Biogeosciences Discussions

Biogeosciences Discuss., author comment AC2
https://doi.org/10.5194/bg-2022-68-AC2, 2022


# Reply on RC2

Karen M. Brandenburg et al.

Author comment on "Physiological control on carbon isotope fractionation in marine phytoplankton" by Karen M. Brandenburg et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2022-68-AC2, 2022

The manuscript by Brandenburg et al. presents a compilation of stable carbon isotope fractionation data in phytoplankton experiments grown under various culture conditions (day length, nutrient availability, temperatures..) and teases out their contribution to the theoretically straightforward and expected demand/supply relationship on ep. The manuscript is an important contribution to the field, very well written and the data is presented nicely. I have therefore only a few comments and suggestions.

*We thank the reviewer for his/her kind words and will address the comments and suggestions below.*

General comments:

1) The authors should clarify the statistical approach. If I understood correctly, their linear models predicting ep had three factors, i.e. POC production/CO2, one influential condition (light, irradiance,.....), and species. While the influential condition and species factors have categorical or discrete factor levels, POC production/CO2 has not. Is that something the lmer function in R can handle? I was under the impression that all levels would need to be categorical or distinct (not a continuum without groups), as it is basically an ANOVA. Please clarify.

*The reviewer is correct that we used three predictor factors in our models (namely POC production/$CO_2$, one influential condition, and species). Lmer has no problem with fitting both continuous and discrete data as fixed predictor variables, as illustrated in the examples from Bates et al. (2015; https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf). What the reviewer maybe refers to is that it does not make sense to use a continuous variable for generating a random intercept in R. However, we used only the discrete variable "species" to provide a random intercept, and no continuous variables, so our model structure is correct. Here is also the R syntax for clarification: lmer(Ep ~ POCproduction/CO2 + [influential factor] + POCproduction/CO2: [influential factor] + (1|Species), data). To make this more clear, we now specify that we used a random intercept also in the text (L112).*

2) The authors have chosen to test POC production/CO2 as the main driving factor for ep

(please see also comment 1). From a pale-reconstruction perspective, that would require estimating two physiological parameters, i.e. POC per cell and instantaneous growth rate, to infer ep. What about the more simple growth rate/CO2 approach? The authors could test if they come to the same conclusions. I reckon they would but better to check.

*The reviewer is right that this more simple approach would be easier to apply from a paleo-reconstruction perspective. This is why we now include this analysis on instantaneous growth rate/CO2 in the supplementary ($\varepsilon_p$ versus $\mu_i$ /[$CO_2$], Fig. S2 and S3) and also mention it in the main text (L126-130; 159-161). While these data confirm our conclusion, these figures clearly stress the need to also make estimations for cellular POC contents in the paleo-domain, as especially for haptophytes this makes a big difference with regard to the explanatory power.*

3) Again, from a reconstruction perspective, the authors could calculate how much explanatory power a multiple linear regression approach would generate. Of course, some of the factors would not work as being categorical (unless a generalised linear model would be used instead), but some could be retained (e.g. light, temperature) or changed over (nutrient concentration, e.g. nitrate as being a proxy for the degree of limitation). That could be done group-specific, and looking at the simple linear regression presented in Figure 2, I could imagine that it would be quite a success.

*We thank the reviewer for his/her suggestion, and tested how much explanatory power we could generate with a multiple linear regression approach using different environmental variables (L114-116). As explained above, this approach can use both continuous and discrete data. We found that the inclusion of the light regime and whether there was nutrient limitation yielded highest explanatory power in all groups, and mention these findings now in the results (L131-133) and in the discussion (L156-157; 280-281).*

Specific comments:

1) L 241: either 'these systems' or 'this system'.

*Thanks for noting this mistake. We changed it accordingly.*