Review II.

The manuscript describes a simple resource competition model designed to allow in- cluding genomic information in the model formulation. The simulations are performed in a chemostat-like setting intended to allow comparison with conditions in the sur- face ocean. Genomic information is implemented as trade-offs between resource-use abilities and maximum growth rate. The authors claim that their simulations are con- ceptually consistent with observations from marine plankton systems and that their approach should lead to better predictions about the response of plankton systems to environmental change.

General evaluation.

Including genomic information in a plankton model is a novel and intriguing approach. I also liked the authors' attempt to keep the model as simple as possible. However, the model formulation is fraught with severe problems and in my eyes just plain wrong. Also, the discussion of the results is overly optimistic and this appears to show a lack of sufficient critical distance of the authors to their work. Nevertheless, I expect that the main outcomes would remain the same if the model's problems were corrected. Thus, a thoroughly revised manuscript together with a reworked model could eventually be publishable.

My main concern here is with the model formulation itself. Problems with the text are given further below. While I very much agree with the approach of keeping the model simple, where the model does a good job, the formulations should nevertheless reflect the processes they are supposed to represent, and here the model fails.

To begin with, the different nitrogen and phosphorus compounds enter a phytoplankton cell on very different routes, the main difference in the present context being that ammonium can be taken up passively by diffusion through the cell wall, whereas all the other molecules (except N2) can only be acquired via active uptake by specialised (channel) enzymes at the cell surface. The difference is very important because all phytoplankton species can use ammonium, whereas some (Prochlorococcus, although Martiny et al. (2009), PNAS 106:10787, find that some Prochlorococcus types can use nitrate) may be unable to utilise nitrate and nitrate, but not ammonium, uptake is affected by iron limitation. Thus, the ability for ammonium use should be treated differently, as common for all species.

All phytoplankton in the model do use ammonium when the concentration is greater than 1 μM. In this case, uptake of ammonium is by passive diffusion as the reviewer states. When the concentration of ammonium drops below 1 μM cells without the gene for the amt transporter can no longer take up ammonium, whereas those that possess the amt cluster can. Uptake of nitrate and N2 are treated differently in that we assume uptake is diffusion limited at all concentrations and that uptake is inhibited at high ammonium concentrations — we appreciate that uptake of nitrate and N2 is via facilitated diffusion process, but we use a simplifying assumption that the process is diffusion limited. We

appreciate that nutrient uptake by cells is complex, with different processes and rate constants being required. However, in this proof-of-concept model, our aim was (as the reviewer appreciates) to generate a simple model that captures the role of the genes in regulating community structure and biogeochemistry, rather than concentrating on the kinetics of uptake. We agree that more realistic representations of uptake need to be considered for the model to be predictive. We are currently developing the model further to allow for more realistic kinetics and incorporating micronutrients such as iron. The goal is to compare model predictions with recent data collected from the Amazon River plume.

The next problem is the treatment of nutrient uptake. Firstly, the authors state that they describe nutrient uptake as a reaction-diffusion process (p. 822, l. 17), but then the equations only describe diffusion.

The reviewer is correct, we treat the process as diffusion limited process but modified so that the concentration gradient driving diffusion changes once the amt gene cluster has been up-regulated (in the case of ammonium, for example). Ideally we would use a full reaction-diffusion model but the reaction constants for the different transporters are currently known for only very few transporters. We have changed the description of this component of the model to read

"*Through the further assumption that nutrient elements and specialized enzymes immediately react as soon as they encounter each other at the cell surface (Atkin, 1998; Eigen and Hammes, 2006), the nutrient uptake is limited only by diffusion process and calculated as the nutrient diffusive flux to the surface of a cell of radius R (Jumars et al., 1993; KarpBoss et al., 1996) by*:"

Secondly, and more importantly, Eqs. (5) and (6) describe a sawtooth function if the corresponding gene is present. I am ready to accept neglecting the feedback between uptake and diffusion (constant surface concentration) as a simplification. However, the sawtooth function used here does not appear to make any sense. Worse, because uptake switches quickly between close to zero if Cj is slightly above C0,j, this introduces a positive feedback between nutrient concentration and uptake. Both this and the previous problem could be solved if the uptake was described like this (but there are probably many other ways to solve these problems): if the gene for a nutrient is present, potential uptake (diffusion) is described by Eq. (6), otherwise there is no uptake for this nutrient, and potential ammonium uptake is always described by Eq. (6). These first two concerns about the model formulation must be convincingly resolved in the revision if the manuscript is to be published.

We agree that the application of sawtooch function is overly simplified. We also realize that there is a positive feedback between nutrient concentration and uptake for phytoplankton that possesses "amt" genes. However, the positive feedback was somehow balanced by the penalty received by the phytoplankton possessing "amt" gene cluster (penalty: the reduction of their maximum growth rate (1%) no matter whether the "amt" gene is activated or not). The advantage for phytoplankton to possess "amt" gene is that

they can further assimilate NH4 according to Eq (6) when NH4 concentration drops below $C0_{,}j$.

To avoid the sawtooth function when C_j is very close to $C_{0,j}$, (Figure 1.a) mentioned by the reviewer we propose to apply a gradually shifting concentration gradient (smooth transition, Figure 1.b) to describe the nutrient flux for phytoplankton possessing "amt" gene. This modification will change the quickly switching two-phase diffusion process ([NH4] = 1 µM) into a smooth transition when [NH4] is lower than 2 µM.

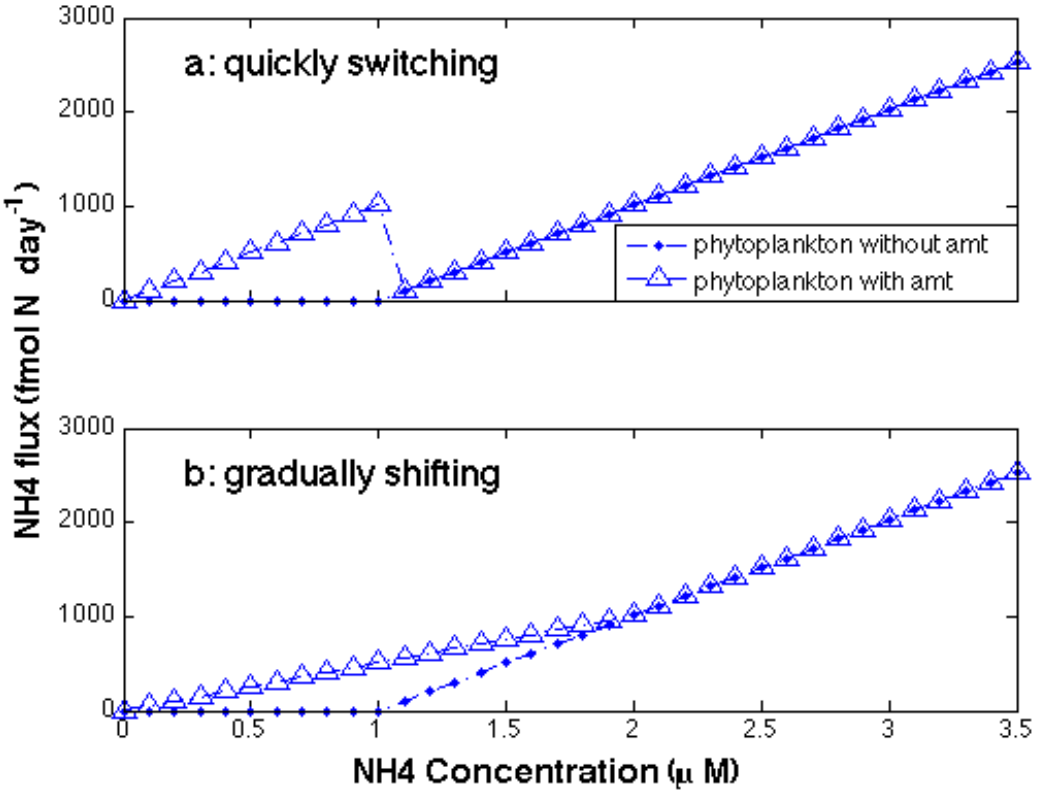

Figure 1: NH_4 flux reaching the cell surface (cell diameter: 1 um) for two types of phytoplankton (amt gene is absent and present).

We ran the model with the new uptake formulation and used the relative difference of total biomass in the last year to quantify the impact of the changes we made in the NH_4 uptake mechanism. In general relative change in biomass was less than 5% and the maximum was about 15% in the beginning of the year (Figure 2).

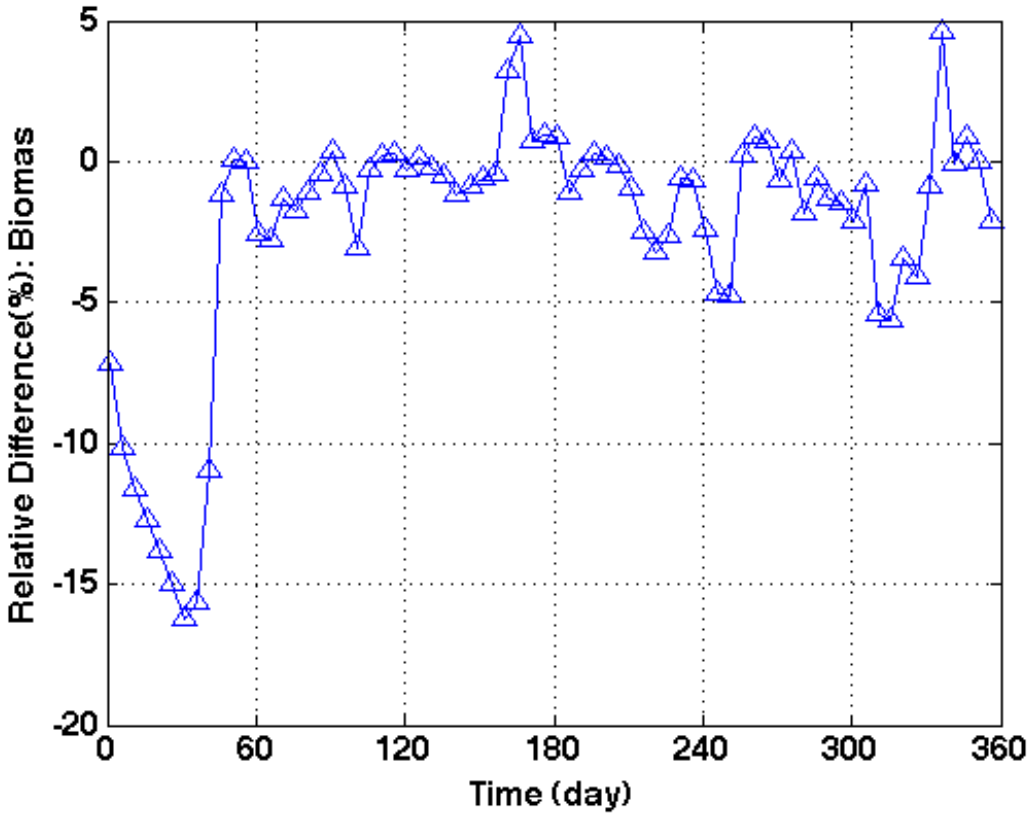

Figure 2: Relative difference of total biomass of phytoplankton from two simulations: quickly switching and gradually shifting concentration gradient.

I also find the gene regulation by extracellular nutrient concentrations somewhat questionable. Nutrient concentrations can and do change quite rapidly in the surface ocean, so that a gene regulation based thereon could often lead to a situation where the genes are constantly switched on and off as the nutrient concentration fluctuates around the critical concentration. I think the cell quota is a more useful (because integrated) quantity. Defining critical N and P cell quotas would also not make the model more complex in any way. It would, however, be less arbitrary than the choice of ammonium and phosphate for gene regulation, as it avoids the question: why not COP (for example)?

It should at least be noted in the text that phytoplankton cells usually generate only very weak gradients (C C0) and that C0 increases with increasing C in reality.

We have added the sentence "In reality, the critical extracellular concentrations (C_0) usually increases while the extracellular nutrient concentrations (C) increases. But we treat C_0 and C independent in our model." after line 14 on p. 822.

The reviewer raises a valid point: what controls gene regulation, internal cellular conditions (e.g. cell quota), external conditions (e.g. extracellular nutrient concentrations) or a mixture of both. As we mention in the discussion, regulation has been found to be controlled by multiple factors (e.g., the *nif* gene cluster is controlled by oxygen, iron,

temperature and irradiance (Gallon, 1981, 1992; Carpenter and Capone, 1992; Capone et al., 1997; Howard and Rees, 1996)). In addition, synthesis of phosphate-binding proteins in the cell wall of *Synechococcus* has been found to depend on extracellular concentrations of inorganic phosphate (Scanlan et al., 1993; 1997).

While the text is generally well written and clear, I find the placement of the results in general biogeochemistry and the discussion with respect to observations and application to the real ocean troubling. Also, it was not clear to me how the model was actually run. Are the figures representing a steady state? If so, the steady state does not appear to hold, e.g., in Fig. 5. If not, how were the initial conditions chosen and how long was the model run? This is important information for me to judge the validity and relevance of the model results.

The following paragraph is added on "p. 826, l. 4" to make the manuscript clearer. *"The initial values of nutrient concentrations are set to zero among all scenarios. The initial values for phytoplankton cell density and cell quota are 10 cells/l and average value of maximum and minimum cell quota respectively. The model run for 5 five years and after a quasi-steady state (i.e. repeating annual cycle) is established for all variables in the model, the results from the final year of the model runs are presented in the next section."*

Our application of the model to the different scenarios was meant to be illustrative and used very simple formulations to examine how the model communities changed under different regimes of nutrient input. We realize that our representation is highly simplistic, but this manuscript was meant to show that the modeling framework gave something that was broadly sensible. We are now working on applying this framework to a situation where we have omics data, biogeochemical data and community structure data.

All figures represent results from a quasi-steady state (i.e a repeating annual cycle).

At the end of the model description the authors state that this study was intended as a proof of concept (p. 11, l. 25), and as such I could live with the simplistic setup of the simulations. But then I would suggest to remove most of the references to biogeochemistry in the discussion and section 3.3.

We have significantly reduced the discussion about the biogeochemistry in the discussion. We have retained some discussion of this because we do believe that the model demonstrates a methodology for connecting omic information to changes in biogeochemistry. This aspect of the model is being actively pursued using recently collected data from the Amazon River Plume.

The comparison of the biodiversity patterns went rather wrong. For example, Pommier et al. (2007) reported higher diversity at higher latitudes whereas the present model predicts the opposite. But the authors made it sound as if the model produced a pattern similar to the observed (p. 16, ll. 13–15). Similarly, the observations of Treusch et al. (2009) indicate an inverse relationship between diversity and temperature, which is

contrary to the model prediction, but the text on p. 16, ll. 2–8 alleges a good agreement between model and observations.

Pommier et al. (2007) reported higher diversity at lower latitudes. Secondly, there is a typographical error in the manuscript: line 13 on page 830 should read "decrease from equator to pole". Our ocean shows an obvious gradient in the biodiversity of phytoplankton, from high biodiversity in the equator to low biodiversity in the polar region (Pommier et al., 2007, Furhman et al., 2008). In principle, our model results match the above observed patterns. Scenario I, which represents surface ocean at higher latitude, has lower biodiversity, but Scenario II as a representative of the surface ocean at lower latitude has higher biodiversity (p. 847, figure 2). The sentence (p. 847, ll. 13-15) might need to be altered as the following "This pattern is echoed from the model results: lower diversity in the surface ocean from higher latitude (Scenario I) and higher diversity from lower latitude (Scenario II)" to make it clear.

Regarding the observations of Treusch et al. (2009), it is true that the data from northwestern Sargasso Sea show an inverse relationship between diversity and temperature. However, the authors realized the apparent contradiction between their observation and the geographical pattern concerning the richness of marine microbial communities and made a comment "*However, in our data set, temperature and richness were inversely related, suggesting that factors other than temperature control richness on a vertical scale in the oceans.*" So far our model has not included temperature and it is hard to assess the relationship between temperature and diversity. However, the observation from Treusch et al. (2009) that eukaryotic phytoplankton and several cyanobacteria species alternate between spring and summer is exactly what we found in Scenario I. We feel it provides at least some indirect support concerning diversity results from our model although our model has not been applied to that particular region.

I think it would be best to rewrite the discussion and focus it on the proof-of-nature type of this study rather than questionable comparisons with observations and equally questionable biogeochemical implications.