Response to reviewer 2

We thank the reviewer for insightful comments and suggestions for our manuscript. Please find attached all your comments and our responses (comments are in *italic*, our responses are in blue).

General Comments:

1. *This study addresses the very important topic of stoichiometric variability in marine phytoplankton. Understanding the magnitude and drivers of this variability as well as its taxonomic variation are essential for developing new and more accurate global biogeochemical models. The authors take a novel approach to this problem by performing a meta-analysis through which they calculate a sensitivity factor for major stoichiometries (N:C, P:C, and N:P) in response to a suite of environmental drivers. The goal of such a quantitative approach - to estimate the group-specific response of these stoichiometries to expected changes in ocean conditions - is laudable.*

   Thank you for these encouraging comments.

2. *However, there are several major flaws in how this approach is applied and how studies are selected and screened for this meta-analysis that would need to be addressed for this to be published in Biogeosciences. Additionally, these major flaws in approach receive little or no discussion throughout the manuscript.*

   The major issues we are going to address in the revision are: 1) study selection criteria (comment #19), 2) S-factor calculation (comment #20), 3) and the overall discussion of the methodological limitations. Please refer to the responses to specific comments for more detail. We note that what are referred to as major flaws (i.e., application of the power law metric to studies with 3 environmental levels) are deliberate and justifiable choices we made given our motivation to develop possibly nonlinear stoichiometric formulations for use in global biogeochemical models. As discussed below, there are tradeoffs in selecting studies with 2 or 3 levels. But in our revision, we will heed the suggestion of the reviewer and consider the additional the selection criteria.

3. *The authors present their approach to estimating a response to an environmental condition as more nuanced and informative than simply calculating a response between two end points or experimental treatments. While those simplistic, past approaches have numerous limitations, they were generally acceptable for meta-analyses due to two major challenges: 1) the high variability in experiment conditions of individual studies; and 2) the fact that some environmental drivers may produce linear or at least monotonic responses within a range of natural variability (e.g. the response to nutrient availability), while other drivers produce responses that are distinctly antitonic (e.g. temperature and irradiance). Essentially the authors have suggested a more complex metric for such meta-analyses without addressing these two major challenges. As a result, ambient nutrient concentrations are treated as a measure of a study species' nutrient status that is comparable across different experiment types (semi-continuous batch vs. chemostat), which is inappropriate for several reasons (addressed below in my specific comments). The flaws of this approach are not discussed in the manuscript and the approach is used to make the study's strongest conclusion, that diatom P:C and N:C are particularly sensitive to N and P availability. It should be added that this result is based on meta-analysis of only four studies, one of which was on a dinoflagellate and incorrectly categorized.*

Thank you for this suggestion. As suggested, we will conduct meta-analysis using two end points which should resolve the two major challenges mentioned here: 1) the high variability in experiment conditions of individual studies; and 2) the fact that some environmental drivers may produce linear or at least monotonic responses.

4. *This approach also results in deeming a given stoichiometry as sensitive to a driver like irradiance or temperature if that stoichiometry has a monotonic response to these drivers. Considering that the responses of phytoplankton to light and temperature are distinctly non-linear and antitonic (usually displaying a clear central optimum), this approach seems very flawed.*

   As mentioned in the previous point (#3), we will conduct two point meta-analysis which should satisfactorily resolve the issues mentioned here.

5. *Considering its novelty and potential value, the approach used by the authors should not be discarded, but refinement and far more discussion of its limitations would be necessary to present it in a manuscript.*

   We will keep our s-factor (fractional change in P(N):C over fractional change in independent variable) as our effect size but will make a refinement to data selection by choosing two end points. We will also discuss caveats and limitations more extensively in the revised manuscript.

6. *The computational needs of the sensitivity factor that the authors use (requiring experiments where the response to at least 3 levels of an environmental driver were measured) also seems to have resulted in a meta-analysis of a somewhat limited number of studies. While this criteria is strict, there is no study selection criteria mentioned that address the many other confounding factors that could differ among studies and little or no discussion of such factors.*

   There is a tradeoff between using two points (more studies but linear response) and three points (fewer studies but possibly nonlinear response). We were focused on the latter but acknowledge the merits of using just two points as well. As suggested, we will add a new selection criteria to include studies with 2 levels, which should increase in the number of studies available for meta-analysis.

7. *Along with this lack of evaluation of the original studies used in the meta-analysis, there is also little comparison of the results of this work to the findings of several other narrative reviews and quantitative meta-analyses of phytoplankton stoichiometry, most of which considered a larger number of original studies. These past studies are generally just mentioned for comparison of approaches, but not their results are not critically evaluated in light of the authors' contributions to this topic.*

   Thank you for this suggestion. We will more critically compare and discuss our results with those from previous meta-analysis studies.

8. *As mentioned above, there also seems to be several studies that were incorrectly categorized, with non-diatom species appearing to be grouped with diatoms in the group-specific meta-analyses.*

Thank you for picking out our errors. We will update our database and will conduct meta-analysis with corrected classification.

Specific comments:

Abstract:

9. *Line 18-20: It seems overly simplistic to imply that the temperature response of cyanobacteria is responsible for global P:C patterns without acknowledging the effect of macronutrient availability, which you have also shown to have a strong effect on P:C and N:C. The global patterns in C:N:P (lower P:C and N:C in subtropics, higher in subpolar and upwelling regions) has also been attributed to macronutrient availability and phytoplankton biogeography with the relative impact of all three drivers being a rich and contentious area of research. Linking your findings to this on-going area of study should either be excluded from the abstract or addressed in a more complete fashion by noting that the macronutrient sensitivity of diatom C:N:P and the temperature sensitivity of cyanobacteria C:N:P you observe are both helpful in explaining the persistent global patterns in C:N:P.*

   Our intent was that temperature is possibly an important factor along with other factors such as macronutrients in explaining the subtropical cyanobacteria C:N:P. As pointed out by the reviewer, we neglected to note the other factors. We will thus modify the sentence to read: "Along with other oceanographic conditions of the subtropical gyres (e.g., low macronutrient availability, low light attenuation), elevated temperature may explain why P:C is consistently low in the cyanobacteria-dominated subtropical oceans."

Introduction:

10. *Line 43-45: This sentence should be supported by citations. It is not clear which of the citations in the previous sentence (if any) are the sources for this information.*

    Information comes from the review paper by (Moreno and Martiny, 2018). We will cite this paper in the revised manuscript.

11. *Line 53-55: This statement is vague and detailed specific support for this should be given.*

    The main message of this sentence is that environmentally induced trait change is variable because it is driven by both plasticity (acclimation) and adaptation, which differs amongst species (Collins et al., 2020; Ward et al., 2019). We will rephrase this sentence to make the meaning clearer.

    It's worth clarifying why previous studies have not yielded a broader understanding of how phytoplankton C:N:P varies across taxa and environmental conditions (and thus justifying your meta-analysis).

    As the field of marine ecological stoichiometry itself is new (i.e., transition from the traditional Redfieldian view), fundamentally, there is not yet many studies that give **broad yet quantitative** views on how marine environmental factors affect plankton C:N:P. Our main motivation for this work therefore was to build a database that could be used to calibrate power-law based flexible C:N:P model of phytoplankton (Tanioka and Matsumoto, 2017) that can easily be incorporated into marine biogeochemical models. We also aim to build on previous phytoplankton cellular models (e.g., Pahlow and Oschlies, 2009) that are usually calibrated with very few selected studies from 20-30 years ago (e.g., Laws and Bannister, 1980).

Also, the inherent genetic differences among taxa don't simply correspond to differences in environmental responses, they correspond to inherent differences in steady-state C:N:P under ideal conditions among major phytoplankton groups (Quigg et al. 2003; Garcia et al. 2018) that likely reflect basic differences in cellular structure and size (Finkel et al. 2016a; Finkel et al. 2016b). See references below.

- *Quigg, A., Finkel, Z. V., Irwin, A. J., Rosenthal, Y., Ho, T. Y., Reinfelder, J. R., ... & Falkowski, P. G. (2003). The evolutionary inheritance of elemental stoichiometry in marine phytoplankton. Nature, 425(6955), 291.*
- *Garcia, N. S., Sexton, J., Riggins, T., Brown, J., Lomas, M. W., & Martiny, A. C. (2018). High variability in cellular stoichiometry of carbon, nitrogen, and phosphorus within classes of marine eukaryotic phytoplankton under sufficient nutrient conditions. Frontiers in microbiology, 9, 543.*
- *Finkel, Z. V., Follows, M. J., Liefer, J. D., Brown, C. M., Benner, I., & Irwin, A.J. (2016a). Phylogenetic diversity in the macromolecular composition of microalgae. PLoS One, 11(5), e0155977.*
- *Finkel, Z. V., Follows, M. J., & Irwin, A. J. (2016b). Size-scaling of macromolecules and chemical energy content in the eukaryotic microalgae. Journal of Plankton Research, 38(5), 1151-1162.*

Thank you for this insight. We will touch on the fact that difference in cellular structure and size also leads to variability in C:N:P at steady-state given. We note however that our study is more concerned with C:N:P under transient condition (i.e., future climate change scenario).

12. *Line 55-58: In addition to the point made in the previous comment, there are many reasons why it is hard to draw consensus from the various studies of phytoplankton C:N:P, but an inconsistency of statistical analyses seems like one of the least compelling of these reasons. What about the differences in how experimental treatments are applied, particularly for macronutrient limitation (e.g. steady-state vs batch cultures and differences in the duration of nutrient stress)? What about confounding experimental conditions (e.g. bacterial contamination, low CO2 availability/high pH in dense batch cultures)? Or more simply, the fact that many studies only measure one or two of the three major elements and few measure the biochemical components that determine elemental quotas. These are all factors that make understanding how phytoplankton C:N:P varies across taxa and conditions difficult when using existing literature and seem much more important than the selection of statistical analyses. Not mentioning these factors in the introduction and, more importantly, in the methods section when considering selection criteria is a major omission in this paper.*

    Thank you for this suggestion. Although it is inherently impossible to consider all of the factors mentioned above, we will add in our database some additional information. These include growth mode, axenic/non-axenic nature of the culture, salinity, culture medium (natural or artificial seawater), acclimation (# of generations), and optimality (temperature and salinity). We will mention in the introduction and in the methods section that there are considerably more factors that affect C:N:P than the five factors that we considered in our study.

13. *Line 59-67: This paragraph seems mostly unnecessary. The value of a quantitative meta-analysis is self-evident for the audience and can be stated by a simple statement of the goal of this work later in the introduction. Shortening this also leaves more room for more helpful introductory information regarding the causes of phytoplankton C:N:P variability or the factors that make this meta-analysis challenging (see previous two comments).*

    We agree that this paragraph is a little lengthy. In the revised manuscript, we will shorten this paragraph to make it more succinct.

14. *Line 69-72: While previous meta-analyses that focus on only one environmental driver are indeed limited, these studies must still have some value or informative conclusions. This introduction contains no mention of the actual findings or major conclusions of these previous studies. Addressing the findings and relative value of previous, similar work should be a fundamental part of any introduction. Again, addressing this omission seems more helpful than the paragraph explaining why meta-analyses are valuable.*

    Thank you for this suggestion. In the revised manuscript, we will address the findings and relative values of previous meta-analysis studies in more depth here (possibly as a table).

15. *Line 76: The sentence contained here is incomplete and seems like a typo.*

    We will connect lines 76 and 77 with a word "and" to make it a complete sentence.

Methods:

16. *Line 93: For readers who might not be familiar with search operators, you should define "TS" as in its first usage as a field tag for "topic" (or some other appropriate definition).*

    Thank you for this suggestion. We will define TS as a field tag for topic in the modified manuscript.

17. *Line 94-100: As with the previous comment, it would be good to explain the meaning of "*" as a wildcard search operator.*

    We will describe " * " as a wild card search in the modified manuscript.

18. *Line 94-100: These descriptions of search terms are not accessible when listed in a paragraph. This information should be placed in a table.*

    Thank you for this suggestion. We will place these search terms in a table either in the main text or in the supplementary information.

19. *Study Selection Criteria: The way in which studies were selected for the meta-analysis and the lack of analysis or discussion of the confounding factors that various studies present are where some of my strongest critiques lie. I've presented these critiques as a list below:*

a) *Limitation of 3 experimental levels: The value of setting the study selection criteria to 3 experimental levels for each environmental factor of interest seems overstated. The terms X and Y (the fractional response and fractional change in conditions) could be calculated with just two* *experimental levels for each experimental unit. Granted this does not allow the error associated with a linear regression of 3 X and Y values to be used or for a non-linear response to be detected, but I would question the value of such an error term or description of a non-linear response that was based on a linear regression of only three points. Give the limits of this additional explanatory power, this criterion seems unnecessarily limiting (see next points).*

It is true that our effect size of meta-analysis (i.e., s-factor; the fractional response over fractional change in conditions) can in theory be calculated with just two points. There is a tradeoff between using two points (more studies but linear response) and three points. Given that one of our main motivations was to incorporate a possibly nonlinear stoichiometric response in a global ocean model, the three point metric was originally selected. However, we acknowledge the merits of using just two points as suggested. In the revised manuscript, we will thus newly provide results of meta-analysis using 2 levels in addition to 3.

b) *Excluding valuable studies: Having only two levels of an experimental factor is not the major failing of most studies of phytoplankton elemental composition. There are many studies that I would deem of high quality that would have made excellent additions to this metaanalysis that only use two experimental levels for a given type of nutrient stress (e.g. Bertilsson et al. 2003; Fu et al. 2007 J. Phycol.). Considering this, the criteria of 3 levels unnecessarily diminishes the data density of the meta-analysis. Again, perhaps a better explanation of the selection criteria and meta-analysis calculations is needed if I am mistaken. It seems like a meta-analysis that utilizes a greater number of individual experimental units by including experiment with only two levels would have much greater breadth and power.*

As mentioned, in the previous point #19a), there is a tradeoff but that we will carry out new analysis using two levels. Since we already have the full list of studies that needed to be included, this process of regathering and reprocessing data should not be too time consuming and achievable.

c) *Not addressing major confounding factors: The more important failing in studies of phytoplankton C:N:P is the lack of consistent experimental conditions or poorly described conditions. Many studies do not offer verification that some desired growth state was successfully applied, particularly in the case of N or P stress. For example, many studies do not describe the growth rate at a given experimental level of a limiting nutrient. How an author defines N or P starvation or to what extent these conditions were applied (e.g. were they applied until growth ceased, or just until growth slowed) can greatly affect the observed response. Additionally, many nutrient starvation experiments are done in dense batch cultures where the additional stressors of light limitation, high pH, and low carbon availability arise as cultures increase in density and coincide with the onset of nutrient starvation. I mention this not to say that the authors should have determined such confounding factors in every study (in many cases, experimental conditions are not described well enough to do this), but rather to point out that such factors are not addressed at all in the selection criteria. In other words, a poorly executed study that did not fully apply nutrient starvation (even across 3 levels) would be included, but a well-described and well-executed experiment across only 2 levels (e.g. nutrient replete vs. nutrient starved) would be excluded. Again, this gets to the point that basing s-factors on a linear regression of 3 or more experimental levels has applied a major constraint on the meta-*

analysis and the value of this constraint is unclear, yet other major confounding factors are not addressed in the selection criteria.

Since all the experiments are "unique" in the sense that they do not have all the same controlled variables, it is pragmatically impossible to take into the account of all the factors (e.g., pH, carbon availability etc.). That being said, we are currently updating our database to include these extra factors such as salinity. Again, as mentioned above in #19a and #19b, we will do additional analysis with a modified selection criteria for nutrient experiments with two levels. For elucidating a linear response, the new analysis will provide more data points and should make our analysis more robust with greater statistical power.

20. *S-factor Calculation for Meta-Analysis: My other major critiques pertain to how s-factor was calculated, particularly for macronutrient stress experiments. Again, I've presented these critiques as a list below:*

a) *How was standard error propagated when calculating s-factors? Does the error reflect both the error associated with each P:C or N:C measurement and the error associated with the regression of X and Y for each experimental unit? How the error associated with the original measurements was accounted for and propagated must be described (if this was done).*

We did not account for errors associated with original measurements but only the error associated with the regression. In the revised manuscript using two experimental levels, we will account for errors associated with the original measurements as well. Error propagation will be described in more depth in the revised manuscript.

b) *With respect to the error associated with the weighted mean s-factors, I realize that the metafor R package is used for this calculation, but some general description of how this package calculates error should be provided. In other words, you should be explicit about what the error bars shown in the figures actually mean.*

The mean s-factor is weighted with respect to the standard error from each individual experimental unit. We will describe this in more depth in the revised manuscript.

c) *It is not at all clear how the fractional change in nitrate or phosphate stress was calculated. Was this simply based on the ambient nitrate or phosphate concentration reported for each experimental level? If so, how can the level of N stress be determined if ammonium or nitrate are not accounted for?*

For chemostat or semi-continuous batch experiments, **inflow phosphate (or nitrate) concentrations** at each dilution are manipulated and we calculate fractional change based on that. For example, a study by Leonardos and Geider (2004) used 5 levels of inflow phosphate concentration (20, 6.7, 3.3, 2.2, and 1.1 μmol/L) while keeping inflow nitrate concentration constant at 100 μmol/L.

For batch experiments, we calculate fractional change based on the **initial phosphate (or nitrate) concentrations in fresh media** at the start of the experiment**.**

In either case, we are not using ambient concentration so that we do not need to take into the account of ammonium or nitrite produced through organic matter remineralization.

d) *Batch, semi-continuous batch, and continuous chemostat experiments were used in the meta-analysis of macronutrient response. I do not understand how a simple measurement of ambient inorganic nutrient concentration can be used to determine experimental levels of N or P stress across these different experiment types. Even between a semi-continuous batch experiment where authors claim cultures are in balanced growth and a chemostat experiment, the measured nutrient concentrations or nutrient concentrations in fresh or inflow media mean different things with respect to extent of nutrient stress. In other words, moving from a nitrate concentration of 1.0 to 0.2 would mean very different things depending on whether they are in semi-continous or continuous mode, the concentration of other forms of dissolved inorganic nitrogen (ammonium, nitrite) or what the concentration of other potentially limiting nutrients are. The extent of nutrient stress cannot be compared between these different growth modes based on dissolved nutrient concentrations alone. Some would argue the extent of nutrient stress cannot be compared across these growth modes at all, and thus they can't be pooled into one type of meta-analysis. Again, a strict criterion of 3 experimental levels has been applied in this meta-analysis to serve a computational need, but other major confounding factors have been ignored. Additionally, these 3 experimental levels have been used to calculate a fractional change in conditions that does not have a consistent meaning across experiment types. The only way to deal with these problems while still using the current meta-analysis approach (sfactors, based on 3 experimental levels) would be to separate experimental units based on their growth mode and apply a more rigorous means of determining experimental levels of nutrient stress (i.e. growth rate) in the semi-continuous and continuous growth experiments.*

Ideally speaking, it would be best to use chemostat experiment for assessing the macronutrient availability on C:N:P because this growth mode can achieve constant growth rate. However, there are two issues: 1) there are significantly fewer chemostat studies compared to batch/semi-continuous, 2) different chemostat studies use different dilution rates. Therefore, we had to decide to put all the three form of experiments (batch, semi-continuous, chemostat) together in our analysis. In the revised manuscript with more data available, we should be able to analyze three growth mode separately.

Firstly, as we are going to increase the number of studies by including 2 level experiments, we should have enough studies to separate out "P-limited → P-replete studies" (for calculating s-factor with respect to change in PO4) from "N-limited → N-replete studies" (for calculating s-factor with respect to change in NO3) in the case of batch and semi-continuous batch experiments.

For chemostat experiments, where dilution rates and inflow concentrations are both manipulated, we will **select the case with the highest dilution rate** so that the effect of growth rate on C:N:P is minimalized (Klausmeier et al., 2004). Although different studies use different chemostat dilution rates we think this would be the most viable option given the fact that the number of studies using chemostat is small.

e) *Similar problems with the s-factor calculation of using a linear fractional change in growth conditions also apply to temperature and irradiance. Such a formulation ignores the growth optimum of a particular species or strain and thus treats an extremely non-linear response as* *something that can be compared across studies and taxa with a simple linear relationship. Consider a scenario where an experiment measured N:C at four temperatures in a species with growth optimum of 22C and had the following result: 15C = 0.14, 20C=0.154, 25C=0.156, and 30C=0.14. An s-factor calculated as a linear regression of X and Y from this experiment would be very small in magnitude and imply that this species is insensitive to temperature changes, when in fact these are actually large changes in N:C with respect to global conditions and what is generally observed in temperature responses. This experiment also shows that N:C declines at supraoptimal temperatures, the most relevant result with respect to climate change scenarios, but something that would be missed by the s-factor. In other words, the s-factor is a poor metric for a biological variable that does not have a monotonic response to some condition as is the case with light and temperature responses.*

A prior meta-analysis study by Yvon-Durocher et al. (2015) did find a linear relationship between temperature and C:P, hence the assumption of monotonic relationship is justifiable. In addition, it is not possible to deduce that P:C and/or N:C should be unimodal simply because the growth-temperature curve is unimodal. In the revised manuscript with 2 level experiments, as a tradeoff from 3 level experiments, we would not be able to depict such unimodal relationship nor get an information on optimal growth temperature. However, our new, 2 point linear approach should be justifiable at least as a first approximation.

Also, depending on the light or temperature levels selected in a given experiment with respect to the study species growth optimum, a fractional change in these conditions means very different things and are not directly comparable.

We agree that results are going to be more comparable if all the studies are compared at optimal temperature or at the temperature when the growth rate is highest amongst others. We will update our database so that this selection criteria is applied across all studies.

21. *Line 147-149: the symbol used to denote dissolved iron should be a mathematical prime symbol(′), not an apostrophe or single quotation mark.*

    Thank you for pointing this out. We will use the correct symbol.

22. *Line 150-151: "only selected experiments where NO3 concentrations were kept constant." This is either a writing error or a misunderstanding of the experiments selected. The non-limiting macronutrient was not kept constant in many of the experiments selected and this is rarely achieved even in chemostat experiments (see the nutrient concentrations described in Leonardos and Geider 2004 for example). Again, the selection criteria and calculation of fractional change for macronutrient stress experiments is either poorly described, problematic, or both.*

    For chemostat experiments (e.g., Leonardos and Geider, 2004) NO3 concentration is referring to the "inflow" NO3 concentration and this was kept constant at 100 μmol/L for experiment conducted by Leonardos and Geider (2004).

    For **chemostat experiments and semi-continuous experiments, we used inflow concentration** and for **batch, we use the initial nutrient concentration**. We will make this point clearer in the revised manuscript.

Results:

23. *Figures 2 – 5: The structure of the figures seems likely to confuse readers. Tables are often arranged such that inclusive categories are listed above subcategories. When first looking at figure 1, I see "Diatoms" in bold and then genus names for various eukaryotes below it and was disoriented for a moment. The figures may be more intuitive if you listed an inclusive group (e.g. "Diatoms") and then listed taxa within that group immediately below it with an indentation.*

    Thank you for this suggestion. We will come up with a way to make figures more intuitive.

    Also, why are the figures arranged as nutrient limitation (Fig. 2), Light (Fig. 3), Temperature (Fig. 4), nutrient-limitation (Fig. 5). I understand if this was done because there is very limited data for Iron limitation, but a more logical arrange of the figures would be better for comparison.

    For better comparison, we will place iron after nutrient limitation.

24. *There also appears to be a few taxonomic assignment errors in the meta-analysis based on the figures. Alexandrium minutum (a dinoflagellate) is listed among the diatoms in the Figure 2, Chlorella sp. (a chlorophyte) is listed among the diatoms in Figure 3, and Phaeocystis (a haptophyte) is listed among the diatoms in Figure 4. Does this error extend to the meta-analysis or was this an error in figure preparation?*

    Thank you for pointing this point. These are indeed misclassification and analyses will be redone with the correct classification.

25. *There also seems to be errors or inconsistencies in how studies were characterized with respect to N or P limitation. For example, why is Leonardos and Geider 2004 only listed among "Phosphate" experiments. This is a chemostat study that spanned both N-limited, balanced growth and P-limited balanced growth and thus could also be included with the "Nitrate" and "Nitrate/Phosphate" meta-analyses. The fact that these chemostats were controlled by manipulating inflow phosphate is irrelevant and does not make them simply "phosphate" experiments. Neither nitrate or phosphate values were constant across experimental levels in this experiment, what matters is that these were chemostats where inflow N:P was manipulated. I did not closely examine every study in the meta-analysis, but I am concerned that other such inconsistencies are present.*

    The reason Leonardo and Geider (2004) study was included in "Phosphate" experiments was because inflow phosphate concentration was manipulated while inflow nitrate concentration was kept constant. It was categorized "Phosphate" not because the experiment is P-limited but purely based on the fact that inflow phosphate was the only variable that was manipulated. Similarly, inflow (for chemostat/semi-continuous) or initial (for batch) phosphate concentration is kept constant for "Nitrate" experiments. As for "Nitrate/Phosphate" experiments, both inflow phosphate and nitrate were manipulated. However, since this last "Nitrate/Phosphate" classification is redundant and not clear, we will remove this category in the new analysis.

Discussion:

26. *Line 230: the word "the" before "chemical" should be removed*

Will be changed as suggested.

27. *Line 241: "making of … reductase". Do you mean "reductant" (i.e. NADPH) rather than reductase (an enzyme)?*

    Thank you for pointing out our mistake. We will change this to reductants.

28. *Line 243-246: These are specific statements that should be supported with references.*

    We decided to remove this sentence as it is not relevant to the rest of this paper.

29. *Line 237-238 and other parts of paragraph: There seems to be a misunderstanding of the term "balanced growth". A natural population or culture can be both nutrient-limited and in steady-state, balanced growth if the limiting nutrient is supplied at a consistent rate. Despite the various factors that limit phytoplankton growth and the natural conditions that represent clearly unbalanced growth (spring blooms), a balanced growth model of natural populations (the "steady-state ocean") is still very relevant for the vast subtropical oceans where consistent and actively growing populations occur amidst apparent chronic nutrient limitation.*

    Thank you for pointing this out. We will modify this paragraph regarding "balanced growth" as suggested.

30. *Line 282: This should be corrected to "we observe a consistent trend" or "we observe consistent trends"*

    Will change to "we observed a consistent trend".

31. *Line 296: I think "… the level…" should be changed to "…the same level…". If this is not just a typo, than this sentence should rewritten and clarified*

    This was a typo. It will be corrected to "… the same level…"

32. *Line 298: the phrase "number of…" or "abundance of…" should be placed before "…ribosomes"*

    The phrase "number of …" will be added.

33. *Line 300: revise to "… in a cell, resulting in … " or "… in a cell and result in…"*

    Will be changed to "… in a cell, resulting in …"

34. *Line 309: The Garcia reference is not appropriate here. References that actually describe this mechanism should be cited:*

    - *Dortch, Q., Clayton, J. R., Thoresen, S. S., & Ahmed, S. I. (1984). Species differences in accumulation of nitrogen pools in phytoplankton. Marine Biology, 81(3), 237-250.*
    - *Lourenço, S. O., Barbarino, E., Lavín, P. L., Lanfer Marquez, U. M., & Aidar, E. (2004). Distribution of intracellular nitrogen in marine microalgae: calculation of new nitrogen-to-protein conversion factors. European Journal of Phycology, 39(1), 17-32.*

- *Grover, J. P. (1991). Resource competition in a variable environment: phytoplankton growing according to the variable-internal-stores model. The American Naturalist, 138(4), 811-835.*
- *Tozzi, S., Schofield, O., & Falkowski, P. (2004). Historical climate change and ocean turbulence as selective agents for two key phytoplankton functional groups. Marine Ecology Progress Series, 274, 123-132.*
- *Talmy, D., Blackford, J., Hardman-Mountford, N. J., Polimene, L., Follows, M. J., & Geider, R. J. (2014). Flexible C: N ratio enhances metabolism of large phytoplankton when resource supply is intermittent.*

Thank you for these references. We will cite these papers instead.

35. *Line 320: The word "and" should be inserted after "significantly"*

    Will be changed as suggested.

36. *Line 328: "Large stoichiometry sensitivity…" should be changed to "The larger stoichiometric sensitivity…" or "The larger sensitivity of P:C…"*

    Will be changed to "The larger stoichiometric sensitivity…"

37. *Line 339-340: "Excess carbon…" – this sentence is a non-sequitur and should be modified to connect with the topic of irradiance effects.*

    Will be changed to "Excess carbon that is fixed under high irradiance condition is …"

38. *Line 349-350: This statement may not be true and should be supported by some reference. The light harvesting apparatus will still be expected to be down-regulated under N-replete conditions in order to avoid oxidative stress and photodamage and also to maximize growth rate and N allocation.*

    We will remove lines 349-350. It is true that down-regulation of light harvesting apparatus is expected under nutrient-replete condition as well (Geider et al., 1996; Laws and Bannister, 1980) and our original description here was not correct.

39. *Line 351-355: Amidst all these explanations of why irradiance has little effect on C:N:P, there is a fundamental explanation that has not been addressed. Although N-content may be expected to decline as irradiance increases due to a down regulation of the light harvesting apparatus, one could also expect an increase in N allocation to other cellular functions including nutrient uptake, biosynthesis, and repair of the light harvesting apparatus in order to match an increase in C-fixation. This shift in N allocation from light harvesting content to nutrient acquisition and biosynthesis is essential to an increase in growth rate with irradiance and could be expected at light levels that are below some photoinhibitory level. I don't know if this reallocation of N is sufficient to offset the expected decline in N content due down regulation of the light harvesting apparatus, but at least this is based on fundamental biological processes rather than critiques of experimental conditions that are not followed by any details or substantiation.*

This is a very good suggestion. As suggested, we will include these factors to illustrate why irradiance could have muted effect on C:N:P.

1) Increase in N allocation to nutrient uptake apparatus following the "chain model" concept (Ågren, 2004; Pahlow and Oschlies, 2009).
2) Increased N requirement for Rubisco at high irradiance offsetting reduction in N-content due to a down regulation of the light harvesting apparatus (Li et al., 2015).
3) Increased demand of proteins (e.g., D1 protein) for the repair of light harvesting apparatus at high irradiance (Demmig-Adams and Adams, 1992; Li et al., 2015; Talmy et al., 2013).

We agree that these fundamental biological processes are probably more important than critiques of experimental conditions and the lines 351-355 will be modified.

40. *Line 351-364: It seems odd that the variation in experimental conditions is invoked here to explain the limited the effect of irradiance on C:N:P, but this was not addressed with respect to macronutrient limitations. It seems logically inconsistent to note these methodological issues only when a clear effect is not found.*

    It is true that the variation in experimental conditions (batch vs semi-continuous batch vs continuous) should be addressed with respect to macronutrient limitations as well. We will move this discussion to an earlier section either in the Methods or in the early part of Discussion.

41. *Line 359-360 and 372-373: "We speculate…" – Aren't these concepts easy to verify or discuss further considering the small number of studies used in the meta-analyses rather than just speculate? Were the experiments used for the irradiance meta-analysis diel or continuous light. What proportion were continuous light?*

    The experiments used in our meta-analysis was a mixture of diel and continuous light. Of 76 experimental units, 10 were using continuous light and 66 were using diel light. As the reviewer suggested, we will verify if the length of light vs dark hours does lead to statistically significant difference in s-factors. If it does not, we will remove such speculation from the manuscript.

    Were these experiments mostly done at optimal temperature?

    We had not previously gathered information on optimal temperature but we are currently gathering such information from the studies included.

    Also, I thought your selection criteria examined studies where irradiance was manipulated, but nutrient status was not. How can nutrient status then be invoked as a possible confounding factor?

    For a given unique experimental unit the nutrient condition was constant. However, some studies from the same paper were conducted both under high nutrient and low nutrient conditions (e.g., "*Thalassiosira_psedonana_HN* (Li17) and *Thalassiosira_psedonana_LN* (Li17) are from high-nutrient and low-nutrient conditions, respectively). In such situations nutrient status can be invoked as a confounding factor for explaining variability in s-factor.

It seems more reasonable and conservative to assume that irradiance simply does not have strong effect on P:C?

As for P:C, we agree that is more reasonable to assume that irradiance does not have a strong effect on P:C. We will rewrite the last sentence beginning with "we speculate that…".

42. *Line 420-422: The time range of selected studies seems like a very weak argument. Wouldn't the selection criteria for the studies used in each meta-analysis also have a strong effect on the result. Also couldn't you simply split your analysis between these time ranges to see how it compares to the Yvon-Durocher study? This seems like another speculation that could be very easily examined.*

    We will be carrying out analysis as suggested by splitting out analysis ranges (pre-1996 and post-1996) to test whether time range is indeed the only reason that leads to a different conclusion.

43. *Line 432-434: The sentence here is incomplete or a fragment and should be revised.*

    We will remove the word "that" early in the sentence to make the meaning clearer.

44. *Line 436: This seems like an erroneous assumption. Couldn't a non-significant effect of iron on stoichiometry also be due to variable and contrasting effects of iron on cellular C and N or reflect the small number of studies examined!?*

    It is possible that non-significant effect of iron was indeed due to a small sample size of meta-analysis. As described earlier, we will expand our database by including studies that only had 2 points to validate/nullify our original assumption.

45. *Line 467-470: Cause and effect seem to be mixed up here. Sea surface warming is driven by air temperature, which in the long-term is driven by radiative forcing (greenhouse effect) rather than visible light. Also changes in incident irradiance at the sea surface are expected to be far smaller than changes in sea surface temperature due to climate change. Surface warming drives stratification, which then results in greater overall light intensity and lower nutrient availability for phytoplankton trapped in a more shallow surface mixed layer. Also some references should be provided in this section.*

    We will rephrase the sentence accordingly and will cite Hutchins and Fu (2017) (see Figure 1 therein) and Boyd et al. (2015).

46. *Line 474: The word "out" should be placed after "carried"*

    Will be changed as suggested.

47. *Line 482-493: This discussion of organic matter decoupling is a bit muddled and unclear. I point out specific problems below. Generally, the value of this paragraph and its connections to the main point of this work are not clear. Is point here simply that P:N:C of cultured phytoplankton analysed here do not directly correspond to ocean particulate matter P:N:C due to the presence of detritus and decomposition?*

Thank you for clarification and yes, that is exactly our main point. We will simplify this paragraph to make our message clear that C:N:P of cultured phytoplankton analyzed here do not directly correspond to ocean particulate matter C:N:P.

48. *Line 484: "organic matter accumulation and remineralization". Are implying that detritus plays a role in bulk organic matter P:N:C? If so, this should be stated directly. Amongst the possible causes of decoupling between expected phytoplankton stoichiometry and measured bulk organic matter stoichiometry, detrital material is likely very important and not addressed. Some helpful references:*
    - *Karl D.M., Dobbs F.C. (1998) Molecular Approaches to Microbial Biomass Estimation in the Sea. In: Cooksey K.E. (eds) Molecular Approaches to the Study of the Ocean. Springer, Dordrecht*
    - *Verity, P. G., Williams, S. C., & Hong, Y. (2000). Formation, degradation, and mass: volume ratios of detritus derived from decaying phytoplankton. Marine Ecology Progress Series, 207, 53-68.*

    We will rephrase this sentence from " … stoichiometry of organic matter accumulation and remineralization" to "stoichiometry of detrital material and decomposition (Karl and Dobbs 1998; Verity et al., 2000; Zakem and Levine, 2019)".

49. *Line 485-488: This sentence is unclear. One point of Martiny et al 2013a is the increase in C:N (or rather a decrease in N:C) of sinking organic matter (see Figure 4 therein). Aside from that point, it is not clear how sinking organic matter being close to Redfield composition predicts low N:C in phytoplankton.*

    We agree that this sentence was unclear and will be removed.

50. *Line 494-505: The study by Moreno et al. 2018 would be good to include here. It not only supports your point about the value of flexible stoichiometry in global biogeochemical models, it particularly highlights the more flexible P:C of diatoms as an important driver of global patterns*
    - *Moreno, A. R., Hagstrom, G. I., Primeau, F. W., Levin, S. A., & Martiny, A. C. (2018). Marine phytoplankton stoichiometry mediates nonlinear interactions between nutrient supply, temperature, and atmospheric CO2. Biogeosciences (Online), 15(9).*

    Thank you for this suggestion. We will cite Moreno et al. 2018.

51. *Line 508: This point seems overstated and not in accordance with your results. Didn't you show that irradiance has no clear effect on P:C and only a weak effect on N:C?*

    We will change this sentence to "… temperature and light dependencies are also important for determining C:P and C:N, respectively."

52. (*Lines) 510-516: This is an interesting suggestion. You have made other predictions based on your meta-analysis, so you should actually present a prediction using this powerlaw function if you are going to suggest it. Or at least use this function to highlight what terms need to be better constrained and/or what terms should be added (e.g. detrital contribution, decomposition) in order to properly apply a power-law formulation to ocean stoichiometry.*

We have actually presented a prediction using power-law in Table 3 but it was not explicitly stated in lines 510-516. We will reorganize the order of paragraphs so that equation (2) comes before the prediction.

53. *Line 520: remove the word "on"*

    Will be changed as suggested

54. *Line 521: "… evolve under the climate change." is grammatically incorrect or a typo. "under the" could just be changed to "with" or one of many other revisions could be applied*

    Will be changed from "under the" to "with".

55. *Line 525: Remove the word "the".*

    Will be changed as suggested

References:

56. *Be sure to double-check reference formatting. Reference titles should not be in all caps (a frustrating result of citing articles from Journal of Phycology).*

    Thank you for pointing out. We now have corrected all the reference formats.

<u>Cited literature:</u>

Ågren, G. I.: The C:N:P stoichiometry of autotrophs - Theory and observations, Ecol. Lett., 7(3), 185–191, doi:10.1111/j.1461-0248.2004.00567.x, 2004.

Boyd, P. W., Lennartz, S. T., Glover, D. M. and Doney, S. C.: Biological ramifications of climate-change-mediated oceanic multi-stressors, Nat. Clim. Chang., 5(1), 71–79, doi:10.1038/nclimate2441, 2015.

Collins, S., Boyd, P. W. and Doblin, M. A.: Evolution, Microbes, and Changing Ocean Conditions, Ann. Rev. Mar. Sci., 12(1), annurev-marine-010318-095311, doi:10.1146/annurev-marine-010318-095311, 2020.

Demmig-Adams, B. and Adams, W. W.: Photoprotection and other responses of plants to high light stress, Annu. Rev. Plant Physiol. Plant Mol. Biol., 43(1), 599–626, doi:10.1146/annurev.pp.43.060192.003123, 1992.

Geider, R. J., MacIntyre, H. L. and Kana, T. M.: A dynamic model of photoadaptation in phytoplankton, Limnol. Oceanogr., 41(1), 1–15, doi:10.4319/lo.1996.41.1.0001, 1996.

Hutchins, D. A. and Fu, F.-X.: Microorganisms and ocean global change, Nat. Microbiol., 2(6), 17058, doi:10.1038/nmicrobiol.2017.58, 2017.

Klausmeier, C. A., Litchman, E. and Levin, S. A.: Phytoplankton growth and stoichiometry under multiple nutrient limitation, Limnol. Oceanogr., 49(4part2), 1463–1470, doi:10.4319/lo.2004.49.4_part_2.1463, 2004.

Laws, E. A. and Bannister, T. T.: Nutrient- and light-limited growth of Thalassiosira fluviatilis in continuous culture, with implications for phytoplankton growth in the ocean, Limnol. Oceanogr., 25(3), 457–473, doi:10.4319/lo.1980.25.3.0457, 1980.

Leonardos, N. and Geider, R. J.: Responses of elemental and biochemical composition of Chaetoceros muelleri to growth under varying light and nitrate : phosphate supply ratios and their influence on critical N: P, Limnol. Oceanogr., 49(6), 2105–2114, doi:10.4319/lo.2004.49.6.2105, 2004.

Li, G., Brown, C. M., Jeans, J. A., Donaher, N. A., McCarthy, A. and Campbell, D. A.: The nitrogen costs of photosynthesis in a diatom under current and future pCO 2, New Phytol., 205(2), 533–543, doi:10.1111/nph.13037, 2015.
Moreno, A. R. and Martiny, A. C.: Ecological Stoichiometry of Ocean Plankton, Ann. Rev. Mar. Sci., 10(1), 43–69, doi:10.1146/annurev-marine-121916-063126, 2018.
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.
Talmy, D., Blackford, J., Hardman-Mountford, N. J., Dumbrell, A. J. and Geider, R. J.: An optimality model of photoadaptation in contrasting aquatic light regimes, Limnol. Oceanogr., 58(5), 1802–1818, doi:10.4319/lo.2013.58.5.1802, 2013.
Tanioka, T. and Matsumoto, K.: Buffering of Ocean Export Production by Flexible Elemental Stoichiometry of Particulate Organic Matter, Global Biogeochem. Cycles, 31(10), 1528–1542, doi:10.1002/2017GB005670, 2017.
Ward, B. A., Collins, S., Dutkiewicz, S., Gibbs, S., Bown, P., Ridgwell, A., Sauterey, B., Wilson, J. D. and Oschlies, A.: Considering the Role of Adaptive Evolution in Models of the Ocean and Climate System, J. Adv. Model. Earth Syst., 1–19, doi:10.1029/2018MS001452, 2019.
Yvon-Durocher, G., Dossena, M., Trimmer, M., Woodward, G. and Allen, A. P.: Temperature and the biogeography of algal stoichiometry, Glob. Ecol. Biogeogr., 24(5), 562–570, doi:10.1111/geb.12280, 2015.