We thank anonymous Referee #2 for his/her constructive criticism and valuable comments. In the following we address the points brought up, with referee comments in boldface and author responses in normal typeface.

**In various places throughout the manuscript (e.g. Page 9993, line 26, Page 994, line 11 etc.) the authors refer to limitation when making inferences on the basis of ratios of available or supplied inorganic nutrients. Actually their own experiments suggest that this link is far from straight forward and I would encourage them to clarify where possible, maybe stating that the dissolved ratios indicate the 'potential for one nutrient to becoming limiting before the other' or sticking to the use of terms like 'deficiency', 'deficit', 'excess' etc. (see e.g. Page 9995, line 1).**
We agree and revised the manuscript as recommended. For example, the sentence on Page 9993, line 18 was changed as follows: „Nevertheless, the nitracline tends to be deeper than the phosphocline in the ETNA (Hauss et al., 2013; Sandel et al., 2015), which also points towards a deficiency of N over P in the euphotic zone.

**Experimental methods and statistical analysis need to be further described in places. In particular, although clear through consulting Table S1, the number of replicate mesocosms for individual treatments should be more clearly indicated to the reader, e.g. through stating in the text on Page 9997. Additionally, on Page 10001 the authors introduce a complex statistical model for the interpretation of the data without providing any justification for why this was required or chosen. Overall I was not sure why the statistical model was required as it appeared to largely be used just for the analysis of the nifH gene/transcript data and it wasn't clear that it added much to the interpretation of this data. Additionally it wasn't clear to me whether the analysis presented in Figure S1 was based on the GLM modelling performed or simple correlation analysis? Additionally, why is Figure S1 in supplementary rather than within main body of manuscript?**
We added information about treatment replicates to the Material and Method section as follows: „In the first experiment, the P supply was changed at constant N supply (*varied P*) in thirteen of the sixteen units, while in the second experiment the N supply was changed at constant P supply (*varied N*) in twelve of the sixteen units. Each of these nutrient treatments was replicated 3 times. In addition to this, "cornerpoints" were chosen, where both the N and P supply was changed. The „cornerpoints" were not replicated."

Due to the fact that both referees don't see the added benefit in introducing a model to interpret our data, we decided to remove the model from our study and instead show the original transcript data (please also see the comment to referee 1).

**Although an entirely feasible explanation, I think any potential causal link between the accumulation/availability of DOP and enhanced N2 fixation needs to be treated with caution on the basis of the data presented and experiment(s) performed. e.g. Page 10005, lines 11-20, an alternative interpretation might be that both the accumulation of DOP and the enhancement of N2 fixation are occurring within the 'varied P' experiments independently simply as a result of the addition of inorganic P. The authors may argue that the time series of DOP, P, POP, N2 fixation might argue against this (e.g. Figure 10), but given only 2 sampling time points for N2 fixation I would argue this remains equivocal. I would suggest the authors may simply wish to acknowledge this potential caveat.**
We agree that the existence of only two sampling points for $N_2$ fixation has to be emphasized more when interpreting and discussing our data set. The text now reads:

"In our experiments a significant increase in $N_2$ fixation rates was only measured in *varied P*. In mesocosms with highest $N_2$ fixation rates, DIP was depleted after day 5 or 6 while POP increased until the end of the experiment. After DIP depletion, DOP concentrations declined, which indicates that DOP served as P source until the end of the

experiment. It has to be noted that $N_2$ fixation rates were only measured at the beginning and the end of our experiment and possible fluctuations over time cannot be accounted for. However, increasing diazotrophic transcript abundances of *Richelia intracellularis* in symbiosis with the diatom *Rhizosolenia* (Het I) were also detected over the course of the *variable P* experiment. While the diatom abundance was probably favored by replete amounts of silicate added at the beginning of the experiment, no increase in diatom-diazotroph associations (DDAs) was detected in the *varied N* experiment. Measured $N_2$ fixation rates and transcript abundances leads us to speculate that DDAs were favored in the *varied P* experiment, where diazotrophs in the mesocosms utilized DOP resources in order to supply P to themselves and/or their symbiont."

Concerning the alternative explanation suggested by the referee, we do not believe that $N_2$ fixation was solely enhanced by the addition of inorganic P in our experiment, since $N_2$ fixation was not measured in all treatments in *varied P*, but only in those treatments were inorganic P was depleted after a couple of days and DOP served as an alternative P source.

**Given the extensive measurements of the P pools (see e.g. Figure 10), it would have been useful to see an attempt at mass balance.**
We addressed this issue by data presented in Fig 10. Mass balances were not subject of this manuscript but will be addressed in a follow up study. More extensive presentation of mass balances would be beyond the scope of this study.

**Page 9995, line 26: '. . .are regarded as key factors. . .'**
This was changed.

**Page 10004, line 11-15: This text does not appear to be fully consistent with the content of Figure 8? i.e. nifH Fil do not appear to be dominant for either experiment in this figure?**
Due to the removal of the model from the manuscript (see comment above), Fig. 8 and 9 were dismissed. A new Fig. 8 was added to show the original transcript data. We ensured that the text describes the figure appropriately.

**Page 10006, line 12 (and elsewhere): it is worth noting that the POC, PON, POP data reported will not just reflect that of 'primary producers' but actually will represent average values for the whole microbial community.**
We agree and made the appropriate changes. For example, the sentences on Page 10006 now reads: "There is a large difference between the supply ratio of inorganic nutrients and the PON:POP ratio of the plankton community in our study."

**Page 10008, line 24: The authors could be more specific here. They are specifically discussing excess inorganic P. Related, the authors should use the more specific term DIP to refer to dissolved inorganic P when appropriate throughout (compare Page 9999 line 13 with Page 10008, line 24).**
As already stated in the response to referee 1, we made the use of DIP, $PO_4^{3-}$ and P consistent throughout the manuscript in order to avoid confusion.

**Page 10009, line 9: do the authors mean P* here? i.e. DIP – DIN/16 or some similar definition c.f. Deutsch et al. 2007? If so I don't think the term has been defined to this point in the manuscript.**
Yes, we refer to P* as described by Deutsch et al, 2007. The term P* has been introduced in the Introduction, Page 9993, line 16.

**Page 10009, line 29: '. . .locally prior to offshore transport.'**
This has been corrected.

**A number of the figure captions (and associated statistics) require work and/or better description and figures could be clarified in places:**

**Figure 1: please explain error bars (standard deviations? Standard errors?)**
The figure caption now reads: “Experimental design and initial nutrient supply conditions during *varied P* (blue circles) and *varied N* (red diamonds). “Cornerpoints” during *varied P* and *varied N* are depicted as grey circles and white diamonds, respectively. Error bars denote the standard deviation of replicated (n=3) treatments.

**Figure 2: shaded areas were a bit difficult to make out**
We decreased the transparency of the shaded areas to make them better visible.

**Figures 3, 4 & 6: error bars for data points need explanation, regression lines also need to be described in caption. Also were the fits model I or model II type regressions?**
The figure captions now read:
“Figure 3: Maximum POC, PON and POP build-up as a function of the initial supply of N, P and N/P. Maximum δPOM is defined as peak POM concentration subtracted by the initial (day 1) POM concentration. Treatments in *varied P* are depicted as blue circles; treatments in *varied N* are depicted as red diamonds. Error bars denote the standard deviation of replicated (n=3) treatments. Regression lines (continuous lines) indicate linear correlations between the initial nutrient supply and POM accumulation.”

“Figure 4: PON/POP stoichiometry during (A) the exponential growth phase and (B) the stationary growth phase of the experiment. The grey line visualizes the Redfield Ratio. The color code, symbols and lines are the same as in Fig. 3.”

“Figure 6. Positive linear correlation between maximum DOP build-up (defined as peak DOP concentration subtracted by the initial DOP concentration) and initial P supply during *varied P* (blue circles) and *varied N* (red diamonds).”

The information about the type of regression was added to the Material and Methods section.

**Figure 7: error bars again need description. Additionally what statistical test was being used here?**
The figure caption now reads: “Figure 7: Mean $N_2$ fixation rates measured on day 2 and day 8 of both experiments. Because of the high variance between replicates we omitted $N_2$ fixation rates from un-replicated treatments. Asterisks indicate a significant difference between day 2 and day 8 (paired t-test). Error bars denote the standard deviation.”

**Figure 10: error bars.**
The figure caption now reads: “Dynamics of $PO_4^{3-}$, POP and DOP in all mesocosms. Because of the high variance between replicates we omitted $N_2$ fixation rates from un-replicated treatments. Error bars denote the standard deviation.”

**Figure S1, caption and figure do not appear to match. Caption refers to ‘a’ and ‘b’ parts when there only appears to be one part in figure?**
Please see comment above.