Supporting Information

Methods

1. Global ocean ecosystem model
The ocean ecosystem model used in this study has been described previously by Follows et al. [1]. Changes in the concentration of the $ith$ nutrient ($Ni$), $jth$ phytoplankton group ($Pj$), $kth$ zooplankton group ($Zk$) and pools of particulate organic matter ($POMi$) and dissolved organic matter ($DOMi$) are modeled with the following equations:

\[
\frac{\partial N_i}{\partial t} + \nabla \cdot (uN_i) = \nabla \cdot (k\nabla N_i) - \sum_j \left[ \mu_j P_j R_{ij} \right] + S_{N_i} \quad \text{(SE1)}
\]

\[
\frac{\partial P_j}{\partial t} + \nabla \cdot (uP_j) = \nabla \cdot (k\nabla P_j) - \frac{\partial (w^p P_j)}{\partial z} - m^p_j P_j - \sum_k \left[ g_{jk} P_j Z_{k,i=i=1} + \mu_j P_j \right] \quad \text{(SE2)}
\]

\[
\frac{\partial Z_{ki}}{\partial t} + \nabla \cdot (uZ_{ki}) = \nabla \cdot (k\nabla Z_{ki}) - m^z_{ki} Z_{ki} + \sum_j \left[ \zeta_{jk} g_{jk} P_j R_{ij} \right] \quad \text{(SE3)}
\]

\[
\frac{\partial POM_i}{\partial t} + \nabla \cdot (uPOM_i) = \nabla \cdot (k\nabla POM_i) - r_{POM} POM_i - \frac{\partial (w_{POM} POM_i)}{\partial z} + S_{POM_i} \quad \text{(SE4)}
\]

\[
\frac{\partial DOM_i}{\partial t} + \nabla \cdot (uDOM_i) = \nabla \cdot (k\nabla DOM_i) - r_{DOM} DOM_i + S_{DOM_i} \quad \text{(SE5)}
\]

The parameters used in these equations are defined below (Table S1) and by Follows et al. [1] and Dutkiewicz et al. [2]. Inorganic nitrogen, phosphorus, iron and silicon nutrients ($Ni$) are taken up by phytoplankton. The rates of uptake are determined by the abundance ($P_j$, expressed in terms of phosphorus biomass), growth rate ($\mu_j$) and elemental ratios ($R_{ij}$) of the phytoplankton. The inorganic nutrients are supplied by different processes that are represented here by $S_i$ and can include remineralization, supplies from external sources, and abiotic transformations. Phytoplankton growth rates ($\mu_j$) are determined by irradiance, nutrient availability, and temperature (see below).

Phytoplankton sink through the water column according to a sinking rate, $w^p_j$, and are lost by mortality $m^p_j$ and grazing by zooplankton. The rate of grazing by the $kth$ zooplankton on the $jth$ phytoplankton is represented by $g_{jk}$, which in turn is given by the function $g_{jk} = g_{mjk} \frac{\eta_{jk}}{A_k + k^p_k}$. Here $g_{mjk}$ is the maximum rate at which the $kth$ zooplankton can graze the $jth$ phytoplankton per unit of phytoplankton biomass, $\eta_{jk}$ is the palatability of the $jth$ phytoplankton to the $kth$ zooplankton, $A_k$ is the sum of phytoplankton weighted by their palatability, $A_k = \sum_j \left[ \eta_{jk} P_j \right]$, and $k^p_k$ is the value of $A_k$ at which grazing by the $kth$ zooplankton reaches half its maximum rate. Zooplankton grow according to their rate of grazing, $g_{jk}$, and feeding efficiency, $\zeta_{jk}$, and are lost by mortality, $m^z_{ki}$.

The terms $S_{POM_i}$ and $S_{DOM_i}$ represent the rates of supply of $POM_i$ and $DOM_i$ (respectively) from the decomposition of biomass, excretion, and inefficient feeding by
zooplankton on phytoplankton. Fixed proportions of decomposing and excreted material \( (f_{POM}) \), and of grazed but unconsumed material \( (g_{POM}) \), pass to the POM pool, and the remaining proportions pass to the DOM pool. The terms \( r_{POM} \) and \( r_{DOM} \) represent the rate of loss of \( POM_i \) and \( DOM_i \) (respectively) by remineralization to inorganic compounds.

The constants \( u \) and \( \kappa \) represent advection and diffusion (respectively) in the model ocean and were provided by the ECCO (“Estimating the Circulation and Climate of the Ocean”) state estimates [3]. These equations were integrated in a 1° X 1° resolution ocean circulation model with 23 vertical layers [4]. The advection scheme used for these simulations was the ‘superbee’ second order flux limiter [5]. We also applied an additional flux limiter to prevent spurious negative concentrations that occur due to the enforcement that mixing occur along isopycnal surfaces.

These simulations were performed with 33 different types of phytoplankton, and 2 different types of zooplankton. The physiological properties and nutrient requirements of different phytoplankton groups were assigned randomly, subject to some simple taxonomic and allometric constraints, using the approach of Follows et al. [1]. In each integration, we initialized 15 ‘small’ phytoplankton types and 18 ‘large’ phytoplankton types. Small phytoplankton have slower maximum growth rates than large phytoplankton, but are able to grow at lower levels of light and nutrient availability (Fig. S1). Initial concentrations of nutrients and grazers were determined using previous simulations [1]. An ensemble of 10 integrations was performed, using different randomly generated phytoplankton communities.

2. Phytoplankton growth and nutrient use
Phytoplankton growth rates, \( \mu_j \), are determined by multiplying the maximum growth rate \( \mu_j \) for a phytoplankton group by three factors that represent the degree of limitation by temperature \( (\gamma_j^T) \), irradiance \( (\gamma_j^I) \), and nutrient availability \( (\gamma_j^N) \), such that
\
\[ \mu_j = \mu_{\text{max}} \gamma_j^T \gamma_j^I \gamma_j^N. \]

Values of \( \gamma_j^T \), \( \gamma_j^I \) and \( \gamma_j^N \) are each scaled to have a value between 0 and 1, and are dimensionless. The response of growth to local temperature \( (T) \) is represented by the function
\
\[ \gamma_j^T = \frac{1}{\tau_i}(A^\tau \exp(-B(T - T_o)^C) - \tau_2), \]

where \( A, B, C \) and \( T_o \) regulate the shape of the dependence of growth on temperature, and \( \tau_i, \tau_1 \) and \( \tau_2 \) are used to normalize \( \gamma_j^T \) to a maximum value (see Fig. S1A). The response of growth to local irradiance \( (I) \) is represented by the function
\
\[ \gamma_j^I = \frac{1}{F_m}(1 - \exp(-k_{\text{PAR}}I)) \exp(-k_{\text{inhib}}I), \]

where \( k_{\text{PAR}} \) controls the saturation of growth as a function of irradiance, \( k_{\text{inhib}} \) controls the inhibition of growth by high irradiance, and \( F_m \) is a factor that is used to normalize \( \gamma_j^I \) to a maximum value of 1 (Fig. S1B). The degree of nutrient limitation is determined by the ambient nutrient concentration that supports the lowest rate of growth, or \( \gamma_j^N = \min(N_j^I) \).
For iron, phosphorus and silica (diatoms only), \( N^{L}_i = \frac{N_i}{N_i + k_{Ni}} \), where \( k_{Ni} \) is the concentration of the \( i \)th nutrient at which the \( j \)th phytoplankton reaches half its maximum growth rate (Fig. S1C). Nitrogen is available in three inorganic forms, NO\(_3\), NO\(_2\) and NH\(_4\). The use of NO\(_3\) and NO\(_2\) are repressed when NH\(_4\) is available. This is represented in the model using the expression

\[
N^{L}_N = \frac{NO_3 + NO_2}{NO_3 + NO_2 + k_{NO}} \exp(-\psi NH_4) + \frac{NH_4}{NH_4 + k_{NH}}.
\]

During the course of the simulations, we introduce phytoplankton mutants (see below) that are unable to use NO\(_3\), but retain the ability to use NO\(_2\) and NH\(_4\). For these mutants, we use the expression

\[
N^{L}_N = \frac{NO_2}{NO_2 + k_{NO}} \exp(-\psi NH_4) + \frac{NH_4}{NH_4 + k_{NH}}.
\]

Other mutants lose the ability to use both NO\(_3\) and NO\(_2\), but retain the ability to use NH\(_4\). For these mutants, we represent limitation by nitrogen as

\[
N^{L}_N = \frac{NH_4}{NH_4 + k_{NH}}.
\]

### 3. Mutation

In the present study, we introduce mutations for the loss (and gain) of specific nitrogen use abilities. We apply these mutations to each of the 15 ‘small’ phytoplankton types, beginning after the third year of the simulation. At this point the model has settled into a relatively steady annually repeating cycle of phytoplankton abundances. Initially, all biomass in a specific phytoplankton type can use NO\(_3\), NO\(_2\) and NH\(_4\) as nitrogen sources. We refer to this biomass as comprising the ‘parent’ population. The process of mutation (see Fig. 1B) is simulated by transferring biomass between the parent population and an additional tracer that represents the biomass of ‘mutant’ phytoplankton that have identical properties to their parent, except for a specific difference that corresponds to the mutated phenotype.

The changes in the concentration of mutant (\( P_M \)) and parent (\( P_P \)) biomass are modeled as follows:

\[
\frac{\partial P_p}{\partial t} + \nabla \cdot (u P_p) = \nabla \cdot (k \nabla P_p) - L_p + \mu_p P_p - \Lambda + \Gamma
\]

\[
\frac{\partial P_m}{\partial t} + \nabla \cdot (u P_m) = \nabla \cdot (k \nabla P_m) - L_m + \mu_m P_m + \Lambda - \Gamma
\]

where the growth of a specific parent population by cell division is described by \( \mu_p P_p \), and for a specific mutant by \( \mu_m P_m \), as in Eq. SE2 and described in Section 2 of the Methods. A fraction, \( \Lambda_f \), of cell divisions result in a mutation (such as the loss of the ability to use nitrate) producing mutant biomass (\( P_M \)) at a rate of

\[
\Lambda = \Lambda_f \log_2(e) \mu_p P_p,
\]
where \( \log_2(e) \) adjusts the growth rate to cell divisions per day. We also allow a fraction, \( \Gamma_f \), of cell divisions of mutant biomass to result in mutation back to the parent phenotype. This results in the production of parent biomass at a rate of

\[
\Gamma = \Gamma_f \log_2(e) \mu MP_M .
\]

\( L_P \) and \( L_M \) represent the change in the concentration of parent and mutant phytoplankton (respectively) per unit biomass from sinking, mortality and grazing (as in Eq. SE2), or

\[
L_P = \frac{\partial(w^P_P P_P)}{\partial Z} + m^P_P P_P + \sum_k [g_{PK} P_P Z_{k,i=1}] \text{, and}
\]

\[
L_M = \frac{\partial(w^P_M P_M)}{\partial Z} + m^P_M P_M + \sum_k [g_{MK} P_M Z_{k,i=1}] .
\]

We use this approach to produce three types of mutants for each of the 15 different groups of small phytoplankton. These are (i) NO\(_3\) loss mutants, which lose the ability to use nitrate, (ii) NO\(_3\)/NO\(_2\) loss mutants, which lose the ability to use both nitrate and nitrite, and (iii) null mutants, which retain the ability to use ammonium, nitrite and nitrate.

For each of the three mutant types, the proportion of replications of the parent population resulting in mutation (\( \Lambda_f \)) was set to a value of \( \Lambda_f = 10^{-8} \). Observed mutation rates of microbes with DNA genomes are often approximately 0.003 per genome per replication [6,7]. For picocyanobacteria with genomes of size 1.75 Mbp – 3 Mbp, we would therefore expect mutation rates of approximately \( 10^{-9} \) per base pair per replication. Typically, mutations at a number of locations (bases) in the genome can lead to the loss of function of a particular gene or trait. This is called the ‘target size’ of a mutation. We do not know the target size for the loss of nitrate or nitrite use abilities in picocyanobacteria. We chose a value of \( \Lambda_f \) that is ten-fold greater than the expected mutation rate (\( \Lambda_f = 10^{-8} = 10^{-9} \times 10 \)), corresponding to a target size of approximately 10 bases. It is possible that target sizes for the loss of nitrogen use traits are substantially larger than this value. If so, we can think of the mutants in the model as representing cells with a mutation at a location that is part of a specific subset (i.e., 10 specific bases) of the locations in the mutational target for a phenotype. For instance, these might represent a specific group of mutations that can cause (i) the loss of function of nitrate reductase (NO\(_3\) loss mutant phenotype) (ii) the loss of function of nitrite reductase or (NO\(_3\)/NO\(_2\) loss mutant phenotype) and (iii) no change in nitrogen use abilities (null mutant phenotype). Given the considerable uncertainty concerning mutation rates (per base pair) and the target size for nitrogen use targets, we performed additional simulations using a greater mutation rate (\( \Lambda_f = 10^{-6} \) see Results, Section 2. ii), to make sure our observations were robust to variation in this parameter.

We set the rate of mutations that return mutants to the parent phenotype to a value of \( \Gamma_f = 10^{-12} \). It is not clear how often nitrate and nitrite use abilities might be newly acquired by cyanobacteria that lack them, though we anticipate it would be much smaller than the rate at which they are lost. We therefore chose a value that was several orders of magnitude smaller than \( \Lambda_f \). At this value, \( \Gamma_f \) is so small (and values of \( P_M \) remain so small) that is expected to have a negligible affect on the model solutions.
4. Supply of inorganic nutrients

The supply of inorganic nutrients is represented in equation SE1 by the term $S_{N_i}$, which includes different processes for the inorganic nutrients of different elements. The equations describing the supply of inorganic phosphorus, nitrogen (NO$_3$, NO$_2$ and NH$_4$), iron and silica are as follows:

\[
S_{PO4} = r_{DOM}DOP + r_{POM}PPO
\]

\[
S_{Si} = r_{POM}POSi
\]

\[
S_{Fe} = r_{DOM}DOFe + r_{POM}POFe - c_{scav}Fe' + \alpha F_{atmos}
\]

\[
S_{NO3} = \zeta_{NO3}NO2
\]

\[
S_{NO2} = \zeta_{NO2}NH4 - \zeta_{NO3}NO2
\]

\[
S_{NH4} = r_{DON}DON + r_{PON}PON - \zeta_{NO3}NH4
\]

Here $r_{DOM}$ and $r_{POM}$ represent the rate of remineralization of DOM and POM (respectively) forms of the $i$th element. The term $c_{scav}$ is the rate of scavenging of free iron, whose concentration, $Fe'$, is modeled following the approach of Parekh et al. [8]. The term $F_{atmos}$ represents the rate at which iron is deposited at the surface of the ocean in dust, and $\alpha$ represents the solubility of this iron. The oxidation of NO$_2$ to NO$_3$ is described by the term $\zeta_{NO3}$, and the oxidation of NH$_4$ to NO$_2$ is described by the term $\zeta_{NO2}$. In addition, we assume that the oxidation of NH$_4$ to NO$_2$ is photoinhibited, such that it does not occur if the ambient light is above 10 $\mu$E m$^{-2}$ s$^{-1}$.

Results

1. Ambient nitrogen concentrations

The annual average concentrations of nitrate, nitrite and ammonium in surface waters (upper 10 m) of the modeled ocean are illustrated in Fig. S2 A, C and E. Annual average concentrations of nitrate, nitrite and ammonium along Atlantic Meridional Transect [9] 13 (AMT 13) in the modeled ocean are illustrated in Fig. S2 D, E and F.

2. Sensitivity of mutant accumulation to model parameters

i. Time dependence of mutant accumulation

In the main text, our analyses concentrate on the abundance of mutants during the fifth year of a model integration (Figs 2 and 3). However, we expect the abundance of mutants in the model to change over time, as they are produced by parent populations. In Fig. S3, the distribution of phytoplankton in surface waters of the global ocean is illustrated for the tenth year of the integration. The distributions of small (parent) and large phytoplankton are similar in the fifth and tenth years. In the tenth year of the integration, mutants have accumulated to greater abundances than in the fifth year, but similar patterns are observed for the relative abundance of NO$_3$/NO$_2$ loss mutants and null mutants in different ocean regions. That is, while NO$_3$/NO$_2$ loss mutants reach abundances similar to null mutants in topical oligotrophic regions, null mutants are
substantially more abundant than NO$_3$/NO$_2$ loss mutants at higher latitudes. This means that the salient observations reported in the main text are not an artifact of the time at which the distribution of the mutants was studied.

**ii. Mutation rate**

We do not know the mutation rates for the loss of NO$_3$ use abilities or NO$_3$/NO$_2$ use abilities in marine cyanobacteria (see Methods, Section 3). We therefore wanted to study the distribution of mutants in the model using a different value for the mutation rate, and performed a simulation where the proportion of cell divisions resulting in mutations ($\Lambda_f$) was set to $\Lambda_f = 10^{-6}$, instead of $\Lambda_f = 10^{-8}$. In the simulation where $\Lambda_f = 10^{-6}$ (Fig. S4) the abundance of NO$_3$/NO$_2$ loss mutants and null mutants are approximately two orders of magnitude greater than in the simulation where $\Lambda_f = 10^{-8}$ (Fig. 2). However, changing the parameter $\Lambda_f$ to $10^{-6}$ did not substantially change the relative abundance of NO$_3$/NO$_2$ loss mutants and null mutants in different ocean regions. This suggests that our observations are robust to variation in the mutation rate over the range of values that were studied.

**iii. Oxidation rates of NH$_4$ and NO$_2$**

We used a model that parameterizes important biogeochemical transformations of nitrogen (see Methods, section 4). There is considerable uncertainty surrounding the rates at which some of these transformations occur in the environment. We therefore performed simulations using different values for two such parameters. Specifically, we used smaller values for the rate of (light-inhibited) oxidation of ammonium to nitrite ($\zeta_{NO_2} = 0.1$ d$^{-1}$) and the rate of oxidation of nitrite to nitrate ($\zeta_{NO_3} = 0.033$ d$^{-1}$). In Fig. S5, ambient concentrations of nitrate, nitrite and ammonium are illustrated for this integration, and in Fig. S6, the distribution of phytoplankton is illustrated for this integration. The lower values of $\zeta_{NO_2}$ and $\zeta_{NO_3}$ result in small changes in the distribution of ammonium and nitrite, including greater concentrations of ammonium in the lower euphotic zone, and near the surface in several regions with high productivity. However, changing the rates of oxidation of NH$_4$ and NO$_2$ did not substantially alter the observation that NO$_3$/NO$_2$ loss mutants reach similar abundances to null mutants in topical oligotrophic regions, but at higher latitudes, null mutants reach much higher abundances than NO$_3$/NO$_2$ loss mutants.

**iv. Randomly generated phytoplankton communities**

We performed integrations of the model with 10 randomly generated phytoplankton communities. In the main text, results are presented for one of these communities. In Fig. S7, the distribution of phytoplankton in global surface waters is illustrated for the remaining nine phytoplankton communities. There is some variation among these integrations, for example in the dominance of particular regions by large versus small phytoplankton. However, some important features of the distribution of NO$_3$/NO$_2$ loss mutants and null mutants are common to each of these integrations. In each case, NO$_3$/NO$_2$ loss mutants accumulate to similar abundances to null mutants in tropical oligotrophic regions, but at higher latitudes, null mutants occur at substantially greater abundances than NO$_3$/NO$_2$ loss mutants, indicating that NO$_3$/NO$_2$ loss mutants are disadvantaged in these regions.
Table S1. Parameters used in the model.

| parameter | description                        | value                  | units         |
|-----------|------------------------------------|------------------------|---------------|
| $\mu$     | Maximum phytoplankton growth rate  | Small: 1.4             | d$^{-1}$      |
|           |                                    | Large: 2.5             |               |
| $m^p$     | Phytoplankton mortality rate       | Small: 0.1             | d$^{-1}$      |
|           |                                    | Large: 0.1             |               |
| $k_{PAR}$ | PAR saturation coefficient         | Small: mean 0.012, std 0.02 | (µE m$^{-2}$ s$^{-1}$)$^{-1}$ |
|           |                                    | Large: mean 0.012, std 0.006 | (µE m$^{-2}$ s$^{-1}$)$^{-1}$ |
|           |                                    | (absolute value)       |               |
| $k_{inhb}$| PAR inhibition coefficient         | Small: mean 6×10$^{-3}$, std 1×10$^{-4}$ | (µE m$^{-2}$ s$^{-1}$)$^{-1}$ |
|           |                                    | Large: mean 1×10$^{-3}$, std 5×10$^{-5}$ | (µE m$^{-2}$ s$^{-1}$)$^{-1}$ |
|           |                                    | (absolute value)       |               |
| $A$       | Temperature curve coefficient       | 1.04                   |               |
| $T_o$     | Temperature optimum coefficient    | -2 to 30               | °C            |
| $B$       | Temperature range coefficient       | Small: 1×10$^{-3}$     | °C$^{-1}$     |
|           |                                    | Large: 3×10$^{-4}$     |               |
| $C$       | Temperature decay coefficient       | 4                      |               |
| $\tau_1, \tau_2$ | Temperature normalization coefficients | 3, 0.3 |               |
| $K_{PO4}$ | Phosphate half saturation          | Small: 1.0×10$^{-2}$ to 1.5×10$^{-2}$ | µM P          |
|           |                                    | Large: 3.5×10$^{-2}$ to 5.5×10$^{-2}$ | µM P          |
| $K_{NO3}$ | Nitrate half saturation            | Small: 0.16 to 0.24    | µM N          |
|           |                                    | Large: 0.56 to 0.88    | µM N          |
| $K_{NO2}$ | Nitrite half saturation            | Small: 0.16 to 0.24    | µM N          |
|           |                                    | Large: 0.56 to 0.88    | µM N          |
| $K_{NH4}$ | Ammonium half saturation           | Small: 0.08 to 0.12    | µM N          |
|           |                                    | Large: 0.28 to 0.44    | µM N          |
| $K_{Si}$  | Silicic acid half saturation       | Non-diatom: 0          | µM Si         |
|           |                                    | Diatom: 1              |               |
| $K_{Fe}$  | Iron half saturation               | Small: 1×10$^{-5}$ to 1.5×10$^{-5}$ | µM Fe         |
|           |                                    | Large: 3.5×10$^{-5}$ to 5.5×10$^{-5}$ | µM Fe         |
| $R_{Si:P}$| Phytoplankton elemental ratios     | 16                     |               |
| $R_{N:P}$ |                                    | 16                     |               |
| $R_{Fe:P}$|                                    | 1×10$^{-3}$            |               |
| $\psi$   | Ammonia/nitrite inhibition         | 4.6                    | (µM N)$^{-1}$ |
| $w^p$    | Phytoplankton sinking rate         | Small: 0               | m d$^{-1}$    |
|           |                                    | Large: 0.5             |               |
| $f_{POM}$ | Decomposing material – proportion to POM | Small phytoplankton: 0.2 |               |
|           |                                    | Large phytoplankton: 0.5 |               |
|           |                                    | Small zooplankton: 0.2 |               |
|           |                                    | Large zooplankton: 0.7 |               |
| $g_{POM}$ | Grazed, unconsumed material – proportion to POM | 0.8 |               |
| $g_m$    | Zooplankton maximum grazing rate   | 0.5                    | d$^{-1}$      |
| parameter | description                                      | value               | units              |
|-----------|--------------------------------------------------|---------------------|--------------------|
| η         | Palatability of phytoplankton to zooplankton     |                     |                    |
|           | Palatability                                    | Large phyto, large zoo: 1 |
|           | to zooplankton                                  | Small phyto, small zoo: 1 |
|           |                                                 | Large phyto, small zoo: 0.2 |
|           |                                                 | Small phyto, large zoo: 0.2 |
|           |                                                 | Diatom, large zoo: 0.7 |
|           |                                                 | Diatom, small zoo: 0.14 |
| m²        | Zooplankton mortality rate                       | 0.033               | d⁻¹                |
| k²        | Phytoplankton half saturation                    | 0.025               | μM P               |
| rDOP      | DOM remineralization rate                        | 0.01                | d⁻¹                |
| rDON      |                                                  | 0.01                |                    |
| rDOFe     |                                                  | 0.01                |                    |
| rPOp      | POM remineralization rate                        | 0.02                | d⁻¹                |
| rPON      |                                                  | 0.02                |                    |
| rPOFe     |                                                  | 0.02                |                    |
| rPOS      |                                                  | 3.3×10⁻³            |                    |
| wPOM      | POM sinking rate                                 | 10                  | m d⁻¹              |
| ζNO2      | NH₄ to NO₂ oxidation rate                        | 2                   | d⁻¹                |
| ζNO3      | NO₂ to NO₃ oxidation rate                        | 0.1                 | d⁻¹                |
| αFe       | Iron solubility constant                          | 0.04                |                    |
| cFe       | Iron scavenging rate                             | 1.1×10⁻³            | d⁻¹                |
| βFe       | Ligand binding strength                           | 2×10⁵               | (μM Fe)⁻¹          |
| kₐ        | PAR attenuation coefficient                       | 0.04                | m⁻¹                |
| kphyto    | PAR attenuation coefficient from phytoplankton    | 0.64                | (μM P)⁻¹ m⁻¹       |
Supplementary References

1. Follows MJ, Dutkiewicz S, Grant S, Chisholm SW (2007) Emergent biogeography of microbial communities in a model ocean. Science 315: 1843-1846.
2. Dutkiewicz S, Follows MJ, Bragg JG (2009) Modeling the coupling of ocean ecology and biogeochemistry. Global Biogeochemical Cycles 23: GB4017.
3. Wunsch C, Heimbach P (2006) Practical global state estimation. Physica D 230: 197-208.
4. Marshall J, Hill C, Perelman L, Adcroft A (1997) Hydrostatic, quasi-hydrostatic, and non-hydrostatic ocean modeling. Journal of Geophysical Research 102: 5733-5752.
5. Roe PL (1985) Some contributions to the modelling of discontinuous flows. Proc. 1983 AMS-SIAM Summer Seminar on Large Scale Computing in Fluid Mechanics. Lectures in Applied Mathematics 22: 163-193.
6. Drake JW (1991) A constant rate of spontaneous mutation in DNA-based microbes. Proceedings of the National Academy of Sciences of the United States of America 88: 7160-7164.
7. Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998) Rates of spontaneous mutation. Genetics 148: 1667-1686.
8. Parekh P, Follows MJ, Boyle E (2004) Modeling the global ocean iron cycle. Global Biogeochemical Cycles 18: GB1002.
9. Aiken J, Rees N, Hooker S, Holligan P, Bale A, et al. (2000) The Atlantic Meridional Transect: overview and synthesis of data. Progress in Oceanography 45: 257-312.
Fig. S1. Growth of modeled phytoplankton as a function of (A) temperature (B) irradiance and (C) ambient phosphorus concentration. Here growth curves are illustrated for the 15 small phytoplankton types (red) and 18 large phytoplankton types (blue) from the simulation presented in the main text.
Fig. S2. Ambient availability of nitrogen. Annual average concentrations (μ M) of nitrate (A, B), nitrite (C, D) and ammonium (E, F) in surface waters of the global ocean (0-10 m; A, C, E), and along a transect in the Atlantic Ocean (AMT 13; B, D, F).
Fig. S3. The distribution of phytoplankton in surface waters of the model ocean during the tenth year of a model integration. Phytoplankton abundance (biomass, in units of μM P, 0 to 10 M, annual average) is plotted for (A) picophytoplankton and (B) large phytoplankton. The distribution of mutant picophytoplankton (biomass in units of μM P, 0 to 10 m, annual average) is plotted for (C) null mutants and (D) NO₃/NO₂ loss mutants.
Fig. S4. The distribution of phytoplankton in surface waters of the model ocean during the fifth year of a model integration, with the proportion of cell divisions resulting in mutations ($\Lambda_f$) set to $\Lambda_f = 10^{-6}$. Phytoplankton abundance (biomass, in units of $\mu$M P, 0 to 10 M, annual average) is plotted for (A) picophytoplankton and (B) large phytoplankton. The distribution of mutant picophytoplankton (biomass in units of $\mu$M P, 0 to 10 m, annual average) is plotted for (C) null mutants and (D) NO$_3$/NO$_2$ loss mutants.
Fig. S5. Ambient availability of nitrogen in an integration with $\zeta_{\text{NO}_2} = 0.1 \text{ d}^{-1}$ and $\zeta_{\text{NO}_3} = 0.033 \text{ d}^{-1}$. Annual average concentrations ($\mu M$) of nitrate (A, B), nitrite (C, D) and ammonium (E, F) in surface waters of the global ocean (0-10 m; A, C, E), and along a transect in the Atlantic Ocean (AMT 13; B, D, F).
Fig. S6. The distribution of phytoplankton in surface waters of the model ocean during the fifth year of a model integration with $\zeta_{NO_3} = 0.1 \text{ d}^{-1}$ and $\zeta_{NO_2} = 0.033 \text{ d}^{-1}$. Phytoplankton abundance (biomass, in units of $\mu$M P, 0 to 10 M, annual average) is plotted for (A) picophytoplankton and (B) large phytoplankton. The distribution of mutant picophytoplankton (biomass in units of $\mu$M P, 0 to 10 m, annual average) is plotted for (C) null mutants and (D) NO$_3$/NO$_2$ loss mutants.
Fig. S7. The distribution of phytoplankton in surface waters of the model ocean during the fifth year of nine additional model integrations (integrations ii-x) using different, randomly generated, phytoplankton communities. Phytoplankton abundance (biomass, in units of μM P, 0 to 10 M, annual average) is plotted for (A) picophytoplankton and (B) large phytoplankton. The distribution of mutant picophytoplankton (biomass in units of μM P, 0 to 10 m, annual average) is plotted for (C) null mutants and (D) NO$_3$/NO$_2$ loss mutants.
integration iii

A

B

C

D

integration iv

A

B

C

D

Fig. S7. continued.
Fig. S7. continued.
Fig. S7. continued.
integration ix

A

B

integration x

C

D

Fig. S7. continued.