Referee #1 Reply

Major issues

1) Experiments were conducted with nitrate and ammonium added at at least 10 fold higher concentrations than in situ values (100 μM vs. 10 μM and 10 μM vs. 1 μM, respectively, i.e., 1000% above ambient, and not 90% as stated in the text p. 11 l. 6).

We rephrased the text to avoid misunderstanding (p 12 line 8)

This means that the measured rates must be treated as potential rates unless the authors can establish an argument for 0th-order kinetics for both denitrification and nitrification. In turn, this implies that the estimated sponge-ground rates may be vastly (10-fold) overestimated. This issue should be discussed and the conclusions modified accordingly.

We agree that these are potential rates. We make this now more clear in the discussion, beginning of chapter 4.4.

^{15}N incubations were based on standard methods (Dalsgaard et al., 2003 and Hannig et al., 2007) with minor modifications as per Hoffmann et al. 2009. These methods allow us to estimate the rates at ambient NO3 concentrations based on 15N incubations. They are not the rates measured directly by the 15N labelled N2 production. This is clarified in Section 2.3. Although concentrations of labelled $^{15}N^-$ exceeded background ^{14}N 10-fold, potential denitrification rates were well within the range of those previously reported for cold and warm water sponges, where ^{15}N amendments were more reflective of ambient NO_3^- concentrations (Hoffmann et al., 2009;Schläppy et al 2010a.

In the oxic experiments, denitrification rates could, in principle, be calculated using the classic isotope pairing calculations for sediment cores (D14 sensu Nielsen 1992), but then the incubations should have been performed without addition of unlabelled ammonium and with maintenance of steady state

Labelled ammonium was added to the annamox incubations. No ammonium was added to the denitrification experiment. We now discovered that this was not clearly explained in the method section, which have led to confusion. The text is now corrected.

Calculations of coupled nitrification-denitrification were based on the approach of Hoffmann et al., 2009, where a similar experimental set-up was employed. The potential rates of denitrification and proportion of coupled nitrification-denitrification are comparable for Geodia barretti – 92 nmol N cm-3 sponge day-1 ; 26% coupled nitrification-denitrification (Hoffmann et al., 2009) relative to 96 nmol N cm-3 sponge day-1; 16% coupled nitrification-denitrification (this study).

2) Nitrification-based denitrification rates are calculated from the accumulation of single labelled 29N2. Firstly, it is not entirely clear how these rates and relative contributions were calculated, and I suggest to include the essential equations in Methods.

The abundance of $^{28}N_2$, $^{29}N_2$ and $^{30}N_2$ were analysed from gas samples using a continuous flow isotope ratio mass spectrometer (CF/IRMS). Calibrations were achieved by injecting lab air and additional in house reference gas samples. Denitrification rates were calculated from the production of ^{15}N isotopes (see below) according to the method described by Nielsen (1992).

The rate of denitrification was measured from ^{15}N isotope production (equations 1 and 2). D_{14} and D_{15} represent denitrification of labelled $^{15}NO_3^-$ and $^{14}NO_3^-$. p ($^{14}N^{15}N$) and p ($^{15}N^{15}N$) are the production rates of the 2 labelled N_2 species $^{14}N^{15}N$ and $^{15}N^{15}N$ (Rysgaard et al.1995). D_{15}

is indicative of denitrification of labelled $^{15}NO_3^-$ and D_{14} represents in situ denitrification of $^{14}NO_3^-$.

$$D_{15} = p\,(^{14}N\,^{15}N) + 2p\,(^{15}N\,^{15}N) \quad (1)$$

$$D_{14} = \frac{p\,(^{14}N\,^{15}N)}{2p\,(^{15}N\,^{15}N)}\,D_{15} \quad (2)$$

To estimate denitrification of NO_3^- from ambient sea water (D_w), in terms of D_{14}, the following calculation was applied (equation 3):

$$D_w = D_{15}\,[^{14}NO_3]_w / [^{15}NO_3]_w \quad (3)$$

where $[^{14}NO_3]_w$ and $[^{15}NO_3]_w$ represent the concentration of unlabelled and labelled NO_3^- in ambient seawater.

In situ coupled denitrification (D_n), in terms of D_{14}, was calculated using equation 4 (see below).

$$D_n = D_{14} - D_w \quad (4)$$

These equations are now given in the method section, page 15.

Secondly,
the concept of water-based and nitrification-based denitrification was developed by Nielsen for sediment cores with steady state distributions of oxygen and nitrate (and it was challenged by Middelburg in L&O 41:1839). In the present study, oxygen was clearly not at steady state during the oxic incubations, and it also seems likely that new formed nitrate may have leaked from the sponge tissue thus gradually decreasing the C2 labelling of the ambient nitrate pool, and increasing 29N2 production from the ambient water. Moreover, the data presented in Fig. 1, for one of the six sponges, suggests that there is an issue with the mass balance of unlabelled N in the incubations. Thus, at the end of the anoxic incubations, excess 29N2 dominated over 30N2 in two of three incubations despite the stated _90% labelling of the nitrate pool, and the accumulated 29N2, reaching up to _23 μM, exceeds the amount of unlabelled nitrate initially available (10 μM in situ + 1 μM from the 99% 15N tracer). Also during the first 24 h, 29N2 production in the anoxic incubations seems higher than predicted by nitrate labelling in the absence of nitrification. Altogether, these uncertainties and discrepancies undermine the conclusion concerning the role of nitrification. Plots of excess 29N2 vs. excess 30N2 could potentially help the authors to evaluate and constrain some of these issues.

Ambient seawater was filtered to remove water column bacteria and or phytoplankton, thus reducing the potential for background nitrification. Although conditions were not at steady state, previous application of this method was considered suitable for dissected sponge explants (Hoffmann et al., 2009). The data has been re-analysed as suggested and errors in the calculation have been corrected. Issues with mass balance and $^{29}N_2$ production have now been resolved.

Specific comments 3, 8-12: The final statement is highly speculative and does not belong in an abstract.
We agree that this statement sounds provocative, but we still consider this a valid interpretation of our data, see justification below for comment on 25,11
4, 16-7: The statement about nif genes seems out of context. *Agree, removed*
6, 14: Science should never aim to show specific results but rather test hypotheses! *Rephrased*
7, 11
9, 4-5: "Upper few centimetres" is vague – considering the negative result, the question is whether only the oxic surface layer was sampled.
As the reviewer pointed out, the "upper few centimetres" sediments are very likely oxic (our microsensor measurements indicate that oxygen can penetrate to ~75 cm below seafloor). However, denitrification (at also other anaerobic processes) is most active in the upper most sediments, due to the widespread of bioturbation and bioirrigation in marine surface sediments (e.g. see the recent paper regarding sulfate reduction rates in the Aarhus Bay sediments (Andrew Dale et al, 2019, GCA)). Therefore, measurements made using the most surface sediments are believed to represent the majority of denitrification activities in a marine sediment column. (Of course, this is supported by our reaction-transport model!)
9, 20: There was no "atmosphere" in the vials? However, incubation with a helium/ oxygen headspace would have kept the incubations oxic throughout.
No atmosphere in the exetainers as described on p11. We followed a standard protocol here.
10, 7-8: This seems a very shaky assumption. Respiration rates must vary with species, temperature, and trophic state.
They certainly do, but we do not have these details for all investigated species and needed to make the best possible estimation.
10, 18-9: Some oxygen is likely introduced during transfer – did you test the water in the Exetainers?
To verify the absence of oxygen in the de-gassed water, an anaerob strip test (colour change from pink to white under anaerobic conditions; Sigma Aldrich) was performed prior to transfer into 12mL exetainers. The caps were then replaced and the gas tight vials were carefully sealed to exclude any air bubbles. An anaerob strip was added to control exetainers (seawater only) to verify the absence of oxygen in anaerobic incubations. See p 11, line 19
11, 6: The values are _1000% above ambient.
10 times above, corrected.
11, 12: According to 7, 11 the in situ temperature was below 0 _C! How would the higher incubation temperature affect the rates?
Lab experiments can never perfectly mimmick in situ conditions, and in our case, there were no cool room available at 0 C. We are aware that this may have led to over-estimation of the Arctic rates and made a comment in the discussion.
12, 18: The accumulations in Fig. 1 look only approximately linear – which test gave p < 0.05? Did the same apply to the linearity of the anoxic rates (13, 4)? *Figures modified after re-calculation of rates*
13, 15: Please specify the equations used here (see major issue #2).
See above, our reply to major issue #2.

16, 3-5: The opening of the Results is very confusing with the first two sentences referring to two different treatments. Delete the first sentence.
The first two sentences explain why our results show the absence of anammox, they refer to the same treatment. We rephrased for more clarity.

16, 22-3: The sediment experiment has little value. The origin of the sediment is unclear, and it does not seem representative of Arctic sediments.
The origin of the sediment samples (next to Arctic sponge ground at Schulz Massiv) is clearly stated in chapter 2.2. The sediment itself is of pelagic origin (ultimately from the primary production in the surface ocean rather than terrestrial origin), this is obvious from the geographic position of the sampling site and does not need to be mentioned.
18, 5: See 6, 14.
Rephrased
18, 18-9: Metabolisms in sponges or what? Please clarify/reference.
Rephrased for clarification
18, 20-5: The presence of denitrification genes and isolation of denitrifiers cannot prove "the presence of denitrification activity".
Rephrased

20, 11: How would the "pulse of organic matter in the water column" (where in the water column?) affect potential denitrification in the sponges' tissue?
Section rephrased

21, 16: "proves" is an overstatement.
Changed to "shows"
22, 1-2: It is not the in situ concentration but the 10 μM ammonium added, that is of relevance here.
Disagree. Ammonium was only added to the annamox experiments, not to the denitrification experiments. So we have a point here.
22, 13-5: Please provide a reference for the single copies.
This is text book knowledge, we do not see a need to provide a reference
22, 16-20: The curve in Fig. 3 does not look like an exponential function. It there statistical support for this relationship?
The figure changed after re-calculating the data. The point is that there is a clear positive correlation between anaerobic denitrification rates and nir genes for most species, we made this more clear now.
22, 20: What is meant by "optimized"?
Rephrased

23, 9-10: With 6 orders of magnitude variation, this is not very telling.
Our point is that there were not less nir genes in the sediment than in some of the sponges, but no denitrification.
23, 19 on: The calculations of sponge ground rates need explanation, but see Major issue #1.
See text added and rephrased at beginning of chapter 4.4
Furthermore, it seems that results of population density surveys are presented here for the first time. If this is the case, the methods and results should be specified i the appropriate sections. Otherwise, a reference should be included.
This is described in the methods, p 8 l 21-22. We rephrased also in section 4.4 to make clear that these are careful estimates.

24, 24: What was the frequency of non-pumping?
Not possible to say something in general because this varies between species, environmental conditions etc. In the particular study quoted here, there was one non-pumping period of 1-2

hours during the experiments of 12-20 hours, this can easily be looked up in the cited reference.
25, 11-2: Is this a short-term or permanent effect? Would reduced pumping rates/increased anoxia not result in reduced growth, reduced biomass, and thereby reduced nitrogen removal in a longer perspective?
Reduced growth would first lead to reduced remineralisation activity – so it would first of all weaken the classical DIN source function of the sponge. We know for sure that reduced pumping leads to reduced oxygen in sponge tissue, but we do not know if reduced pumping leads to reduced growth, so there is no point speculating about it here.
The system effect of the stressors seems speculative.
We rephrased to make clear that this is not speculation, but a potential scenario based on valid data interpretation.

Table 1: The number of significant digits should be adjusted.
Adjusted to what? There are two digits per value, what is wrong?

Fig. 1: Different triangles are used for 29N2 and 30N2.
Resolved