is critical because it governs vertical sinking velocities. Two different approaches are being pursued to include so-called size-structure in models: desegregation of the boxes into distinct size classes (Moloney and Field 1991), and continuous size distributions and allometric relationships that mimic the shift to larger cells and detritus particles under higher biomass and/or productivity levels (Armstrong 2003). A related trend is the simulation of distinct "geochemical functional groups" that alter ocean biogeochemical distributions in unique ways (Moore *et al.* 2002). Specific examples are diatoms that produce biogenic silica, calcifiers such as coccolithophores (Iglesias-Rodriguez *et al.* 2002), and nitrogen-fixing bacteria.

At the other extreme, individual-based models are growing more computationally feasible, allowing researchers to quantify the interactions of thousands to millions of discrete organisms. Numerical methods include both microfluidic (Siegel 1998) and cellular automata approaches. Although current simulations can only be conducted over relatively small time and space scales (days and centimeters), the results can be used to provide a mechanistic basis for developing model functions of biological processes (eg grazing) for use in larger scale models (akin to turbulence modeling in physical oceanography).

All of these modern numerical models have a number of limitations. Many of the key processes in the simulations are difficult (if not impossible) to directly measure. Examples include phytoplankton mortality rates (which can come from viral infection and, perhaps, programmed cell death) and the ultimate controls on the abundance of higher trophic levels. Model parameters for these processes are often adjusted so that overall model performance is improved, but in the end this is a rather unsatisfying exercise. Second, the models aggregate a tremendous amount of ecological information and complexity into a small number of bulk compartments. Although the recent move to a dozen or so compartments is a far cry from representation of the full diversity of the ocean, even that level of complexity often exceeds our ability to constrain model parameters or validate model behavior, given our current observation base (Denman 2003). There are, to put it simply, too many "free parameters", and a reasonable result can be obtained for the wrong reasons.

© The Ecological Society of America

Ideally, the predictive skill of models is assessed by comparisons against oceanographic data, such as time series, process studies, and regional distributions of the state variables. More advanced techniques attempt to improve upon the predictive skill of the "forward model" equations by formally incorporating select observations within the model equations. Data assimilation, a numerical approach adapted from weather prediction, creates a hybrid model/observational dataset that is consistent (within some error bounds) with both the observations and a set of ecological/physical model equations. However, ocean biological data are often sparse, noisy, and incomplete, making model evaluation difficult.

Despite many deficiencies, current state of the art models do a surprisingly credible job of capturing the seasonal cycle and large-scale spatial patterns for bulk ecological measures (eg surface chlorophyll, primary productivity, nutrients and inorganic carbon, export flux, and so on). Comparisons of models and measurements demonstrate the importance, for example, of including phytoplankton iron limitation and an active iron cycle (Moore *et al.* 2002). In turn, model predictions for the time/space variations of different plankton groups are broadly consistent with the available, but admittedly rather limited, taxonomic field data. Historical data document significant shifts in ecosystem rates and function on interannual to decadal scales (Karl *et al.* 2001), offering a window on the ecological responses to climate perturbations; some simulations match qualitatively observed biological variability driven by climate modes such as the El Niño–Southern Oscillation. Although these tests of model skill are necessary, they are not sufficient, especially as the planet moves into new climate regimes.

Bridging genomics and numerical ecosystem simulations

How do we reconcile ecosystem modeling with the wealth of new environmental genomic data? The current generation of NPZD and expanded functional group methods can probably only take us so far. For the most part, current model structures do not allow for direct, straightforward comparisons with genomic data, except at the most basic level of aggregated biogeochemical rates (eg primary production, nitrogen fixation, or calcification) and the presence or absence of major taxonomic groups. Efforts are underway to relate simulated phytoplankton nutrient limitation patterns to the expression of specific genes associated with nutrient stress (eg active transport proteins, or enzymes for alternative metabolic pathways). One can envision similar approaches for other particular metabolic processes.

But few, if any, of the current models represent cellular biochemistry at the required level of detail, and they do not incorporate the variety of recently discovered metabolic pathways. In fact, many aspects of environmental genomic information have no direct analogues within existing models, for example the presence of unexpressed genes. Many of the most exciting genomic findings are linked to the enormous genetic diversity in the sea. This diversity captured within even the most sophisticated models will probably remain, for the foreseeable future, many orders of magnitude lower than that of real ocean ecosystems.

New conceptual models and paradigms will be required. One promising avenue is to back away from the goal of predicting the temporal evolution of species or compartments and simply to simulate the key ecological functions that are present within the system and their responses to internal and external perturbations. Such an approach would be based on determining the environmental factors that regulate specific functions, analogous to genetic regulation; that is, models would start with a simulated ocean genotype, one much richer than presently resolved, and then predict, under a range of conditions, the expression of a simulated phenotype, which could then be compared with genomic data. Climate variability and climate change provide an interesting challenge in this regard, because of their scale and persistence: how resilient will ecosystems be, and how will they adapt, to such large-scale environmental changes?

A genotype–phenotype model is particularly attractive when studying certain groups such as heterotropic bacteria, where bulk measures such as standing stock are less useful because only a subset of the organisms may be active at any particular time, or for minor niche species, whose ecological importance is based on their biogeochemical function (eg nitrogen fixation or trace gas production). Rules-based metabolic models being developed following concepts of non-equilibrium thermodynamics (Vallino 2003) offer a starting point that can be augmented by emerging theoretical and empirical arguments for how evolutionary and ecological pressures shape ecosystems. But designing credible models for pelagic ecosystems and then coupling those into ocean circulation simulations present some major challenges.

From modeling and large-scale ecosystem perspectives, how much it matters which specific organisms, or assemblages of organisms, are carrying out the biogeochemical transformations remains unresolved. Put another way, the ocean phenotype almost certainly depends to some extent on how the genotype is dispersed among different species and individuals within species. The genotype–phenotype approach to ecological modeling will therefore have to include, either implicitly or explicitly, treatment for the ecological effects of competition and interactions between individual organisms. This provides a potentially important bridge between genomic-based modeling and high-resolution, individual-based models.

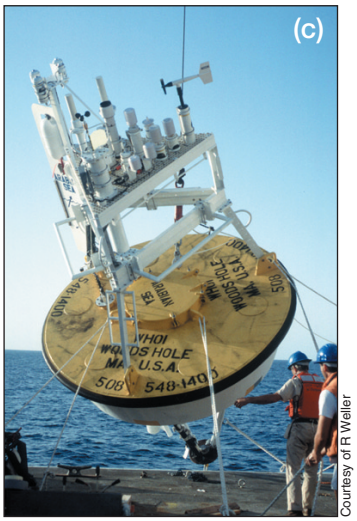
Genomics, oceanography, and climate change

Oceanic biology is an important component of the global climate system, but many of the feedbacks between marine biogeochemistry and climate are only poorly understood. The inorganic carbon inventory of the sea is about fifty

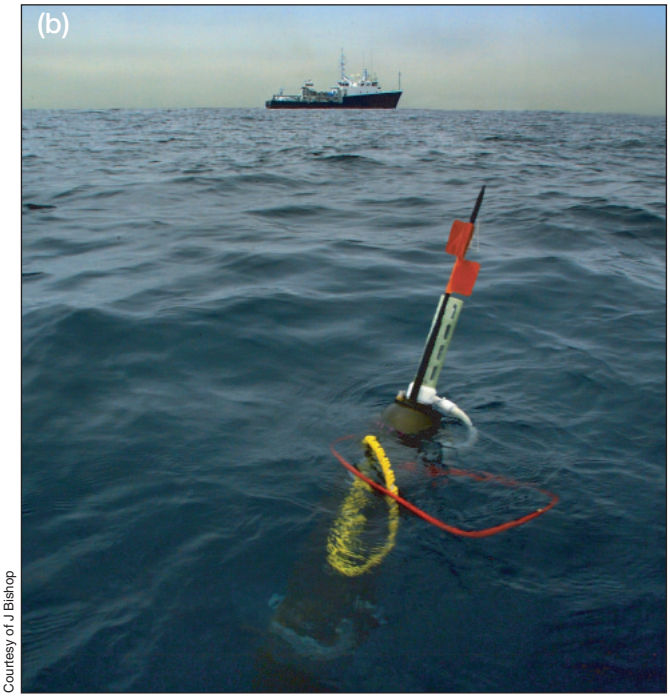
© The Ecological Society of America



Courtesy of US JGOFS Office and K Brink



Courtesy of R Weller



Courtesy of J Bishop

Figure 5. *Elements of an interdisciplinary ocean observing system include: (a) commercial and research ship-based surveys and time-series; (b) moorings and coastal and regional cabled networks (eg ORION http://www.orion.org); and (c) autonomous drifters, floats and gliders, and aircraft and satellite remote sensing (see Figure 4; Doney* et al. *2004).*

times larger than the amount of atmospheric CO_2, a critical greenhouse gas. Ice-core records reflect large variations in atmospheric CO_2 concentrations (about 25%) over glacial–interglacial cycles, and the ocean is the most probable culprit for such oscillations (Sigman and Boyle 2000). At present, the ocean absorbs about a third of the CO_2 released by fossil fuel combustion, primarily through physical circulation, thereby slowing potential greenhouse warming (Sarmiento and Gruber 2002). There are also significant (biological) marine sources of N_2O, also a greenhouse gas and a key compound in stratospheric ozone chemistry, methane (CH_4), and DMS, which can alter cloud and radiation feedbacks (Charlson *et al.* 1987). Climate model projections suggest that the surface ocean will warm and therefore become more stratified over the next several centuries, due to human-induced climate change. Preliminary ecosystem model simulations predict large changes in regional productivity and marine community structure (Boyd and Doney 2002). However, the impacts on ocean biogeochemistry and carbon uptake, not to mention fisheries and ecosystem health, are not yet clear.

At some level, the oceanic genome encapsulates a complete blueprint of potential biological function, and should help to provide answers for such questions as how marine biogeochemistry will respond to global warming. The actual ecological dynamics, however, are determined by the interactions of these genomic capabilities with substrate availability, environmental cues and constraints, other organisms, and physical processes such as sinking and advection. That is, we should be able to work our way up from genes to organisms to ecosystems, using first principles. Because of the daunting complexity of the scaling problem, however, such a reductionist approach is not likely to be the most productive path. Rather, a more holistic, multi-pronged attack on the problem is called for, one where the genomic research complements and leverages efforts on a range of scales.

In particular, we need to continue and expand measurements of biomass, chemical tracers, and *in situ* biological rates, as these are the state variables which describe ecological dynamics. The ocean biogeochemical community has recently completed work on just such a research strategy, *Ocean Carbon and Climate Change* (Doney *et al.* 2004), which is a natural and welcome home for future ocean genomic work. Traditional shipboard sampling will be augmented by satellite remote sensing, which now routinely monitors the biological and physical state of the surface ocean (McClain *et al.* 1998), and in situ sensors, which can measure a wide range of environmental properties in near real time from fixed moorings, submarine cable networks, and

© The Ecological Society of America

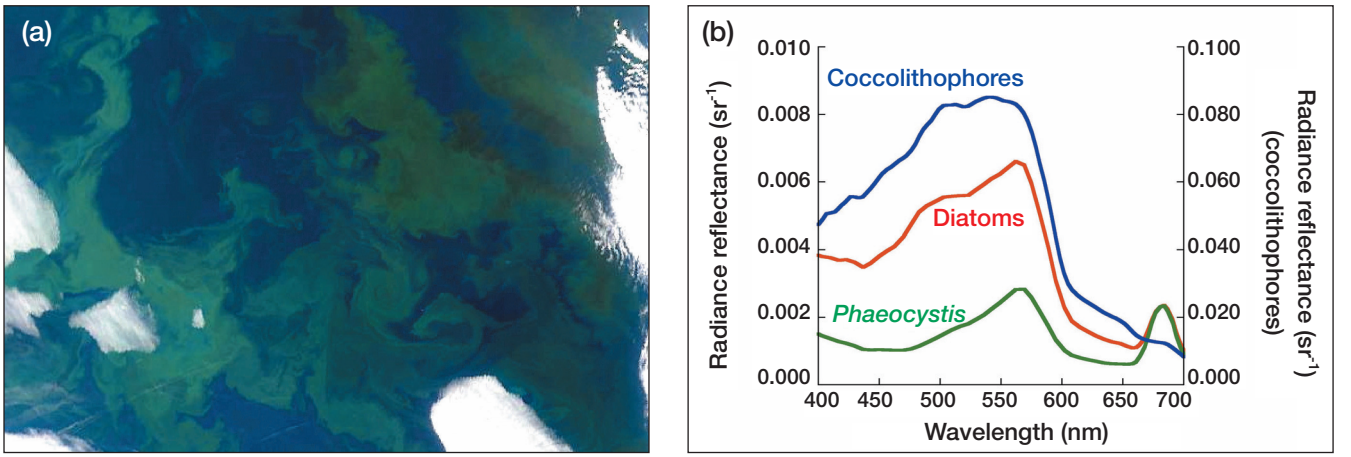


Figure 6. *Under some conditions, measurements of ocean color can reveal much more than the concentration of chlorophyll. (a) True color satellite image of a phytoplankton bloom in the Bering Sea on June 7, 2001 (292 km x 200 km centered near 58.7°N, 177°W). During this period, shipboard sampling indicated blooms dominated by diatoms and the prymnesiophyte* Phaeocystis *in close proximity, probably corresponding to the lighter and darker green features in the image. Coccolithophore blooms are highly reflective and may be responsible for the brighter features in the SW corner of the image. (b) In-water measurements of hyperspectral ocean color (reflectance: the ratio of upwelling radiance to downwelling solar irradiance at the surface) from the Bering Sea reveal striking differences between blooms (note the scale change for the coccolithophores); not only does the brightness of the water (average reflectance) vary significantly, shapes of the spectra differ because pigmentation, cell size, and the quantum yield of sun-induced chlorophyll fluorescence (the peaks near 680 nm) influence the measurements, providing a key for remote sensing of species composition and perhaps physiological condition of phytoplankton. (Data from JJ Cullen and RF Davis.)*

autonomous platforms (Figure 5; Dickey 2003). Figure 6 highlights a bio-optical example of this interdisciplinary synergy (Bissett *et al.* 2001). Combined with advanced numerical models, these observations provide fundamental, quantitative constraints on the ocean budgets and cycling of biogeochemical tracers and the flow of energy that is essential for interpreting and understanding genomic data.

Venter *et al.* (2004) highlight the value of doing more sequencing work, especially since costs will come down. This is certainly an opportunity that the ocean science community must pursue, and a major field effort is now underway as part of a new research initiative funded by the Gordon and Betty Moore Foundation. However, it is very important to emphasize that the value of sequencing is greatly enhanced when accompanied by thorough oceanographic and biochemical sampling, and modeling to test developing and evolving hypotheses. Only then can the community properly place ocean genomic complexity within the context of larger scale processes.

Acknowledgments

The authors are members of the Partnership for Advancing Interdisciplinary Global Modeling (PARADIGM) project that is supported by the National Ocean Partnership Program through funding provided by the Office of Naval Research and the National Science Foundation (NSF/ONR NOPP grant N000140210370). S Doney would like to thank C Carlson, S Hedrick, B Jenkins, G Rocap, M Saito, D Siegel, D Toole, B Ward, and E Webb for helpful comments, suggestions, and assistance with some of the figures. This is Woods Hole Oceanographic Institution contribution #11191.

References

Armstrong RA. 2003. A hybrid spectral representation of phytoplankton growth and zooplankton response: the "control rod" model of plankton interaction. *Deep-Sea Res II* **50**: 2895–2916.

Béjà O, Aravind L, Koonin EV, *et al.* 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–06.

Béjà O, Spudich EN, Spudich JL, *et al.* 2001. Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786–89.

Bissett WP, Schofield O, Glenn S, *et al.* 2001. Resolving the impacts and feedback of ocean optics on upper ocean ecology. *Oceanography* **14**: 30–53.

Boyd PW and Doney SC. 2002. Modelling regional responses by marine pelagic ecosystems to global climate change. *Geophys Res Lett* **29**: doi:10.1029/2001GL014130.

Boyd PW, Watson AJ, Law CS, *et al.* 2000. A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilisation of waters. *Nature* **407**: 695–702.

Charlson RJ, Lovelock JE, Andrae MO, and Warren SG. 1987. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* **326**: 655–61.

Chisholm SW, Olson RJ, Zettler ER, *et al.* 1988: A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**: 340–43.

Cooksey KE (Ed). 1998. Molecular approaches to the study of the ocean. London, UK: Chapman & Hall.

DeLong EF. 1992. Archaea in coastal marine environments. *Proc Natl Acad Sci USA* **89**: 5685–89.

DeLong EF. 2004. Microbial population genomics and ecology: a new frontier. In: Fraser CM, Read T, and Nelson KE (Eds). Microbial genomes. New Jersey: Humana Press.

DeLong EF, Wickham GS, and Pace NR. 1989. Phylogenetic

© The Ecological Society of America

stains: ribosomal RNA-based probes for the identification of single cells. *Science* **243**: 1360–63.
Denman KL. 2003. Modelling planktonic ecosystems: parameterizing complexity. *Prog Oceanogr* **57**: 429–52.
Dickey T. 2003. Emerging ocean observations for interdisciplinary data assimilation systems. *J Mar Syst* **40-41**: 5–48.
Doney SC. 1999. Major challenges confronting marine biogeochemical modeling. *Global Biogeochem Cycles* **13**: 705–14.
Doney SC, Anderson R, Bishop J, *et al.* 2004. Ocean carbon and climate change (OCCC): an implementation strategy for US ocean carbon cycle science. Boulder, CO: UCAR.
Dugdale RC and Goering JJ. 1967. Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnol Oceanogr* **12**: 196–206.
Falkowski PG and de Vargas C. 2004. Shotgun sequencing in the sea: a blast from the past? *Science* **304**: 58–60.
Fasham MJR (Ed). 2003. Ocean biogeochemistry: the role of the ocean carbon cycle in global change. New York, NY: Springer-Verlag.
Fasham MJR, Ducklow HW, and McKelvie SM. 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. *J Mar Res* **48**: 591–639.
Fuhrman JA, McCallum K, and Davis AA. 1992. Novel major archaeabacterial group from marine plankton. *Nature* **356**: 148–49.
Fung IY, Meyn SK, Tegen I, *et al.* 2000. Iron supply and demand in the upper ocean. *Global Biogeochem Cycles* **14**: 281–95.
Gibson G. 2002. Microarrays in ecology and evolution: a preview. *Mol Ecol* **11**: 17–24.
Giovannoni SJ, Britschgi TB, Moyer CL, and Field KG. 1990. Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60–63.
Iglesias-Rodriguez MD, Brown CW, Doney SC, *et al.* 2002. Representing key phytoplankton functional groups in ocean carbon cycle models: coccolithophores. *Global Biogeochem Cycles* **16**: doi:10.1029/2001GB001454.
Karl DM. 1999. A sea of change: biogeochemical variability in the North Pacific subtropical gyre. *Ecosystems* **2**: 181–214.
Karl DM. 2002a. Microbiological oceanography – hidden in a sea of microbes. *Nature* **415**: 590–91.
Karl DM. 2002b. Nutrient dynamics in the deep blue sea. *Trends Microbiol* **10**: 410–18
Karl DM, Bidigare RR, and Letelier RM. 2001. Long-term changes in plankton community structure and productivity in the North Pacific subtropical gyre: The domain shift hypothesis. *Deep-Sea Res II* **48**: 1449–70.
Karner MB, DeLong EF, and Karl DM. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507–10.
Kiørboe T. 1993. Turbulence, phytoplankton cell size, and the structure of pelagic food webs. *Adv Mar Biol* **29**: 1–72.
Legendre L and Le Fèvre J. 1989. Hydrodynamic singularities as controls of recycled versus export production in oceans. In: Berger WH, Smetacek VS, and Wefer G (Eds). Productivity of the ocean: present and past. New York, NY: John Wiley & Sons.
Margalef R. 1978. Life forms of phytoplankton as survival alternatives in an unstable environment. *Oceanol Acta* **1**: 493–509.
Martin JH. 1990. Glacial–interglacial CO_2 change, the iron hypothesis. *Paleoceanography* **5**: 1–13.
McClain CR, Cleave ML, Feldman GC, *et al.* 1998. Science quality SeaWiFS data for global biosphere research. *Sea Technology* **39**: 10–16.
Moloney CL and Field JG. 1991. The size-based dynamics of plankton food webs. I. A simulation model of carbon and nitrogen flows. *J Plank Res* **13**: 1003–38.
Moore JK, Doney SC, Kleypas JA, *et al.* 2002. An intermediate complexity marine ecosystem model for the global domain. *Deep-Sea Res II* **49**: 403–62.
Pace NR. 1996. New perspective on the natural microbial world: molecular microbial ecology. *Amer Soc Microbiol News* **62**: 463–70.
Pace NR. 1997. A molecular view of microbial diversity and the biosphere. *Science* **276**: 734–40.
Palenik B, Brahamsha B, Larimer FW, *et al.* 2003. The genome of a motile marine *Synechococcus*. *Nature* **424**: 1037–42.
Paul JH. 2001. Methods in microbiology: Volume 30 – marine microbiology. San Diego, CA: Academic Press.
Rocap G, Larimer F, Lamerdin J, *et al.* 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042–47
Sarmiento JL and Gruber N. 2002. Sinks for anthropogenic carbon. *Phys Today* **55**: 30–36.
Siegel DA. 1998. Resource competition in a discrete environment: why are plankton distributions paradoxical? *Limnol Oceanogr* **43**: 1133–46.
Sigman DM and Boyle EA. 2000 Glacial/interglacial variations in atmospheric carbon dioxide. *Nature* **407**: 859–69.
Stahl DA and Tiedje J, 2002: Microbial genomics and ecology: a crossroads of opportunity. American Academy of Microbiology, colloquia report, http://www.asm.org, 28p.
Steele J. 1974. The structure of marine ecosystems. Cambridge, MA: Harvard University Press.
Sverdrup HU, Johnson MW, and Fleming RH. 1942. The oceans. Englewood Cliffs, NJ: Prentice-Hall.
Vallino JJ. 2003. Modeling microbial consortiums as distributed metabolic networks. *Biol Bull (Woods Hole)* **204**: 174–79.
Venter JC, Remington K, Heidelberg JF, *et al.* 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.
Zehr JP, Waterbury JB, Turner PJ, *et al.* 2001. Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. *Nature* **412**: 635–38.

 © The Ecological Society of America