

Response to Referee #2

The authors would like to thank Referee #2 for considering this manuscript being interesting and appropriate for publication in Biogeosciences. The manuscript largely improved by including the comments and suggestions provided by Referee #2. We addressed the comments as follows; modifications are highlighted in green in the main text:

General comments

Referee comment (RC) 1: The abstract says that “metagenomic” data are shown, but the molecular markers used in this study are for bacterial/archaeal diversity, and a few functional genes. “Meta” suggests that large portions of the community genes are evaluated, which is not the case. It would be more accurate to edit this phrase. Outside of the abstract, this term is not used, so the body of the text is appropriate in scope.

Author comment (AC)1: We removed the term from the abstract and replaced it with microbial community study.

RC2: Here the statement is made that enhanced primary productivity fuels enhanced export, but there is little to no primary evidence within the manuscript to support this statement. Is export flux greater within the eddy than in nearby regions, and if so, how is the time-space decoupling of productivity and flux resolved?

AC2: Within the framework of this special issue a paper (currently in typeset) has been submitted by Gerhard Fischer (MARUM; Bremen, Germany) and colleagues entitled: “Bathypelagic particle flux signatures from a suboxic eddy in the oligotrophic tropical North Atlantic: production, sedimentation and preservation”. The paper discusses sediment trap data from a low-oxygen eddy observed in 2010 from that same area. The authors document a remarkable impact on all productivity related processes. They estimated a 3-fold higher productivity in the surface layer compared to surrounding waters. In particular they found a multiple times increase in the mass flux in the bathypelagic traps (2300 and 170 m above the seafloor) during the eddy passage. Furthermore, Fiedler et al. (to be submitted to this special issue as well) determined export flux derived from Carbon remineralization rates within the eddy and found a 3-4-fold enhanced export flux compared to background conditions in the open-ocean ETNA. We added these references and the related information to the text.

RC3: The cut-off of 90 $\mu\text{mol/L}$ oxygen concentration to differentiate ‘realm’ effects is not sufficiently supported. Is there evidence in the literature for such a cut-off, for example, are certain microorganisms known to respond differently across this threshold in relation to metabolism/productivity and therefore, it is an ecologically important distinction?

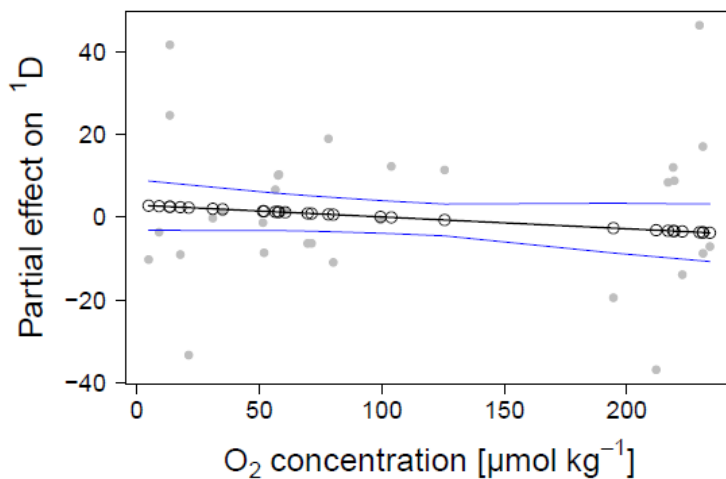
AC3: This range was chosen for two reasons: The first reason was to obtain sample groups of fairly equal size between stations for statistical reasons.

Indeed, there is evidence in the literature, that has not been explained sufficiently in our manuscript is that $90 \mu\text{mol L}^{-1}$ is the highest concentration of O_2 at which denitrification has been described to be active (Gao et al., 2010). We included this information and the reference into the methods section.

Gao, H., Schreiber, F., Collins, G., Jensen, M. M., Kostka, J. E., Lavik, G., de Beer, D., Zhou, H.-y., and Kuypers, M. M. M.: Aerobic denitrification in permeable Wadden Sea sediments, ISME J, 4, 417-426, 2010. <http://www.nature.com/ismej/journal/v4/n3/abs/ismej2009127a.html>

RC4: Relating again to the oxygen concentration cut-off – how do communities compare along the oxygen gradient? Does alpha diversity (or total OTU abundance) decrease with decreasing oxygen concentration?

AC4: As stated on p. 14187, l. 22-23, our analysis does not show a pronounced effect of O_2 on alpha diversity if other effects are controlled (see figure below): The black regression line through the fitted values (circles) is compatible with zero, as indicated by the blue 95% confidence interval.



RC5: Within the methods section, where volumes are given for reagents within assays, it would be more useful to provide the final concentration.

AC5: We agree and added the missing information to the text.

‘Reactions were performed in technical duplicates in a final volume of 12.5 μL using 0.25 μL of each primer (10 $\text{pmol } \mu\text{L}^{-1}$), 3.25 μL nuclease-free water and 6.25 μL SYBR qPCR Supermix W/ROX (Life Technologies, Carlsbad, CA, USA) on a ViiA7 qPCR machine (Life Technologies, Carlsbad, CA, USA) according to established protocols (Ahlgren et al., 2006; West et al., 2011). TaqMan-based qPCRs were performed for picophytoplankton (*Prochlorococcus/Synechococcus*) and bacteria as previously described

(Suzuki et al., 2001) in a final volume of 12.5 μL with primer/probe concentrations as shown elsewhere (Table 1, (West et al., 2011)), but with the addition of 0.5 μL BSA (20 mg mL^{-1}) and 6.25 μL TaqMan Mix (Life Technologies, Carlsbad, CA, USA).'

RC6: On several occasions the phrase "of around" is used to mean "approximately." While generally well written, the manuscript requires some additional editing for increased readability.

AC6: We changed this where adequate; the manuscript was also corrected by a native speaker, now (see also comments to referee 1).

Specific comments:

RC7: p. 14182 l. 18 – DNA and RNA were quantified fluorometrically using a Nanodrop. This instrument is a spectrophotometer.

AC7: This is true, we modified the sentence.

RC8: p. 14185 Statistics section – intent and readability would be improved for each subsection with an initial sentence about the statistical process and its purpose; especially for those less familiar with the exact procedures.

AC8: We addressed this by adding initial sentences to each section, which explain the purpose of the analysis. The section now reads:

'2.5 Statistics

Low-abundance OTUs were removed to reduce noise and computation time. Statistical downstream analysis was performed in R v3.1.3 (R Core Team, 2015) with custom scripts (available from the authors on request). As OTUs of very low abundance only increase computation time without contributing useful information, they were removed from the data set as follows: After transformation of counts in the sample-by-OTU table to relative abundances (based on the total number of reads per sample), OTUs were ordered by decreasing mean percentage across samples. The set of ordered OTUs for which the cumulative mean percentage amounted to 99% was retained in the filtered OTU table.

Distribution of OTUs across samples was modeled by a set of environmental variables with minimal interdependence. The variance in OTU composition (i.e., the extent of change in OTU abundance across samples) explained by the measured environmental variables was explored by redundancy analysis (RDA) with Hellinger-transformed OTU counts (Langfeldt et al., 2014; Stratil et al., 2013; Stratil et al., 2014) using the R package *vegan* (Oksanen et al., 2013). In order to minimize collinearity of explanatory variables in the RDA model, a subset of the recorded environmental variables was chosen according to their variance inflation factor (VIF), employing *vegan*'s functions *rda* and *vif.cca*. Starting with an RDA

model that contained all explanatory variables, the variable with the highest VIF was iteratively determined and removed from the model until all remaining explanatory variables had a VIF <2.5.

OTU distribution was subject to "Realm" depending on O₂ concentration. Model selection started with a full RDA model containing all main effects and possible interactions based on the set of explanatory variables with minimal collinearity. This model was simplified by backward selection with function *ordistep*. The final RDA model exhibited a significant interaction effect "Realm:O₂" (see results section). For plotting and indicator analysis (see below), the continuous variable "O₂" was converted into a factor with two levels "high O₂" (>90 µM) and "low O₂" (≤90 µM); the threshold of 90 µM was chosen to obtain sample groups of fairly equal size between stations, which include low O₂ parts of the water column at all sampling stations in order to enable a comparison between the ETNA OMZ (outside the eddy) and the eddy OMZ.

We determined OTUs typical for a given combination of levels of factors "Realm" and "O₂". OTUs significantly correlated with any axis in the final RDA model were determined using the function *envfit* with 10⁵ permutations, followed by Benjamini-Hochberg correction (false discovery rate, FDR) (Benjamini and Hochberg, 1995). In order to reduce the number of tests in this procedure, OTUs were pre-filtered according to their vector lengths calculated from corresponding RDA scores (scaling 1) by profile likelihood selection (Zhu and Ghodsi, 2006).

OTUs significant at an FDR of 5% were further subject to indicator analysis with function *multipatt* of the R package *indicspecies* v1.7.4 (De Cáceres and Legendre, 2009) with 10⁵ permutations. Indicator OTUs – in analogy to indicator species *sensu* De Cáceres and Legendre (2009) – are OTUs that prevail in a certain sample group (here: a level of factor "Realm" within a chosen O₂ level) while being found only irregularly and at low abundance in other sample groups. In order to remove the effects of the covariate "Depth" in indicator analysis, Hellinger-transformed counts of significant OTUs were first subjected to a linear regression with "Depth"; residuals of this regression were then transformed to positive values by subtraction of their minimum and used as input for indicator analysis.

3D visualizations of the RDA model were produced in kinemage format (Richardson and Richardson, 1992) using the R package *R2Kinemage* developed by S.C.N., and displayed in *KiNG* v2.21 (Chen et al., 2009).

Diversity within samples was related to environmental variables by advanced linear regression. For alpha diversity analysis, effective OTU richness (Shannon numbers equivalent, ¹D, (Jost, 2006, 2007)) was calculated from the filtered OTU table. ¹D was fitted to the set of explanatory variables with minimal collinearity in a generalized least squares (GLS) model using function *gls* of the R package *nlme* v3.1-120 (Pinheiro et al., 2015). The variable "NO₂" was square root-transformed to decrease the potential leverage effect of its two highest values (0.25 µM and 0.28 µM, respectively) on ¹D. Apart from main

effect terms, the interaction term “Realm:O₂” was included into the GLS model for comparability with beta diversity analysis (see results section). The variance structure of the GLS model was chosen to account for both different variances per level of “Realm” and an overall decreasing variance by “Depth”. The resulting model was validated following the recommendations of Zuur et al. (2009). While only the “Realm” effect was significant, the other terms were kept in the model to maintain a valid residual distribution. For visualization of the (partial) effect of only factor “Realm” on ¹D, partial response residuals were extracted from the full GLS model re-fitted without the “Realm” main effect. These partial response residuals were then modelled by the “Realm” main effect alone, using the same variance structure as for the full GLS model.’

RC9: p.14188 l. 13-15 What were the depths at which samples were collected that are considered below the euphotic zone?

AC9: We used the term ‘euphotic zone’ as the depth where photosynthetic available radiation (PAR) is <1% of its surface value. During this survey, this was the case below 60 m water depth, this information has been added to the text.

RC10: p. 14188 l. 16-17 How might carbon fixation measurements be affected if total volumes were not filtered for delta13C enrichment? Could primary productivity estimates be over-estimates?

AC10: Productivity could not be overestimated because the filtrate volume was taken into account for calculating primary productivity. Incubation bottles were also shaken periodically every ten minutes while doing filtration for homogenization of particulate matter.

RC11: p. 14192 l. 12-16 Are cell counts (microscopy and/or flow cytometry) available to conclude that the qPCR Prochloro/Synechococcus data are representative of relative differences in abundance of cyanobacteria and eukaryotes in- and outside of the eddy or within either chl max layer?

AC11: Unfortunately, cell counts are not available. Our approach was to compare the 16S rRNA amplicon dataset with qPCR data, which also includes two different methods. Quantification by qPCR has been performed against a standard dilution series, and then compared to the sequence abundance in the 16s rRNA pool to assure that the respective clusters are key organisms in the samples. As a related comment has been made by Referee 1 we also added a detailed table on cyanobacterial and eukaryotic phytoplankton distribution to the supplement.

RC12: p.14192 l. 17-28 Why might HL-adapted Prochlorococcus ecotypes be abundant below the euphotic zone? This seems counter-intuitive. Suggestions on why this might be would be interesting.

AC12: The detected ecotype is described to cluster among the HL *Prochlorococcus* strains. However, also in the Pacific it has been recovered from waters below the euphotic zone. It could therefore be thought that the 16S rRNA based classification of HL and LL strains may differ from the functional classification. A future approach to obtain information on that could include sampling for flow cytometry. A pigment and biomass analysis following cell sorting may lead to this information in future studies.

RC13: p. 14193 l. 12-13 This paper states that *Prochlorococcus* could contribute up to 40% of the DOC that could support bacterial production. As written, the statement suggests that *Prochlorococcus* is responsible for 40% of bacterial production.

AC13: Thanks for this comment, we corrected this statement, it now reads:

'This may be critical as *Prochlorococcus* is one of the most abundant photosynthetic organisms in the ocean and contributes up to ~40% of dissolved organic carbon for bacterial production (Bertillon et al., 2005).'

RC14: p. 14195 l. 20 I am surprised that *nifH* genes were not quantifiable from the eddy. *nifH* genes have been retrieved from this region. Assay detection limit? High inorganic dissolved nitrogen concentrations and N:P ratios close to Redfield do not exclude the possibility of diazotrophs and/or biological nitrogen fixation.

AC14: It is for sure true that inorganic N compounds do not exclude N₂ fixation in all cases. The argument was used to rather explain the absence of N₂ fixers. The detection limits were determined from no-template controls. Those were run in duplicate for each primer and probe set, and were undetectable after 45 cycles, thus setting the theoretical detection limit of our assay mixture to one *nifH* copy. In reality, the detection limit depends on the amount of seawater filtered per sample, elution volume after extraction, and the amount of sample loaded to the qPCR assay, too. We eluted in 100 µL of elution buffer, therefore, when using 2.5 µL of the eluate, a minimum of 40 copies would be needed for the detection of 1 copy. Based on a filtration volume of 2L seawater, the detection limit would be 20 copies L⁻¹. qPCR efficiencies were calculated using the formula $E = 10^{-1/\text{slope}} - 1$, and were between 95.3% and 96.8%.

Figures

1. no comments
2. For the oxygen concentrations, can the profile of discrete O₂ concentrations be shown?

AC: The discrete profiles are shown in figures 3 and 4.

3. Figures should be larger for easier readability.

AC: We will increase the size of figures; this has also been suggested by Referee 1.

4. Figures should be larger for readability. Greater transparency of the colored bars would make the trends easier to compare across panels. Figure legend reads 'oxygen versus depth,' but this line is oxygen concentration. Colors used for Proteobacteria and Bacteroidetes are very similar. Would be easier to discern with different colors.

AC: We will increase the size of this figure, too. Colors have been modified to make the difference between different phyla more obvious. The legend has been changed to O₂ [μmol kg⁻¹]

5. Dark purple and dark blue points are difficult to discern from one another.

AC: We modified the colors.

6. Check eddy axis labels. Is this correct with 'eddy_2' on the left?

AC: Yes, this is correct.

7. Transparency of bars needs to be greater so data points can be seen. Are these discrete or derived measurements of chl a? Legend edit is required and methods section should include description of chl a measurement methods. Chl a units are missing.

AC: These are discrete measurements, Chl a was measured from filter samples, the unit is μg L⁻¹. We modified the figure and the legend by adding this information; we further increased the transparency of the bars. A description of chl a measurements has been added to the methods section:

'Sea water samples (0.5 – 1 L) for chlorophyll *a* (Chl *a*) analyses were filtered (200 mbar) on GF/F filters (25 mm, 0.7 μm; Whatman, Maidstone, UK). Filters were transferred to a plastic vial and 1 ml of MilliQ water was added. Filters were immediately frozen at -20°C and stored for at least 24 h. Afterwards, 9 ml acetone (100 %) was added to the vials and the fluorescence was measured with a Turner Trilogy fluorometer (Sunnyvale, CA, USA). Calibration took place using a Chl *a* standard dilution series (*Anacystis nidulans*, Walter CMP, Kiel, Germany). Chl *a* concentrations were determined as described by Parsons et al. (1984).'

8. Difficult to discern low concentration areas of the plots. Is the number zero copies, or "not detected?" What is the detection limit of the assays? Symbols are difficult to differentiate as plotted. Increase size of plots and/or data points. X axes' labels should be edited to be consistent in format.

AC: The size of the figure has been increased. Symbols sizes were increased and the colors were modified. The x axis of Fig. 8A was adjusted. The detection limits of the qPCR assays are mainly the same as described above for *nifH*, again determined from no-template controls. Those were run in duplicate for each primer and probe set and were undetectable after 45 cycles with a theoretical detection limit of 1 copy. As described, above, we eluted in 100μL buffer, therefore, when using 2.5μL of the eluate, a minimum of 40 copies would be needed for the detection of 1 copy. Based on a filtration volume of 2L seawater, the detection limit would be 20 copies L⁻¹.

9. One symbol could be 'open' so overlapping data can be more clearly seen. Again, what is the detection limit of the assay? Is the data point zero or 'not detected?'

AC: The data point zero means below detection limit of the assay (<20 copies L^{-1}), therefore it is between zero and 20 copies L^{-1} .