Biogeosciences Discuss., 10, 815–850, 2013
www.biogeosciences-discuss.net/10/815/2013/
doi:10.5194/bgd-10-815-2013


Biogeosciences Discussions



# Incorporating genomic information and predicting gene expression patterns in a simplified biogeochemical model

**P. Wang[1], A. B. Burd[1], M. A. Moran[1], R. R. Hood[2], V. J. Coles[2], and P. L. Yager[1]**

[1]Department of Marine Science, University of Georgia, Athens, GA, 30602, USA
[2]Center for Environmental Science, University of Maryland, Cambridge, MD, 21613, USA

Received: 14 December 2012 – Accepted: 26 December 2012 – Published: 15 January 2013

Correspondence to: P. Wang (paulwang628@gmail.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.







## Abstract

We present results from a new marine plankton model that combines selective biogeochemical processes with genetic information. The model allows for phytoplankton to adapt to a changing environment by invoking different utilization pathways for acquisition of nutrients (nitrogen and phosphorus) in response to concentration changes. The simulations use simplified environmental conditions represented by a continuously stirred tank reactor, which is populated by 96 different types of phytoplankton that differ in their physiological characteristics and nutrient uptake/metabolism genes. The results show that the simulated phytoplankton community structure is conceptually consistent with observed regional and global phytoplankton biogeography, the genome content from the dominant types of phytoplankton reflects the imposed environmental constraints, and the transcription of the gene clusters is qualitatively simulated according to the environmental changes. The model shows the feasibility of including genomic knowledge into a biogeochemical model and is suited to understanding and predicting changes in marine microbial community structure and function, and to simulating the biological response to rapid environmental changes.

## 1 Introduction

A major goal of marine biogeochemistry is the understanding of the role that phytoplankton populations and their community structure play in oceanic biogeochemical pathways (Hood et al., 2006; Sarmiento and Gruber, 2006). Numerical models have provided significant insights into the relationship between microbial community structure, determined largely by environmental factors such as nutrient availability (Landry and Kirchman, 2002), and biogeochemical processes (Gregg et al., 2003; Le Quéré et al., 2005; Moore et al., 2002). Early models used three compartments to simulate the dynamics of nutrients, phytoplankton and zooplankton (Franks et al., 1986; Steele, 1974), with detritus being added later (Fasham et al., 2006). Other models





BGD
10, 815–850, 2013



subdivided these compartments according to size (Moloney and Field, 1991) or characterized groups of organisms by broad biogeochemical function (Chai et al., 2002). Many current marine ecosystem models incorporate phytoplankton community structure by using distinct functional groups representing phytoplankton that perform specific biogeochemical activities such as $N_2$ fixation or silica utilization (Hood et al., 2006; Le Quéré, 2006; Litchman et al., 2006).

The oceans are presently changing at an unprecedented rate, with increasing temperature (Barnett et al., 2005), changes in pH (Doney et al., 2009), increased stratification and corresponding impacts on nutrient availability (Beman et al., 2011; Sarmiento et al., 2004). There is strong evidence that phytoplankton can adapt to these environmental changes (Gienapp et al., 2008; Holt, 1990; Visser, 2008) and models that can incorporate the potential of organisms to adapt to changing conditions will be crucial if we are to predict the consequences of these changes. Unfortunately, the functional relationships used in models to describe different groups are fixed (Hood et al., 2007), implying that the functional groups cannot adapt to changes in the local environment.

Recent modeling studies have worked to capture the interactions between phytoplankton community structure and environmental conditions in a less prescriptive manner (Follows and Dutkiewicz, 2011; Follows et al., 2007). These models are initialized with numerous phytoplankton groups, with the physiological parameters for each group randomly selected from ranges in the literature and competition between phytoplankton groups is determined by the local environmental conditions. Distributions and changes in phytoplankton community structure obtained from these models closely resemble those seen in observational data.

The last twenty years have also seen a dramatic increase in our knowledge of the genetic diversity and biogeochemical functioning of marine phytoplankton and microbial communities (Dupont et al., 2007; Kerkhof and Goodman, 2009). This has been accompanied by an astronomical increase in the amount of available genomic information about these communities. While the genomic data itself provides a library of potential biogeochemical processes, use of transcriptomic data reveals which pathways





BGD
10, 815–850, 2013



are actually being utilized (Frias-Lopez et al., 2008; Hewson et al., 2009a; Poretsky et al., 2005, 2009). However, this genomic and transcriptomic information has yet to be explicitly incorporated into marine biogeochemical models (Hood et al., 2007).

In this paper we present results from a new model that explicitly incorporates genomic information and links gene expression with changing environmental conditions to simulate phytoplankton community structure. Although the model structure builds on that developed by Follows et al. (2007), each type of phytoplankton in our model is able to adapt to available nutrient concentrations by using different pathways of nutrient utilization encoded in its genome. Since the transcriptome reflects the genes that are being actively expressed under a specific environmental condition, the simulated amount of the activation of a gene qualitatively represents its transcription levels. In the methods section we describe the structure of the model, its components, and how competition for available resources and growth between species is represented in the model. In the results we describe the outcome within several idealized scenarios, and finally we discuss the model results in the context of microbial diversity, and potential ecological applications.

## 2 Methods

We develop the model with a continuous stirred-tank reactor (CSTR) system to represent simplified environmental conditions, allowing us to control inputs of nutrients to the system and ignore physical transport and mixing terms. The phytoplankton communities are modeled by extending the methods of Follows et al. (2007) to explicitly incorporate the information of genomics and the environmental conditions that induce gene expression, and allow the organisms to adapt to changing nutrient conditions by switching on different nutrient uptake and utilization pathways. The model tracks the response of 96 different types of phytoplankton, differentiated by their cell size, genetic makeup and physiological constraints (i.e. cell quotas).





BGD
10, 815–850, 2013



## 2.1 Ecosystem model simulation in a CSTR system

The nutrient inputs to the model include three nitrogen-bearing nutrients (ammonia ($NH_4^+$), nitrate ($NO_3^-$), and dissolved nitrogen gas ($N_2$)) and three phosphorus-bearing nutrients (phosphate ($PO_4^{3-}$), dissolved phosphomonoester (COP), and dissolved phosphonate (CP)). The different nitrogen (N) and phosphorus (P) compounds are not distinguished within the cell, so that $PO_4^{3-}$, COP, and CP all contribute to P cell quota. The dynamics of the nutrient concentrations and phytoplankton populations are described by the following system of coupled ordinary differential equations.

$$\frac{dC_i}{dt} = -\sum_{j=1}^{n}(A_{ij} \times B_j) + f_i \times \sum_{j=1}^{n}(m_j \times B_j \times Q_{j(\mathrm{N\,or\,P})}) + d \times (C_{i\_\mathrm{in}} - C_i) \quad (1)$$

$$\frac{dQ_{j(\mathrm{N\,or\,P})}}{dt} = \sum_{i=1}^{3}(A_{ij} - u_j \times Q_{j(\mathrm{N\,or\,P})}) \quad (2)$$

$$\frac{dB_j}{dt} = (\mu_j - m_j - d) \times B_j \quad (3)$$

where $C_i$ is the concentration of $NH_4^+$, $NO_3^-$, $PO_4^{3-}$, COP or CP – $N_2$ is assumed to be constant within the reactor with a value (20 mg $L^{-1}$) determined by $N_2$ solubility at 20 °C and 1 atm. $C_{i\_\mathrm{in}}$ is the concentration of nutrient $i$ in the input flow, $A_{ij}$ is the uptake rate of nutrient $i$ by phytoplankton type $j$, $Q_{j,(\mathrm{N\,or\,P})}$ is the cell quota of N or P, $\mu_j$ is the specific growth rate, $B_j$ is the cell density, $d$ is the CSTR dilution rate, and $m_j$ represents the mortality rate for phytoplankton type $j$. As the contents of the cells are remineralized, a fraction $f_i$ of the cell biomass is returned to the CSTR as nutrient $i$.

## 2.2 Selection and activation of functional gene clusters

We avoid the complexity of a whole-cell model (Kettler et al., 2007) and focus instead only on the effects of adaptation to variable N and P concentrations, and elemental





BGD
10, 815–850, 2013



forms in the environment. These adaptations involve complex metabolic pathways controlled by multiple genes. To simplify the consideration of how active metabolic pathways influence a particular biochemical process, we utilize the concept of a gene cluster (Frias et al., 1997); each cluster represents the collection of genes responsible for the uptake of a specified form of nutrient and the model considers the effects of 6 gene clusters. For example, nitrate assimilation into the cell involves an active transporter and an intracellular two-step reduction to ammonium by nitrate and nitrite reductase. For the sake of simplicity we consider the combination of genes responsible for $NO_3^-$ transport and reduction (e.g. nrt, nap, nir, nar) to be a gene cluster and name them as nr in the model. The nr gene cluster is highly expressed when extracellular $NH_4^+$ concentrations are low (e.g. $0.5 \sim 1.0\,\mu M$) (Eppley et al., 1969; Frias et al., 1997; Kikuchi et al., 1996). The nitrogenase (nif) gene cluster performing $N_2$ fixation among cyanobacteria is highly expressed in cultures depleted with dissolved inorganic nitrogen (Huang et al., 1999; Flores et al., 1999). The expression of cell-wall associated $NH_4^+$ or $PO_4^{3-}$ transporters (amt or pst) has been observed when extracellular nutrient concentrations fall below a certain level; 1 µM for amt, (Montesinos et al., 1998); 50 nM for pst, (Scanlan et al., 1997). The uptake of COP and CP is mediated by P related gene clusters (e.g. phoA, phnCDE) and the expression of these gene clusters is regulated by extracellular $PO_4^{3-}$ concentration (Ray et al., 1991; Kononova and Nesmeyanova, 2002). Thus we consider amt, nr and nif gene clusters to be activated when $NH_4^+$ concentration drops below 1 µM and pst, pho, and phn gene clusters to be activated when $PO_4^{3-}$ concentration is lower than 50 nM in our model.

### 2.3 Assembling phytoplankton communities in the model

We based the development of the phytoplankton community on the methodology used by Follows et al. (2007). First, we set up two phytoplankton size classes, large and small, with diameters of 10 µm and 1 µm respectively and maximum specific growth rates of $2\,d^{-1}$ and $1\,d^{-1}$ respectively. In the model, large phytoplankton are fast-growing, "opportunistic" phytoplankton (*r* strategists), while small, slower growing phytoplankton





BGD
10, 815–850, 2013



are species who are more competitive in minimal resource environments (K strategists) (Barton et al., 2010; Kilham and Hecky, 1988). The small size class is populated with 64 types of phytoplankton, sufficient to represent all combinations of the 6 gene clusters. $N_2$ fixation is restricted to the small size class and so only 32 phytoplankton types are required for the large size class. Among the resulting 96 distinct phytoplankton types, each has a different combination of gene clusters determining the suite of nutrients that an individual phytoplankton type can use.

In what follows, we represent the gene cluster combination for a specific phytoplankton type as an ordered sequence (amt, nr, nif, pst, pho, phn) with a zero or one indicating absence or presence of that specific gene cluster. For example, a phytoplankton type containing amt, nr and pho gene clusters is represented by the sequence "110010".

Phytoplankton types also differ in their cell quotas ($Q$). Maximum and minimum cell quotas for N and P (Table 1) are determined using allometric relationships (Finkel et al., 2004; Raven and Kubler, 2002), assuming that the cells were spherical.

### 2.4 Nutrient uptake rate

We model nutrient uptake as a reaction-diffusion process (Berg and Purcell, 1977; Munk and Riley, 1952; Volker and Wolf-Gladrow, 1999). Nutrient uptake is through either simple diffusion without the help of membrane transporters or facilitated diffusion through membrane transporters after the corresponding gene cluster is expressed. Assuming transport limitation, uptake rate is equal to the nutrient diffusive flux to the surface of a cell of radius $R$ (Jumars et al., 1993; KarpBoss et al., 1996) by:

$$J = 4\pi R S_h D \cdot (C - c(R)) \tag{4}$$

where $C$ is the nutrient concentration in the extracellular medium, $c(R)$ is the concentration at the cell surface, $S_h$ is the Sherwood number and $D$ is the diffusion coefficient.

We define critical extracellular concentrations ($c_{0,i}$) for $NH_4^+$ and $PO_4^{3-}$ which determine the activation of corresponding gene clusters and hence the mode of transport





BGD
10, 815–850, 2013



across the cell membrane; $c_{0,NH_4}(R) = 1\,\mu M$ for $NH_4^+$, $c_{0,PO4}(R) = 50\,nM$ for $PO_4^{3-}$. If the concentration of a given nutrient in the medium, $C_j$, is greater than $c_{0,j}(R)$, nutrient uptake is through simple diffusion according to Eq. (5) (Pasciak and Gavis, 1974)

$$J_j = 4\pi R S_h D_j \cdot (C_j - c_{0,j}(R)) \quad (5)$$

When $C_j$ drops below $c_{0,j}(R)$, the appropriate gene cluster is expressed, producing the corresponding nutrient-binding proteins. These proteins capture available nutrient molecules near the cell surface resulting in a zero concentration of the nutrient at the cell wall. In this case, phytoplankton that possesses those gene clusters are able to take up nutrients through facilitated diffusion (Munk and Riley, 1952) according to Eq. (6):

$$J_j = 4\pi R S_h D_j \cdot C_j \quad (6)$$

Thus $c_0$ serves as a signal of gene activation and a switch between simple and facilitated diffusion.

$NO_3^-$ and $N_2$ act as alternative N sources in the model. The expression of nr and nif gene clusters is assumed to be regulated by the critical concentration of $NH_4^+$ such that when $c_{NH_3} < c_{0-NH_3}$, nr and/or nif is expressed and the cell can utilize $NO_3^-$ and/or $N_2$. In this case, the concentrations of $NO_3^-$ or $N_2$ at the cell surface are set to zero, simulating facilitated diffusion, and nutrients are taken up according to Eq. (6). Uptake of DOP (dissolved organic phosphorus, including COP and CP) is modeled in a similar way. Once the extracellular $PO_4^{3-}$ concentration drops below $c_{0-PO_4}$, the gene clusters facilitating DOP uptake (pho and phn) are expressed, the concentration of DOP at the cell surface is set to zero, and Eq. (6) is used to calculate their uptake rate.





BGD
10, 815–850, 2013



## 2.5 Adjustment of nutrient uptake rate

The nutrient uptake rates calculated above are further adjusted using the maximum cell quota (Lehman et al., 1975; Thingstad, 1987):

$$A_{ij} = J_i \times \left( \frac{Q^{\max}_{j(\mathrm{N\,or\,P})} - Q_{j(\mathrm{N\,or\,P})}}{Q_{j(\mathrm{N\,or\,P})} - Q^{\min}_{j(\mathrm{N\,or\,P})}} \right) \quad (7)$$

($i = \mathrm{NH_4^+}$, $\mathrm{NO_3^-}$, $\mathrm{N_2}$, $\mathrm{PO_4^{3-}}$, COP and CP)

where $J_i$ is the dissolved nutrient flux reaching the cell surface. N and P are modeled separately, and their interaction (Pahlow and Oschlies, 2009) is currently not considered.

## 2.6 Gene-dependent maximum growth rates ($\mu'_{\max}$) and specific growth rate ($\mu$)

Each cell in the model is assigned an intrinsic maximum growth rate according to the cell size class ($\mu_{\max} = 2\,\mathrm{d}^{-1}$ for large cells and $1\,\mathrm{d}^{-1}$ for small cells). These rates are then adjusted as described below for individual cells according to the complement of genes they have and their cell quota.

The first adjustment incorporates the idea that larger genome size tends to lower the maximum growth rates due to the increase in cell maintenance costs (Hessen et al., 2010). We calculate gene-dependent maximum growth rate of individual phytoplankton type according to

$$\mu'_{\max,j} = \left( 1 - \sum_{i=1}^{6} e_i \times r_i \right) \times \mu_{\max,j} \quad (8)$$

where the index $i$ represents the gene cluster, $e_i$ is 1 if that gene cluster is present and 0 otherwise, and $r_i$ is an arbitrary growth reduction constant for the gene cluster





BGD
10, 815–850, 2013



$i$. The determination of $r_i$ is based on the average number of genes making up the gene cluster and we assume that each gene in the cluster will cause a 1 % decrease in maximum growth rate. For example, if a cell possesses the nif gene cluster only (comprised of 14 genes), its $\mu'_{max} = (1 - 0.14) \cdot u_{max} = 0.86 \cdot u_{max}$. Thus phytoplankton types having more functional gene clusters receive a greater penalty in terms of maximum growth rate, but benefit from being able to utilize more forms of N or P as a nutrient supply. This penalty also prevents us from producing "super phytoplankton" that can dominate under all environmental conditions because there are no costs associated with harboring a genetic capability.

Next, specific growth rates of phytoplankton type $j$ for each nutrient $i$ are derived from the values of the intracellular nutrient quota, $Q_j$, and its minimum value for that nutrient, $Q_j^{min}$, according to (Flynn et al., 1997)

$$\mu_{i,j} = \mu'_{max,j} \cdot \left(1 - \frac{Q_{i,j}^{min}}{Q_{i,j}}\right) \quad (9)$$

where $\mu'_{max,j}$ is gene-dependent maximum growth rate. Between N and P, the primary limiting nutrient is determined to be the nutrient that allows the lowest specific growth rate. However, we consider that the growth is not limited by two nutrients if both $Q_j > 0.95 \cdot Q_j^{max}$, where the intracellular nutrient quota for cell $j$ is almost saturated.

Finally, we use the Law of the Minimum to calculate the specific growth rate based on the current limiting nutrient (Elrifi and Turpin, 1985; Zonneveld, 1996).

$$\mu_j = \min(\mu_{N,j}, \mu_{P,j}) \quad (10)$$

### 2.7 Model scenarios

We use four simple scenarios, divided into two groups (I and II) according to the nutrient flux, with the high flux situation further subdivided according to N : P stochiometry, to





BGD
10, 815–850, 2013



examine how the modeled phytoplankton community responds to changes in environmental availability. The rate of water exchange (e.g. seasonal variability due to mixing) is modeled as variations of dilution rate within the CSTR. In each scenario, nutrient concentrations and dilution rates from the input flow are calculated as a cosine curve in time with a zero in the middle of the year and a maximum (Table 2) in the beginning and end of each year (Eqs. 11–13):

$$d_j = (d_{j\_\text{high}} - d_{j\_\text{low}}) \times f\_\text{time} + d_{j\_\text{low}} \quad (11)$$

$$C_{i\_\text{in},j} = (C_{i\_\text{in},j\_\text{high}} - C_{i\_\text{in},j\_\text{low}}) \times f\_\text{time} + C_{i\_\text{in},j\_\text{low}} \quad (12)$$

$$f\_\text{time} = \frac{\cos\left(\frac{2\pi}{365} \times (t-1)\right) + 1}{2} \quad (13)$$

where $d_j$ is the CSTR dilution rate in scenario $j$, $C_{i\_\text{in},j}$ is the concentration of nutrient $i$ ($NH_4^+$, $NO_3^-$, $PO_4^{3-}$, COP or CP) from the input flow in scenario $j$, and $t$ is time (day).

Scenarios I.1–3 have abundant nutrient loading from the input flow and higher dilution rates from late fall to early spring. These scenarios can be thought to represent an idealized surface ocean with the deepening of the mixed layer and abundant nutrient injection from deep water in the spring. Decreasing nutrient input and dilution rate during summer represents increasing stratification in the water column and an oligotrophic system. Such seasonal variation mimicked by the CSTR might represent oceans at higher latitude. The difference among scenarios I.1–3 lies in the ratio of nutrient concentrations (N : P) specified in the input flow so that Scenario I.1 is considered to be P limited whereas Scenario I.2 and I.3 are increasingly N limited.

Scenario II has low nutrient loading and dilution rates throughout the whole year representing constantly stratified and oligotrophic surface water such as those found at lower latitudes.

This effort is meant to be a proof of concept. Other important environmental variables (e.g. light) and pathways (e.g. carbon assimilation) are not included at present, but are able to be incorporated once corresponding physiology characteristics and gene





BGD
10, 815–850, 2013



clusters are assigned to each phytoplankton type and described in a proper mathematical way. All other parameters that are not found in method section are listed in the Table 3.

## 3 Results

### 3.1 Seasonal variation: total biomass, diversity, and dominant species

Because the model at present includes no seasonal light limitation, it does not mimic any real ocean ecosystem through a seaonal cycle. It does, however, illustrate the phytoplankton response to changing nutrients. The patterns of phytoplankton biomass in Scenarios I.1 and I.2 are similar (Fig. 1a, b), decreasing from a maximum in winter, reaching a minimum in summer shortly after the nutrient input becomes zero, and increasing again from early fall to winter. Total biomass in Scenario I.3 also decreases from winter and reaches a minimum in summer (Fig. 1c). In contrast to Scenarios I.1 and I.2 where biomass gradually increases in fall, total biomass in I.3 quickly increases and reaches the yearly peak around day 245. This difference is a result of the rapid growth by $N_2$ fixers activating the nif gene cluster who relieve N limitation during the fall period of low N input. If $N_2$ fixers are removed from the Scenario I.3 by eliminating the nif gene cluster, the pattern of total biomass is similar to that in Scenarios I.1 and I.2 (Fig. 1d). Total biomass in Scenario II, with decreased input of both N and P, is much lower than in Scenarios I.1–3 and patterns are more constant throughout the year (Fig. 1e).

Diversity of phytoplankton types in Scenarios I.1–3 is similar (Fig. 2), with the lowest in the winter and highest in the summer and fall. Scenario II shows a higher overall diversity with less variation throughout the year. Large phytoplankton types with higher maximum growth rates dominate in the highly variable environmental conditions from late fall to early spring of Scenario I.1–3. When dilution rates and nutrient inputs decrease in summer, small phytoplankton starts to compete better with large





BGD
10, 815–850, 2013



phytoplankton and eventually dominate the community. In contrast to Scenarios I.1–3, small phytoplankton types dominate the community over the whole annual cycle in Scenario II (Fig. 1).

### 3.2 Impact of environmental selection: gene cluster combinations and activation ratios

Seasonal changes in the combination of gene clusters among dominant species in Scenario I.1–3 closely follow nutrient variations in the input flow. In Scenario I.1, with a high ratio of N : P (100 : 1) in the input flow, large phytoplankton with the genetic combination of "000111" (absence of amt, nr, and nif; presence of pst, pho, and phn; Table 1) outcompete all other species from late fall to early spring. Possessing three P related gene clusters allows this phytoplankton type to efficiently utilize all sources of P under P limitation (Table 4), and possessing no N related gene clusters imposes no additional penalties on the maximum growth rate. Thus the phytoplankton type with the genetic combination of "000111" has maximal fitness in an N-replete, P-limited environment. Small phytoplankton become dominant during the summer in Scenario I.1 and three types with different genetic combinations contribute high and comparable amounts of biomass (Table 4). Although their growth is limited by P, possessing multiple P related gene clusters among small phytoplankton in summer is less advantageous. For example, the first two dominant species cannot use DOP due to the lack of pho and phn gene clusters, but achieve a higher growth rate compared to the third dominant species.

In Scenario I.2, after P limitation is alleviated by increasing the N : P ratio in the input flow to 16 : 1 (Redfield ratio), the dominant type in the winter and spring has the gene cluster combination "110110" (Table 4), allowing it to utilize 4 different types of nutrients ($NH_4^+$, $NO_3^-$, $PO_4^{3-}$ and COP) out of 5 possible (large phytoplankton types never possess the "nif" gene cluster in the model). Although the most dominant type in summer is large, the phytoplankton community is still dominated (> 50 % biomass) by small phytoplankton types.





BGD
10, 815–850, 2013



In Scenario I.3, after the N : P ratio in the input flow decreases to a value of 1 : 1, thereby causing N-limiting conditions, large phytoplankton possessing the gene cluster combination of "110000" outcompete all other types of phytoplankton. Analogous to Scenario I.1, the fitness of this dominant phytoplankton type is maximized through possessing both possible N related gene clusters (to utilize $NH_4^+$ and $NO_3^-$) while incurring no growth penalties for possessing P related genes (Table 4). Although the model imposes a strong N limitation in Scenario I.3 in summer, the growth of the most dominant species (small type "011010") is actually limited by P. In fact, the assimilation of $N_2$ as an unlimited N source by possessing the nif gene cluster makes it the only species whose growth is limited by P in Scenario I.3.

In Scenario II, with much less variability of dilution rate and lower nutrient input throughout the year (Table 2), the most dominant species are generally small, slow-growing phytoplankton, the exception being during days 186–201 when a large phytoplankton type contributes most of the biomass (Table 4). Although N : P in the input is set to the Redfield ratio, the primary limiting nutrient for all dominant species is P due to the N fixation, except several short periods of time when either does not limit phytoplankton growth. Since P is the primary limiting nutrient, species having more P related gene clusters gradually show more development advantage and become dominant.

Explicitly incorporating N and P related gene clusters into the model allows us to explore the impact of environmental variability on phytoplankton community structure. This impact is quantitatively expressed as the ratio of the activated N : P gene clusters (Fig. 3). We define the number of activated N and P related gene clusters as $N_{act}$, and $P_{act}$ respectively. In Scenario I.1, $N_{act} : P_{act} < 1$, consistent with the conditions of abundant N in the input flow (Fig. 3). In Scenario I.3, $N_{act} : P_{act} > 1$ which is a result of the intensive utilization of N during spring and nitrogen fixing activity during summer. In Scenario I.2, $N_{act} : P_{act}$ fluctuates around the value of 1, and is < 1 for the most of time in Scenario II because of P limitation (Fig. 3).





BGD
10, 815–850, 2013



### 3.3 Response of the phytoplankton community to environmental change

The nif gene clusters in the model illustrate how an expressed gene cluster changes the external environment. $N_2$ fixers become the dominant phytoplankton types as P loading increases from Scenario I.1 to Scenario I.3. They first appear to be the fourth dominant species in Scenario I.2 (gene combination "001100") and become the most dominant type in Scenario I.3 (gene combination "011010", Table 4). In Scenario II, $N_2$ fixers dominate for 9 months and bring in $\sim 0.31\ \mu mol\, L^{-1}\, day^{-1}$ that is relatively evenly distributed throughout the year (Fig. 4).

The succession of phytoplankton types in Scenario II provides an example of how $N_2$ fixers shape the environment, affect phytoplankton community structure, and how the phytoplankton community adapts to the environment changes. First of all, increased dissolved inorganic nitrogen ($DIN = [NH_4^+]+[NO_3^-]$) in the CSTR stimulates phytoplankton growth except $N_2$ fixers due to their relatively low growth rate (note: possession of nif gene cluster causes a 14 % reduction from intrisic maximum growth rate) and the loss of N fixation pathway if their nif gene cluster is deactivated at $[NH_4^+] > 1\ \mu M$. After DIN stops increasing in CSTR, other types of phytoplankton start to decline due to N limitation, but $N_2$ fixers start to grow by taking up $N_2$ (Fig. 5). Eventually, new N brought in through $N_2$ fixers increases DIN, stimulates the growth of other phytoplankton in an echo bloom, but depresses the growth of $N_2$ fixers themselves.

## 4 Discussion

Despite not having light controls, the patterns of phytoplankton community diversity in the model scenarios share some similarities with observed biodiversity patterns. Seasonally, the phytoplankton community in variable environments (Scenarios I.1–3) alternates between dominance by a single, fast-growing phytoplankton type in winter and spring to dominance by a group of small, slow-growing types in the summer and early autumn. Reduced seasonal variability resulting in a permanently oligotrophic water





BGD
10, 815–850, 2013



column (Scenario II) results in less variable phytoplankton biomass accumulation, enhanced diversity, and more stable communities composed mainly of slow-growing phytoplankton types. These patterns compare well to observed phytoplankton community shifts in surface waters of HOT (Hawaii Ocean Time-series) and BATS (Bermuda Atlantic Time-series Study) (Campbell et al., 1997; DuRand et al., 2001). For example, the surface phytoplankton community at BATS alternates between a single dominant eukaryotic phytoplankter in the winter and spring and several cyanobacteria species in the summer (Treusch et al., 2009). The reduced seasonal variability and stratified water column at HOT results in more stable phytoplankton communities throughout the whole year. The dominant species at HOT are quite close to those found during summer stratification at BATS (Shi et al., 2011). At larger scales, higher latitude oceans with strongly seasonal changes tend to select for fast-growing phytoplankton, but biodiversity is observed to decrease from pole to equator (Pommier et al., 2007). This pattern is echoed in the increased diversity between higher latitude Scenarios I.1–3 compared with tropical Scenario II.

The simple ecosystem model is able to adapt to changing environmental conditions. The alternating pattern between $N_2$ fixers and non $N_2$ fixers is consistent with patterns hypothesized by Hood et al. (Hood et al., 2001) and demonstrated by Coles et al. (Coles et al., 2004), and this N transfer from $N_2$ fixers to their partners was later demonstrated in field experiments (Foster et al., 2011). When $N_2$ fixation is a major pathway for N supply at Station ALOHA (Dore et al., 2002), $N_2$ fixation leads to reduced concentration of $PO_4^{3-}$ and the phytoplankton community eventually becomes P-stressed (Hebel and Karl, 2001; Karl et al., 2001). In situ experiments also confirmed that P was a major limiting nutrient for $N_2$ fixers (Moisander et al., 2007; Sanudo-Wilhelmy et al., 2001). Our model results in Scenario II reflect those observations: cell quota ratio of N : P is all above typical Redfield ratio (16) and the phytoplankton community eventually adapts itself to those who possess more P related gene clusters (Table 4). The community gene expression from both culture (Dyhrman and Haley, 2006) and in situ observations (Hewson et al., 2009b; Sowell et al., 2009) also shows that P-regulated transcripts are





BGD
10, 815–850, 2013



elevated to meet the phosphorus demand after N inputs through $N_2$ fixation. Since we lack the concrete evidence of how many transcripts can be produced from an activated gene cluster, our model does not explicitly output transcript numbers. However, if we assume the transcripts should be proportional to the amount of activated gene clusters, the model output of $N_{act}$ and $P_{act}$ can be qualitatively related to the transcriptome. Thus, the result $N_{act} : P_{act} < 1$ among phytoplankton communities in Scenario II (Fig. 3) is consistent with what has been found in the above studies. $N_{act} : P_{act}$ in our model also serves as an indicator of predicting ecosystem changes: P limited if $N_{act} : P_{act} > 1$; N limited if $N_{act} : P_{act} < 1$. This incorporation of genetic information in our model may provide a deeper understanding of how phytoplankton will adapt to and alter biogeochemical scenarios in a future ocean. By incorporating gene clusters for silicic acid transport (sit) (Martin-Jezequel et al., 2000), our model would be suited to studying diatom-diazotroph symbiotic associations (DDAs) which represent a previously overlooked upper ocean silica cycle (Brzezinski et al., 1998) with relevance to carbon sequestration (Karl et al., 2012; Subramaniam et al., 2008).

Traditional marine ecosystem models (e.g. NPZD or PFTs type) have difficulties addressing how phytoplankton will respond to a rapidly changing environment. This has led to the development of more biologically detailed models (Flynn et al., 1997; Pahlow and Oschlies, 2009). However, the crucial parameterizations of physiology are extremely difficult to obtain since quantitative, physiological information from laboratory measurements is scarce. As pointed by Anderson (Anderson, 2005), increasing model complexity without robust parameterization is not helpful. Recent genomic discoveries regarding ecological and biogeochemical pathways can improve the accuracy of ecosystem models to predict climate change (Doney et al., 2004), provided we meet the challenge of incorporating them into models in a realistic fashion. Our study is the first attempt we are aware of to build an ecosystem model that integrates microbial genomics and gene expressions with marine biogeochemical cycling to predict relative transcription levels of nutrient acquisition genes. Although relating biogeochemical functions to genomic information leads us further down the path of increased





BGD
10, 815–850, 2013



model complexity, we ignore intracellular biochemistry and metabolic pathways to reduce model complexity while holding the concept that models should be simple enough to describe the system but complex enough to capture important processes.

A pitfall of the model lies in estimating key physiological parameters. How we determine the gene-dependent maximum growth rate based on the maximum growth rate and genome size for a particular type of phytoplankton is based on qualitative estimation (Hessen et al., 2010). Such trade-offs are used to accommodate the observation that there is no a single organism that can dominate all environments. However, to our best knowledge, there is still no experimental data available to quantify such a relationship and the application of Eq. (8) causes our model to remain heuristic. The regulation of gene activation is also crucial information, but activation regulation in our model has been simply controlled by nutrient concentration in CSTR. More complex regulations and limiting factors have been ignored. For example, the regulation of expression of nif gene clusters by factors such as oxygen (Gallon, 1981, 1992), temperature (Carpenter and Capone, 1992), light (Capone et al., 1997), and Fe (Howard and Rees, 1996) is not included in the model, leading to overestimates $N_2$ fixation rates in the model compared to field estimates (Table 5). The idealized CSTR system used in the model is another simplification, as we focused in this case on exploring the incorporation of genetic information into biogeochemical models. Despite the simplifications, simulation results provided insights regarding what role these selected gene clusters play in changing ecosystem structure and biogeochemical cycles. To be applied in a specific ecosystem, the current model needs to include other important factors such as light, temperature, and predation pressure, and be embedded in a large-scale biogeochemical model.

*Acknowledgements.* We thank M. Follows for useful discussions concerning ecosystem modeling. We also appreciate all our ROCA (the River Ocean Continuum of the Amazon) colleagues who worked in the field and laboratory, especially J. Richey, J. Meng, C. Smith, and B. Satinsky. This work was funded by the Gordon and Betty Moore Foundation.





BGD
10, 815–850, 2013



## References


Anderson, T. R.: Plankton functional type modelling: running before we can walk?, J. Plankton Res., 27, 1073–1081, doi:10.1093/plankt/fbi076, 2005.

Barnett, T., Zwiers, F., Hegerl, G., Allen, M., Crowley, T., Gillett, N., Hasselmann, K., Jones, P., Santer, B., Schnur, R., Scott, P., Taylor, K., and Tett, S.: Detecting and attributing external influences on the climate system: a review of recent advances, J. Climate, 18, 1291–1314, 2005.

Barton, A. D., Dutkiewicz, S., Flierl, G., Bragg, J., and Follows, M. J.: Response to Comment on "Patterns of diversity in marine phytoplankton", Science, 329, 5991, doi:10.1126/science.1190048, 2010.

Beman, J. M., Chow, C.-E., King, A. L., Feng, Y., Fuhrman, J. A., Andersson, A., Bates, N. R., Popp, B. N., and Hutchins, D. A.: Global declines in oceanic nitrification rates as a consequence of ocean acidification, P. Natl. Acad. Sci. USA, 108, 208–213, doi:10.1073/pnas.1011053108, 2011.

Berg, H. C. and Purcell, E. M.: Physics of chemoreception, Biophys. J., 20, 193–219, 1977.

Brzezinski, M. A., Villareal, T. A., and Lipschultz, F.: Silica production and the contribution of diatoms to new and primary production in the central North Pacific, Mar. Ecol.-Prog. Ser., 167, 89–104, doi:10.3354/meps167089, 1998.

Campbell, L., Liu, H. B., Nolla, H. A., and Vaulot, D.: Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991–1994 ENSO event, Deep-Sea Res. Pt. I, 44, 167–192, doi:10.1016/s0967-0637(96)00102-1, 1997.

Capone, D. G., Zehr, J. P., Paerl, H. W., Bergman, B., and Carpenter, E. J.: Trichodesmium, a globally significant marine cyanobacterium, Science, 276, 1221–1229, doi:10.1126/science.276.5316.1221, 1997.

Carpenter, E. J., and Capone, D. G.: Nitrogen fixation in Trichodesmium blooms, in: Marine Pelagic Cyanobacteria: Trichodesmium and Other Diazotrophs, edited by: Carpenter, E. J., Capone, D. G., and Rueter, J., Kluwer Academic Publishers, The Netherlands, 211–217, 1992.

Chai, F., Dugdale, R. C., Peng, T. H., Wilkerson, F. P., and Barber, R. T.: One-dimensional ecosystem model of the equatorial Pacific upwelling system. Part I: model development and silicon and nitrogen cycle, Deep-Sea Res. Pt. II, 49, 2713–2745, 2002.






BGD
10, 815–850, 2013



Coles, V. J., Hood, R. R., Pascual, M., and Capone, D. G.: Modeling the impact of Trichodesmium and nitrogen fixation in the Atlantic Ocean, J. Geophys. Res.-Oceans, 109, C06007, doi:10.1029/2002jc001754, 2004.

Doney, S. C., Abbott, M. R., Cullen, J. J., Karl, D. M., and Rothstein, L.: From genes to ecosystems: the ocean's new frontier, Front. Ecol. Environ., 2, 457–466, 2004.

Doney, S. C., Fabry, V. J., Feely, R. A., and Kleypas, J. A.: Ocean Acidification: The Other $CO_2$ Problem, Annu. Rev. Mar. Sci., 1, 169–192, doi:10.1146/annurev.marine.010908.163834, 2009.

Dore, J. E., Brum, J. R., Tupas, L. M., and Karl, D. M.: Seasonal and interannual variability in sources of nitrogen supporting export in the oligotrophic subtropical North Pacific Ocean, Limnol. Oceanogr., 47, 1595–1607, 2002.

Dupont, S., Wilson, K., Obst, M., Skold, H., Nakano, H., and Thorndyke, M. C.: Marine ecological genomics: when genomics meets marine ecology, Mar. Ecol.-Prog. Ser., 332, 257–273, doi:10.3354/meps332257, 2007.

DuRand, M. D., Olson, R. J., and Chisholm, S. W.: Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea, Deep-Sea Res. Pt. II, 48, 1983–2003, doi:10.1016/s0967-0645(00)00166-1, 2001.

Dyhrman, S. T., and Haley, S. T.: Phosphorus scavenging in the unicellular marine diazotroph *Crocosphaera watsonii*, Appl. Environ. Microbiol., 72, 1452–1458, doi:10.1128/aem.72.2.1452-1458.2006, 2006.

Elrifi, I. R. and Turpin, D. H.: Steady-state luxury consumption and the concept of optimum nutrient ratios – a study with phosphate and nitrate limited selenastrum-minutum (Chlorophyta), J. Phycol., 21, 592–602, 1985.

Eppley, R. W., Coatswor, J. L., and Solorzan, L.: Studies of nitrate reductase in marine phytoplankton, Limnol. Oceanogr., 14, 194–205, 1969.

Fasham, M. J. R., Flynn, K. J., Pondaven, P., Anderson, T. R., and Boyd, P. W.: Development of a robust marine ecosystem model to predict the role of iron in biogeochemical cycles: a comparison of results for iron-replete and iron-limited areas, and the SOIREE iron-enrichment experiment, Deep-Sea Res. Pt. I, 53, 333–366, doi:10.1016/j.dsr.2005.09.011, 2006.

Finkel, Z. V., Irwin, A. J., and Schofield, O.: Resource limitation alters the 3/4 size scaling of metabolic rates in phytoplankton, Mar. Ecol.-Prog. Ser., 273, 269–279, doi:10.3354/meps273269, 2004.





BGD
10, 815–850, 2013



Flores, E., Muro-Pastor, A. M., and Herrero, A.: Cyanobacterial nitrogen assimilation genes and NtcA-dependent control of gene expression, in: The Phototrophic Prokaryotes, edited by: Peschek, G. A., Loffelhardt, W., and Schmetterer, G., Plenum Publishing Corporation, New York, NY, 1999.

Flynn, K. J., Fasham, M. J. R., and Hipkin, C. R.: Modelling the interactions between ammonium and nitrate uptake in marine phytoplankton, Philos. T. Roy. Soc. B, 352, 1625–1645, 1997.

Follows, M. J. and Dutkiewicz, S.: Modeling diverse communities of marine microbes, Annu. Rev. Mar. Sci., 3, 427–451, 2011.

Follows, M. J., Dutkiewicz, S., Grant, S., and Chisholm, S. W.: Emergent biogeography of microbial communities in a model ocean, Science, 315, 1843–1846, doi:10.1126/science.1138544, 2007.

Foster, R. A., Kuypers, M. M. M., Vagner, T., Paerl, R. W., Musat, N., and Zehr, J. P.: Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses, ISME J., 5, 1484–1493, doi:10.1038/ismej.2011.26, 2011.

Franks, P. J. S., Wroblewski, J. S., and Flierl, G. R.: Behavior of a simple plankton model with food-level acclimation by herbivores, Mar. Biol., 91, 121–129, doi:10.1007/bf00397577, 1986.

Frias, J. E., Flores, E., and Herrero, A.: Nitrate assimilation gene cluster from the heterocyst-forming cyanobacterium Anabaena sp strain PCC 7120, J. Bacteriol., 179, 477–486, 1997.

Frias-Lopez, J., Shi, Y., Tyson, G. W., Coleman, M. L., Schuster, S. C., Chisholm, S. W., and DeLong, E. F.: Microbial community gene expression in ocean surface waters, P. Natl. Acad. Sci. USA, 105, 3805–3810, doi:10.1073/pnas.0708897105, 2008.

Gallon, J. R.: The oxygen sensitivity of nitrogenase – a problem for biochemists and microorganisms, Trends Biochem. Sci., 6, 19–23, doi:10.1016/0968-0004(81)90008-6, 1981.

Gallon, J. R.: Reconciling the incompatible – N-2 fixation and O-2, New Phytol., 122, 571–609, 1992.

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A., and Merila, J.: Climate change and evolution: disentangling environmental and genetic responses, Mol. Ecol., 17, 167–178, doi:10.1111/j.1365-294X.2007.03413.x, 2008.

Gregg, W. W., Ginoux, P., Schopf, P. S., and Casey, N. W.: Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model, Deep-Sea Res. Pt. II, 50, 3143–3169, doi:10.1016/j.dsr2.2003.07.013, 2003.





BGD
10, 815–850, 2013



Hebel, D. V. and Karl, D. M.: Seasonal, interannual and decadal variations in particulate matter concentrations and composition in the subtropical North Pacific Ocean, Deep-Sea Res. Pt. II, 48, 1669–1695, doi:10.1016/s0967-0645(00)00155-7, 2001.

Hessen, D. O., Jeyasingh, P. D., Neiman, M., and Weider, L. J.: Genome streamlining and the elemental costs of growth, Trends Ecol. Evol., 25, 75–80, doi:10.1016/j.tree.2009.08.004, 2010.

Hewson, I., Poretsky, R. S., Beinart, R. A., White, A. E., Shi, T., Bench, S. R., Moisander, P. H., Paerl, R. W., Tripp, H. J., Montoya, J. P., Moran, M. A., and Zehr, J. P.: In situ transcriptomic analysis of the globally important keystone N-2-fixing taxon *Crocosphaera watsonii*, ISME J., 3, 618–631, doi:10.1038/ismej.2009.8, 2009a.

Hewson, I., Poretsky, R. S., Dyhrman, S. T., Zielinski, B., White, A. E., Tripp, H. J., Montoya, J. P., and Zehr, J. P.: Microbial community gene expression within colonies of the diazotroph, Trichodesmium, from the Southwest Pacific Ocean, ISME J., 3, 1286–1300, doi:10.1038/ismej.2009.75, 2009b.

Holt, R. D.: The microevolutionary consequences of climate change, Trends Ecol. Evol., 5, 311–315, doi:10.1016/0169-5347(90)90088-u, 1990.

Hood, R. R., Bates, N. R., Capone, D. G., and Olson, D. B.: Modeling the effect of nitrogen fixation on carbon and nitrogen fluxes at BATS, Deep-Sea Res. Pt. II, 48, 1609–1648, 2001.

Hood, R. R., Laws, E. A., Armstrong, R. A., Bates, N. R., Brown, C. W., Carlson, C. A., Chai, F., Doney, S. C., Falkowski, P. G., Feely, R. A., Friedrichs, M. A. M., Landry, M. R., Moore, J. K., Nelson, D. M., Richardson, T. L., Salihoglu, B., Schartau, M., Toole, D. A., and Wiggert, J. D.: Pelagic functional group modeling: Progress, challenges and prospects, Deep-Sea Res. Pt. II, 53, 459–512, 2006.

Hood, R. R., Laws, E. A., Follows, M. J., and Siegel, D. A.: Modeling and prediction of marine microbial populations in the genomic era, Oceanography, 20, 155–165, 2007.

Howard, J. B. and Rees, D. C.: Structural basis of biological nitrogen fixation, Chem. Rev., 96, 2965–2982, doi:10.1021/cr9500545, 1996.

Huang, T. C., Lin, R. F., Chu, M. K., and Chen, H. M.: Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium *Synechococcus* sp. strain RF-1, Microbiol.-UK, 145, 743–753, 1999.

Jumars, P. A., Deming, J. W., Hill, P. S., Karp-Boss, L., Yager, P. L., and Dade, W. B.: Physical constraints on marine osmotrophy in an optimal foraging context, Mar. Microb. Food Web., 7, 121–159, 1993.





BGD
10, 815–850, 2013



Karl, D. M., Bidigare, R. R., and Letelier, R. M.: Long-term changes in plankton community structure and productivity in the North Pacific Subtropical Gyre: the domain shift hypothesis, Deep-Sea Res. Pt. II, 48, 1449–1470, doi:10.1016/s0967-0645(00)00149-1, 2001.
Karl, D. M., Church, M. J., Dore, J. E., Letelier, R. M., and Mahaffey, C.: Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation, P. Natl. Acad. Sci. USA, 109, 1842–1849, doi:10.1073/pnas.1120312109, 2012.
KarpBoss, L., Boss, E., and Jumars, P. A.: Nutrient fluxes to planktonic osmotrophs in the presence of fluid motion, Oceanogr. Mar. Biol., 34, 71–107, 1996.
Kerkhof, L. J. and Goodman, R. M.: Ocean microbial metagenomics, Deep-Sea Res. Pt. II, 56, 1824–1829, doi:10.1016/j.dsr2.2009.05.005, 2009.
Kettler, G. C., Martiny, A. C., Huang, K., Zucker, J., Coleman, M. L., Rodrigue, S., Chen, F., Lapidus, A., Ferriera, S., Johnson, J., Steglich, C., Church, G. M., Richardson, P., and Chisholm, S. W.: Patterns and implications of gene gain and loss in the evolution of Prochlorococcus, PLOS Genet., 3, 2515–2528, e23110.1371/journal.pgen.0030231, 2007.
Kikuchi, H., Aichi, M., Suzuki, I., and Omata, T.: Positive regulation by nitrite of the nitrate assimilation operon in the cyanobacteria *Synechococcus* sp strain PCC 7942 and Plectonema boryanum, J. Bacteriol., 178, 5822–5825, 1996.
Kilham, P. and Hecky, R. E.: Comparative ecology of marine and fresh-water phytoplankton, Limnol. Oceanogr., 33, 776–795, 1988.
Kononova, S. V., and Nesmeyanova, M. A.: Phosphonates and their degradation by microorganisms, Biochemistry-Moscow, 67, 184–195, doi:10.1023/a:1014409929875, 2002.
Landry, M. R. and Kirchman, D. L.: Microbial community structure and variability in the tropical Pacific, Deep-Sea Res. Pt. II, 49, 2669–2693, 2002.
Le Quéré, C.: Reply to Horizons Article "Plankton functional type modelling: running before we can walk" Anderson (2005): I. Abrupt changes in marine ecosystems?, J. Plankton Res., 28, 871–872, doi:10.1093/plankt/fbl014, 2006.
Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Da Cunha, L. C., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., and Wolf-Gladrow, D.: Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, Global Change Biol., 11, 2016–2040, doi:10.1111/j.1365-2486.2005.1004.x, 2005.





BGD
10, 815–850, 2013



Lehman, J. T., Botkin, D. B., and Likens, G. E.: Assumptions and rationales of a computer model of phytoplankton population-dynamics, Limnol. Oceanogr., 20, 343–364, 1975.
Litchman, E., Klausmeier, C. A., Miller, J. R., Schofield, O. M., and Falkowski, P. G.: Multi-nutrient, multi-group model of present and future oceanic phytoplankton communities, Biogeosciences, 3, 585–606, doi:10.5194/bg-3-585-2006, 2006.
Martin-Jezequel, V., Hildebrand, M., and Brzezinski, M. A.: Silicon metabolism in diatoms: implications for growth, J. Phycol., 36, 821–840, doi:10.1046/j.1529-8817.2000.00019.x, 2000.
Moisander, P. H., Paerl, H. W., Dyble, J., and Sivonen, K.: Phosphorus limitation and diel control of nitrogen-fixing cyanobacteria in the Baltic Sea, Mar. Ecol.-Prog. Ser., 345, 41–50, doi:10.3354/meps06964, 2007.
Moloney, C. L. and Field, J. G.: The size-based dynamics of plankton food webs .1. a simulation-model of carbno and nitrogen flows, J. Plankton Res., 13, 1003–1038, doi:10.1093/plankt/13.5.1003, 1991.
Montesinos, M. L., Muro-Pastor, A. M., Herrero, A., and Flores, E.: Ammonium/Methylammonium permeases of a cyanobacterium – identification and analysis of three nitrogen-regulated amt genes in *Synechocystis* sp. PCC 6803, J. Biol. Chem., 273, 31463–31470, doi:10.1074/jbc.273.47.31463, 1998.
Moore, J. K., Doney, S. C., Kleypas, J. A., Glover, D. M., and Fung, I. Y.: An intermediate complexity marine ecosystem model for the global domain, Deep-Sea Res. Pt. II, 49, 403–462, 2002.
Munk, W. H. and Riley, G. A.: Absorption of nutrients by aquatic plants, J. Mar. Res., 11, 215–240, 1952.
Pahlow, M. and Oschlies, A.: Chain model of phytoplankton P, N and light colimitation, Mar. Ecol.-Prog. Ser., 376, 69–83, doi:10.3354/meps07748, 2009.
Pasciak, W. J. and Gavis, J.: Transport limitation of nutrient uptake in phytoplankton, Limnol. Oceanogr., 19, 881–898, 1974.
Pommier, T., Canback, B., Riemann, L., Bostrom, K. H., Simu, K., Lundberg, P., Tunlid, A., and Hagstrom, A.: Global patterns of diversity and community structure in marine bacterioplankton, Mol. Ecol., 16, 867–880, doi:10.1111/j.1365-294X.2006.03189.x, 2007.
Poretsky, R. S., Bano, N., Buchan, A., LeCleir, G., Kleikemper, J., Pickering, M., Pate, W. M., Moran, M. A., and Hollibaugh, J. T.: Analysis of microbial gene transcripts in environmental samples, Appl. Environ. Microbiol., 71, 4121–4126, doi:10.1128/aem.71.7.4121-4126.2005, 2005.





BGD
10, 815–850, 2013



Poretsky, R. S., Hewson, I., Sun, S. L., Allen, A. E., Zehr, J. P., and Moran, M. A.: Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre, Environ. Microbiol., 11, 1358–1375, doi:10.1111/j.1462-2920.2008.01863.x, 2009.
Raven, J. A. and Kubler, J. E.: New light on the scaling of metabolic rate with the size of algae, J. Phycol., 38, 11–16, doi:10.1046/j.1529-8817.2002.01125.x, 2002.
Ray, J. M., Bhaya, D., Block, M. A., and Grossman, A. R.: Isolation, transcription, and inactivation of the gene for an atypical alkaline-phosphatase of *Synechococcus*-sp strain PCC 7942, J. Bacteriol., 173, 4297–4309, 1991.
Sanudo-Wilhelmy, S. A., Kustka, A. B., Gobler, C. J., Hutchins, D. A., Yang, M., Lwiza, K., Burns, J., Capone, D. G., Raven, J. A., and Carpenter, E. J.: Phosphorus limitation of nitrogen fixation by Trichodesmium in the central Atlantic Ocean, Nature, 411, 66–69, doi:10.1038/35075041, 2001.
Sarmiento, J. L. and Gruber, N.: Ocean Biogeochemical Dynamics, Princeton University Press, Princeton, NJ, USA, 2006.
Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S. A., and Stouffer, R.: Response of ocean ecosystems to climate warming, Global Biogeochem. Cy., 18, Gb3003, doi:10.1029/2003gb002134, 2004.
Scanlan, D. J., Silman, N. J., Donald, K. M., Wilson, W. H., Carr, N. G., Joint, I., and Mann, N. H.: An immunological approach to detect phosphate stress in populations and single cells of photosynthetic picoplankton, Appl. Environ. Microbiol., 63, 2411–2420, 1997.
Shi, Y., Tyson, G. W., Eppley, J. M., and DeLong, E. F.: Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean, ISME J., 5, 999–1013, doi:10.1038/ismej.2010.189, 2011.
Sowell, S. M., Wilhelm, L. J., Norbeck, A. D., Lipton, M. S., Nicora, C. D., Barofsky, D. F., Carlson, C. A., Smith, R. D., and Giovanonni, S. J.: Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea, ISME J., 3, 93–105, doi:10.1038/ismej.2008.83, 2009.
Steele, J.: Stability of plankton ecosystems, in: Ecological Stability, edited by: Usher, M. B. and Williamson, M. H., Halsted Press, A division of John Wiley and Sons, New York, NY, USA, York, England, 179–191, 1974.
Subramaniam, A., Yager, P. L., Carpenter, E. J., Mahaffey, C., Bjorkman, K., Cooley, S., Kustka, A. B., Montoya, J. P., Sanudo-Wilhelmy, S. A., Shipe, R., and Capone, D. G.: Amazon





BGD
10, 815–850, 2013



River enhances diazotrophy and carbon sequestration in the tropical North Atlantic Ocean, P. Natl. Acad. Sci. USA, 105, 10460–10465, doi:10.1073/pnas.0710279105, 2008.
Thingstad, T. F.: Utilization of N, P, and organic C by heterotrophic bacteria. 1. Outline of a chemostat theory with a consistent concept of maintenance metabolism, Mar. Ecol.-Prog. Ser., 35, 99–109, doi:10.3354/meps035099, 1987.
Treusch, A. H., Vergin, K. L., Finlay, L. A., Donatz, M. G., Burton, R. M., Carlson, C. A., and Giovannoni, S. J.: Seasonality and vertical structure of microbial communities in an ocean gyre, ISME J., 3, 1148–1163, doi:10.1038/ismej.2009.60, 2009.
Visser, M. E.: Keeping up with a warming world; assessing the rate of adaptation to climate change, P. R. Soc. B, 275, 649–659, doi:10.1098/rspb.2007.0997, 2008.
Volker, C. and Wolf-Gladrow, D. A.: Physical limits on iron uptake mediated by siderophores or surface reductases, Mar. Chem., 65, 227–244, doi:10.1016/s0304-4203(99)00004-3, 1999.
Zonneveld, C.: Modelling the kinetics of non-limiting nutrients in microalgae, J. Marine Syst., 9, 121–136, doi:10.1016/0924-7963(96)00021-8, 1996.





BGD
10, 815–850, 2013



**Table 1.** Size-dependent scaling parameters ($y$) of cell quota, in the form of $y = aV^b$. $a$ and $b$: intercept and size-scaling exponent, respectively. V: cell size in um$^3$. $Q_N^{min}$ and $Q_N^{max}$: minimum or maximum cell quota of nitrogen, respectively. $Q_P^{min}$ and $Q_P^{max}$: minimum or maximum cell quota of phosphorus, respectively.

| Parameters (y) | Parameter units | A | B |
|---|---|---|---|
| $Q_N^{min}$ | fmol N cell$^{-1}$ | 1.5 | 0.72 |
| $Q_N^{max}$ | fmol N cell$^{-1}$ | 3 | 0.85 |
| $Q_P^{min}$ | fmol P cell$^{-1}$ | 0.0934 | 0.72 |
| $Q_P^{max}$ | fmol P cell$^{-1}$ | 0.1867 | 0.85 |





BGD
10, 815–850, 2013



**Table 2.** Variations in nutrient concentrations and dilution rates in the input flow of the CSTR for four different scenarios.

| Scenario (ocean system that can be represented) | Dilution rate ($d^{-1}$) | Nutrient concentrations in the inflow (μM) | | |
|---|---|---|---|---|
| | | nutrient | maximum | minimum |
| Scenario I.1 (oceans at higher latitude) | d_high = 0.1; d_low = 0; | $NH_4^+$ | 150 | 0 |
| | | $NO_3^-$ | 200 | 0 |
| | | $PO_4^{3-}$ | 2.5 | 0 |
| | | COP | 0.5 | 0 |
| | | CP | 0.5 | 0 |
| | | | (TN : TP = 100 : 1) | |
| Scenario I.2 (oceans at higher latitude) | d_high = 0.1; d_low = 0; | $NH_4^+$ | 150 | 0 |
| | | $NO_3^-$ | 200 | 0 |
| | | $PO_4^{3-}$ | 12 | 0 |
| | | COP | 5 | 0 |
| | | CP | 5 | 0 |
| | | | (TN : TP = 16 : 1) | |
| Scenario I.3 (oceans at higher latitude) | d_high = 0.1; d_low = 0; | $NH_4^+$ | 150 | 0 |
| | | $NO_3^-$ | 200 | 0 |
| | | $PO_4^{3-}$ | 150 | 0 |
| | | COP | 100 | 0 |
| | | CP | 100 | 0 |
| | | | (TN : TP = 1 : 1) | |
| Scenario II (oceans at lower latitude) | d_high = 0.01; d_low = 0; | $NH_4^+$ | 0 | 0 |
| | | $NO_3^-$ | 5 | 0 |
| | | $PO_4^{3-}$ | 0 | 0 |
| | | COP | 0.15 | 0 |
| | | CP | 0.15 | 0 |
| | | | (TN : TP = 16 : 1) | |





BGD
10, 815–850, 2013



**Table 3.** Parameters of the ecosystem model.

| Symbol | Descriptions | Value when assigned | Units |
|---|---|---|---|
| $Q$ | intracellular nutrient content or cell-quota | – | fmol cell$^{-1}$ |
| $Q^N$ | N cell-quota | – | fmol N cell$^{-1}$ |
| $Q^P$ | P cell-quota | – | fmol P cell$^{-1}$ |
| $Q_{min}$ | minimum intracellular nutrient content required for survival | – | fmol cell$^{-1}$ |
| $Q_{max}$ | Maximum intracellular nutrient content | – | fmol cell$^{-1}$ |
| $Q^N_{min}$ | minimum intracellular N content required for survival | – | fmol N cell$^{-1}$ |
| $Q^N_{max}$ | maximum intracellular N content | – | fmol N cell$^{-1}$ |
| $Q^P_{min}$ | minimum intracellular P content required for survival | – | fmol P cell$^{-1}$ |
| $Q^P_{max}$ | maximum intracellular P content | – | fmol P cell$^{-1}$ |
| $[NH_4^+]$ | extracellular ammonia concentration | – | μmol N L$^{-1}$ |
| $[NO_3^-]$ | extracellular nitrate concentration | – | μmol N L$^{-1}$ |
| $[N_2]$ | extracellular $N_2$ concentration | – | μmol N L$^{-1}$ |
| $[PO_4^{3-}]$ | extracellular phosphate concentration | – | μmol P L$^{-1}$ |
| [COP] | extracellular phosphomonoester concentration | – | μmol P L$^{-1}$ |
| [CP] | extracellular phosphonate concentration | – | μmol P L$^{-1}$ |
| $u_{max}$ | Maximum growth rate | – | day$^{-1}$ |
| $u'_{max}$ | adjusted maximum growth rate | – | day$^{-1}$ |
| $u_N$ | Nitrogen-limited specific growth rate | – | day$^{-1}$ |
| $u_P$ | phosphorus-limited specific growth rate | – | day$^{-1}$ |
| $u$ | Nutrient-limited specific growth rate | – | day$^{-1}$ |
| $e$ | gene existence coefficient | 1: existence;<br>0: non-existence; | – |
| $r$ | growth reduction constant | $r_{amt}$: 0.01<br>$r_{nr}$: 0.06<br>$r_{nif}$: 0.12<br>$r_{pst}$: 0.04<br>$r_{cop}$: 0.03<br>$r_{cp}$: 0.14 | – |
| $J$ | dissolved nutrient flux reaching the cell surface | – | fmol element cell$^{-1}$ day$^{-1}$ |
| $R$ | cell radius | – | μm |
| $D$ | diffusion coefficient for each nutrient | $D_{NH_4}$: 0.00016<br>$D_{NO_3}$: 0.00015<br>$D_{N_2}$: 0.000017<br>$D_{PO_4}$: 0.000086<br>$D_{COP}$: 0.000093<br>$D_{CP}$: 0.000093 | m$^2$ day$^{-1}$ |
| $C$ | nutrient concentration in CSTR | – | fmol element L$^{-1}$ |
| $C_{in}$ | nutrient concentration in the input flow | – | fmol element L$^{-1}$ |
| $c(R)$ | nutrient concentration at the cell surface | – | fmol element L$^{-1}$ |
| $C_0$ | critical extracellular nutrient concentration for gene reduction | $C_0^{NH_4}$: 1<br>$C_0^{PO_4}$: 0.05 | μmol element L$^{-1}$ |
| $d$ | CSTR dilution rate | – | day$^{-1}$ |
| $S_h$ | Sherwood number | 1 | – |
| $A$ | nutrient uptake rate | – | fmol element cell$^{-1}$ day$^{-1}$ |
| $B$ | cell density | – | cells L$^{-1}$ |
| $m$ | cell motality | 0.1 | day$^{-1}$ |
| $f$ | the ratio of allocation of each nutrient mineralization | $f_{NH_4}$: 0.9<br>$f_{NO_3}$: 0.1<br>$f_{PO_4}$: 0.8<br>$f_{COP}$: 0.1<br>$f_{CP}$: 0.1 | – |





BGD
10, 815–850, 2013



**Table 4.** Seasonal variations of dominant species and their physiological information (the order of gene cluster combination: amt, nr, nif, pst, pho, phn; 0: absence, 1: presence; L: large, S: small types of phytoplankton).

| Scenario | Season (or days) | Contribution to total biomass from dominant species* | Genetic combination for dominant phytoplankton | Limiting nutrients | N : P in cell quota |
|---|---|---|---|---|---|
| I | Spring | 97.5 % | 000111 (L) | P | 44 |
| | Summer | 29.1 % | 000100 (S) | P | 19.7 |
| | | 21.4 % | 100100 (S) | P | 19.5 |
| | | 20.5 % | 110111(S) | P | 60.6 |
| II | Spring | 99.9 % | 110110 (L) | P | 24 |
| | Summer | 43.5 % | 110111 (L) | P | 33 |
| | | 29.3 % | 100100 (S) | None | 16 |
| | | 11.6 % | 110100 (S) | None | 16 |
| | | 11.0 % | 001100 (S) | None | 16 |
| III | Spring | 82.3 % | 110000 (L) | N | 5.6 |
| | Summer | 26.8 % | 011010 (S) | P | 17 |
| | | 25.3 % | 100010 (S) | N | 13 |
| | | 24.9 % | 110000 (S) | N | 13 |
| IV | 1–72 | 62.4 % | 111110 (S) | P | 22 |
| | 73–105 | 36.1 % | 110111 (S) | P or None | 19 |
| | 106–144 | 31.7 % | 011111 (S) | P or None | 18 |
| | 145–185 | 26.1 % | 001111 (S) | P or None | 18 |
| | 186–201 | 21.3 % | 100111 (L) | P | 50 |
| | 202–221 | 22.6 % | 001100 (S) | P | 22 |
| | 222–283 | 42.0 % | 001111 (S) | P | 20 |
| | 284–311 | 33.6 % | 011111 (S) | P or None | 17 |
| | 312–354 | 32.7 % | 110111 (S) | P or None | 18 |
| | 355–360 | 31.0 % | 111110 (S) | P | 17 |

* Number in this column refers to the percentage of total biomass contribution from the dominant species in corresponding period (seasons in Scenario I.1–I.3, and days in Scenario II).





BGD
10, 815–850, 2013



**Table 5.** Compilation of direct estimates of $N_2$ fixation rates from several regions of the ocean.

| Region | $N_2$ fixation rates ($\mu mol\,L^{-1}\,day^{-1}$) | Reference |
|---|---|---|
| Western English Channel (NE Atlantic) | 0.02 | Rees et al. (2009) |
| Tropical Atlantic | 0.04–0.06 | Voss et al. (2004) |
| North Pacific | 0.01–0.02 | Watkins-Brandt et al. (2011) |
| Tropical and subtropic Pacific | 0.001–0.01 | Moore et al. (2009) |





BGD
10, 815–850, 2013



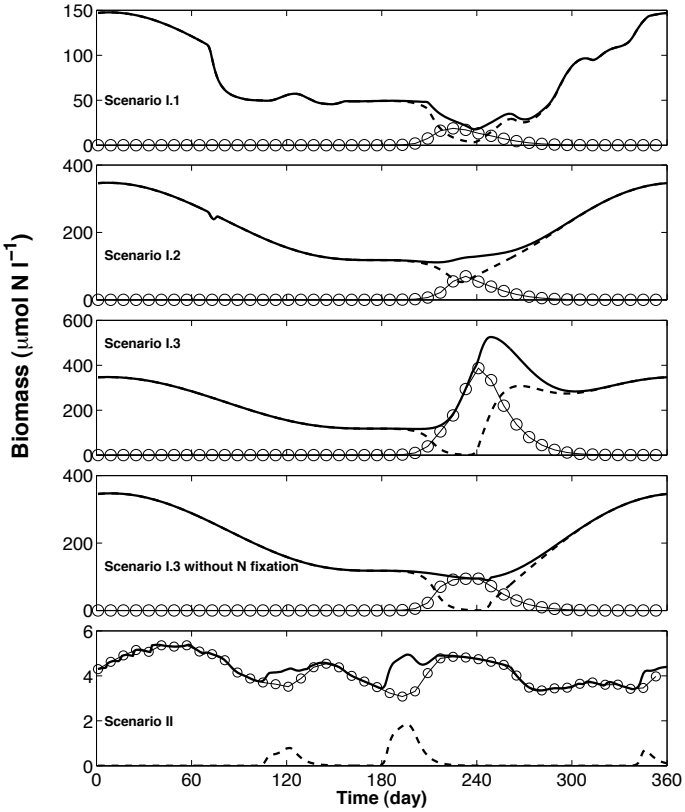


**Fig. 1.** Seasonal variations of total biomass (μmol L$^{-1}$ in units of nitrogen, solid line), biomass from small (circle line) and large phytoplankton (dashed line) in different scenarios.





BGD
10, 815–850, 2013



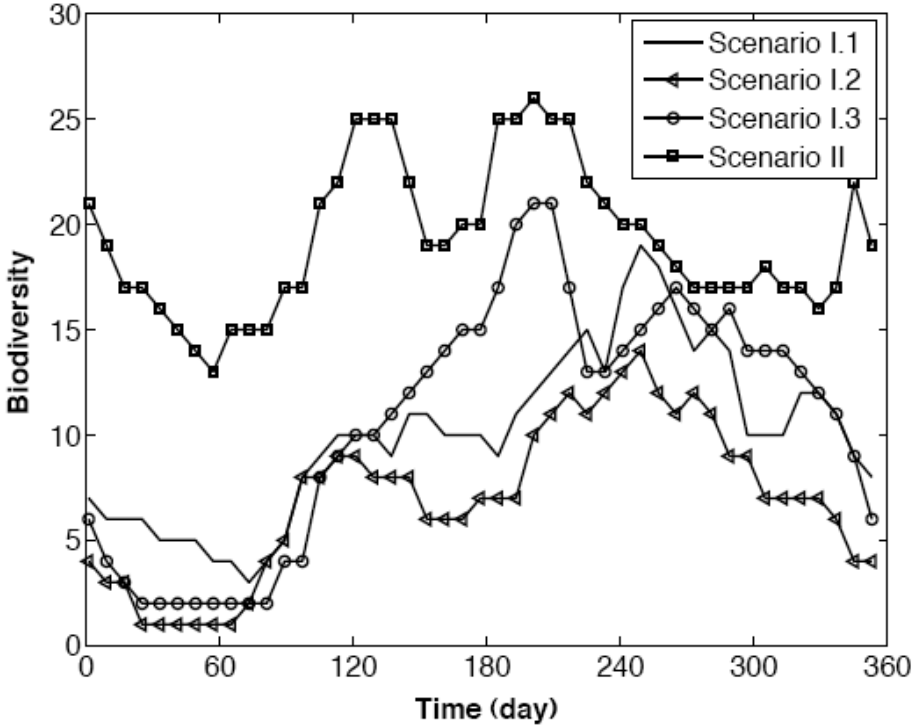


**Fig. 2.** Diversity of modeled phytoplankton. (Diversity is defined as the number of phytoplankton types comprising > 0.1 % of the total biomass.)





BGD
10, 815–850, 2013



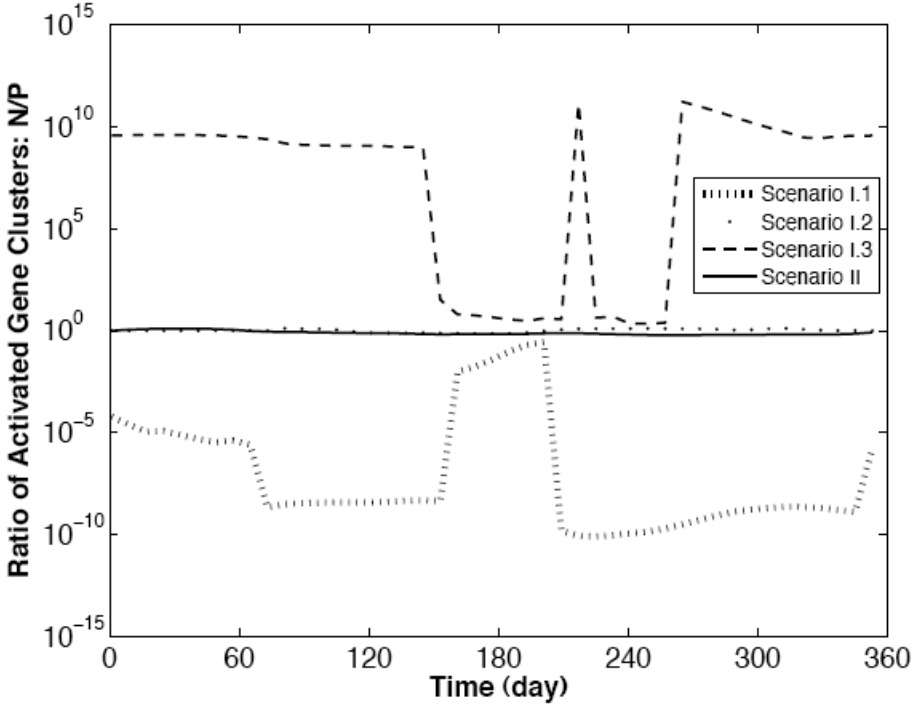


**Fig. 3.** Ratio of the activation of total N (amt, nr, and nif) and P (pst, pho, and phn) gene clusters among four scenarios.





BGD
10, 815–850, 2013



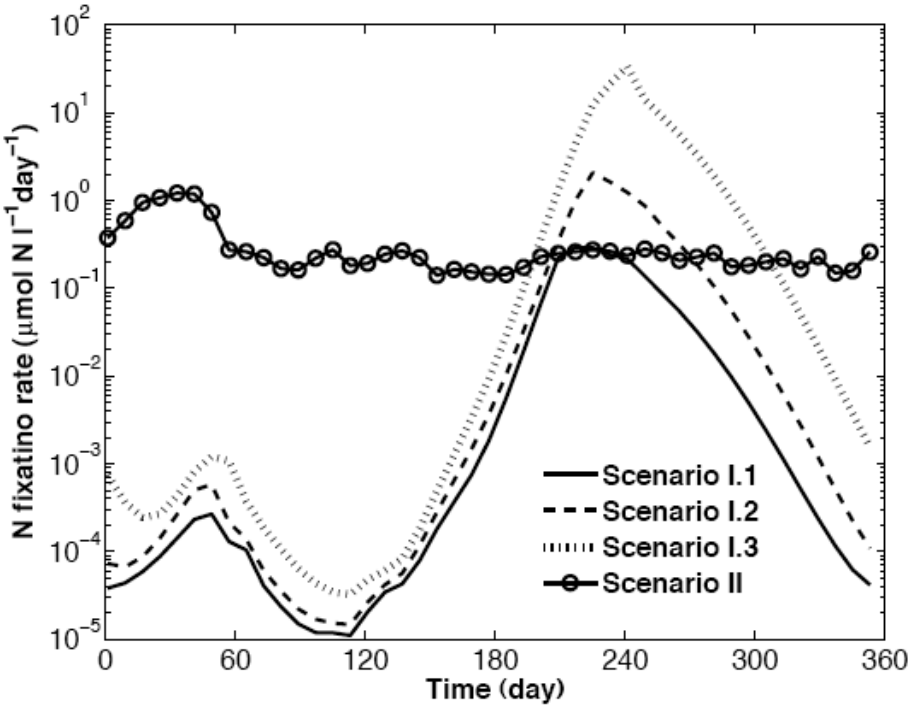


**Fig. 4.** Nitrogen fixation rate among 4 scenarios.





BGD
10, 815–850, 2013



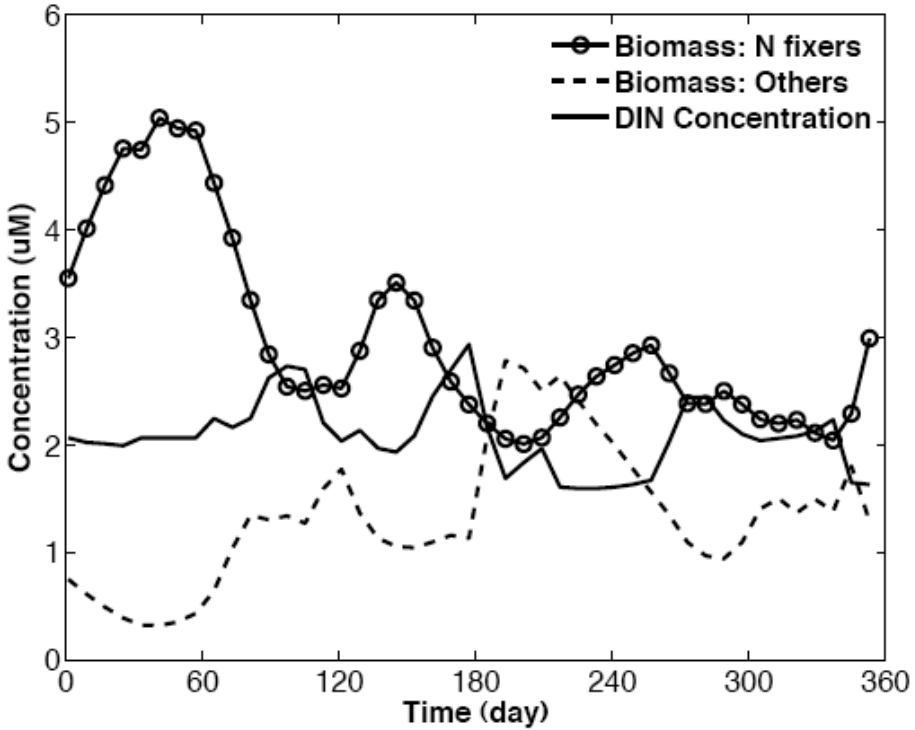


**Fig. 5.** Succession of phytoplankton community in Scenario II.





BGD
10, 815–850, 2013