Final author comments on "Dimethylsulfide (DMS) production in polar ocean may be resilient to ocean acidification" by F.E. Hopkins et al., manuscript number bg-2018-55

We thank both anonymous reviewers for their detailed, constructive, and positive reviews of our manuscript – we greatly appreciate the care and detail that has gone into its assessment. Below we respond to their comments point-by-point. It is worth noting that the related paper by Richier et al. (2018) has now been accepted for publication in Global Change Biology (doi: 10.1111/gcb.14324), providing further substantiation of many of the discussion points in this manuscript. The reviewers comments are shown in italics, with our responses shown in bold. Line numbers in our response refer to the revised version.

1. **Response to Anonymous Referee #1**

General comments

1.1 The paper describes experimental results examining the response of DMS and DMSP concentration, synthesis & production rates to acidification in Southern Ocean & Arctic waters and compares them with previously published results from the NW European shelves. The authors report regional trends in responses, which they attribute to the variability in the carbonate system and its influence on the plasticity of the phytoplankton community and DMS/DMSP response. The analysis is somewhat limited to the carbonate system & phytoplankton size class without consideration of other factors. The paper is clear and well-written, and makes important points including regional variation in response to acidification, and also that different processes occurring at different scales are responsible for the variable responses reported over different timescales (as exemplified by comparison of microcosm versus mesocosm responses). The paper is of publication standard if the points below relating to interpretation & analysis can be addressed.

We thank the reviewer for their positive view of our paper, and we are confident that we have addressed their concerns.

Specific comments

1.2 Title The comparison with the NW European Shelf results is an important part of this paper and merits mention in the title

The title has been changed to:

"Dimethylsulfide (DMS) production in polar oceans may be insensitive to ocean acidification: a meta-analysis of 18 microcosm experiments from temperate to polar waters".

'resilient' has been replaced with 'insensitive' at the suggestion of reviewer #2, see point 2.8 below.

1.3 Introduction Line 122–these microcosm experiments are not long enough to test adaptation. Results from experiments on timescales of < 1 week may give insight into plasticity and acclimation, but not "adaptive capacity"

We acknowledge this point while further noting that a key point we emphasise is that we interpret the variable sensitivity to short term experimental conditions observed as potentially corresponding to pre-existing adaptation to prevailing conditions across sampled populations/systems. Text has been altered accordingly in three places (line numbers correspond to revised version):

L124: "The focus is then on the effect of short-term CO_2 exposure on physiological processes, as well as the extent of the variability in acclimation between communities".

L126: "The capacity of organisms to acclimate to changing environmental conditions contributes to the level of resilience of key ecosystem functions, such as DMS production".

L132: "…our approach can provide insight into the physiological response of a variety of polar surface ocean communities, as well as their potential level of sensitivity to future OA when compared between environments that differ in carbonate chemistry…".

1.4 Results Fig. 3 Error bars are relatively large at 96 hours in Arctic waters – this should be noted and discussed

We feel that this observation does not add any additional insights to the results or discussion so we have simply added into the results text:

L331: "Increased variability between triplicate incubations became apparent in all three Arctic experiments by 96 h but no significant effects of elevated CO_2 on DMS concentrations were observed".

1.5 Line 304-305; Fig 4c Error - "DMSP concentrations were found to DEcrease significantly in response to elevated CO2 AT 48 h for Barents Sea (Fig. 4 C)". Also note that DMSP was not significantly different at 96 hours.

Text has been altered accordingly.

1.6 Discussion 4.1 Regional differences in the response of DMS(P) to OA The interpretation of the treatment effects would benefit from statistical analysis to support the interpretation in:

Line 375-376 "De novo DMSP synthesis and DMSP production rates show a similar relationship with DIC/Alk (Fig. 7 A and B)" - is the difference between 0.91 > and < 0.91 significant? With the exclusion of one station (DIC/Alk ~0.901) there looks to be no difference in Figure 7. Statistical confirmation required.

Statistics have been confirmed and the text now reads:

***"De novo* DMSP synthesis and DMSP production rates show a similar relationship with C_T/A_T (Fig. 7 A and B), with a significant suppression of DMSP production rates in temperate waters compared to polar waters (Fig. 7B, Kruskal-Wallis One Way ANOVA $H = 8.711$, $df = 1$, $p = 0.003$). Although a similar trend was seen for *de novo* DMSP synthesis, the difference between temperate and polar waters was not statistically significant (Fig. 7A)".**

Line 379-380. "At T1, Chl a showed little response to elevated CO2 at polar stations, whereas a strong negative response was seen in temperate waters (Fig. 8A)" – again this description does not really match the data in the figure. The polar stations show a smaller range of treatment effect than temperate stations which show both larger positive and negative effects. Statistical confirmation required

Line 380-382. "A slight positive response in Chl a was seen at most temperate stations by T2, with generally little response at polar stations (Fig. 8 B)." Aren't the highest treatment effects at the polar stations? Statistical confirmation required

Statistics have been confirmed and the text now reads:

"At T_1, a statistically significant difference in response was seen between temperate and polar waters for Chl *a* (Kruskal-Wallis One Way ANOVA *H* = 20.577, *df* = 1, *p*<0.001), with minimal response to elevated CO_2 at polar stations, and in general a strong negative response was seen in temperate waters (Fig. 8A). By T_2, no significant difference in response of Chl *a* between temperate and polar waters was detectable (Fig. 8B), although a slight positive response in Chl *a* was seen at some temperate stations, and polar stations showed a minimal response, with the exception of *Barents Sea* which saw strongly enhanced Chl *a* at T_2 (Fig. 8 B)".

1.7 The analysis is limited to considering the carbonate system & phytoplankton size class as the factors determining regional response. Other factors will have differed between the polar waters and NW European shelf and may have influenced DMS/DMSP response to ocean acidification such as temperature, light, nutrients and phytoplankton community composition. For example, the authors mention "slower microbial metabolism at low water temperatures", so could this explain the observed difference in regional response? Datasets for these variables are most likely available, and a more comprehensive analysis that considers these would benefit the paper and interpretation. This may have already been carried out by the authors, in which case it should be noted that there are no relationships between response and these other variables.

Richier et al. (2018) provide a detailed overview of the role of other potential environmental drivers for the differences in response between temperate and polar waters. This paper has now been accepted for publication in Global Change Biology, so we will refer the reader to it for more detailed analysis of this issue. To address the reviewers concerns, we have also added an appropriate amount of discussion relating to this to the current manuscript.

Firstly, to address the reviewers comment on "slower microbial metabolism at low temperature waters": we failed to observe strong responses to high CO_2 in experiments performed in Arctic waters (cruise JR271), therefore incubation times during the Southern Ocean experiments (cruise JR274) were increased from 4 to 6 to 8 days, whilst including a higher CO_2 treatment (2000 µatm). However, the magnitudes of the responses were found to be independent of overall experimental duration. As Richier et al. (2018) explain, net growth rates may be expected to be 2- to 3 –fold higher in the warmest compared to the coldest waters (Eppley, 1972), and indeed maximum Chl a-normalised photosynthesis rates were indeed 2- to 3-fold higher in polar waters – but the response to experimentally-induced OA in polar waters remained insignificant despite the length of incubation being up to twice the duration of temperate experiments, and up to 4x longer, than the 48 h time point where strong responses were typically observed in temperate waters.

We have added the following to the methods section at L210:

"The magnitude of response was not related to incubation times, and expected differences in net growth rates (2- to 3-fold higher in temperate compared to polar waters (Eppley 1972)) did not account for the differences in response magnitude despite the increased incubation time in polar waters (see Richier et al. (2018) for detailed discussion)".

Secondly, to address the reviewers concerns regarding "other factors [that] will have differed between polar waters and NW European shelf…" that may have influenced the response, we have added the following to the Discussion:

L440: "Across all experiments, the response of net total community Chl *a* and net growth rates of small phytoplankton (<10 µm) scaled with pCO_2 treatment, and strongly correlated with in situ carbonate chemistry, whilst no relationships were found with any of the other wide range of initial physical, chemical or biological variables (Richier et al. 2018). Overall, the observed differences in regional response to carbonate chemistry manipulation could not be attributed to any other measured factor that varied systematically between temperate and polar waters. These include ambient nutrient concentrations, which varied considerably but where direct manipulation had no influence on the response, and initial community structure, which was not a significant predictor of the response (Richier et al. 2018)".

1.8 A minor point here is that methodological differences should also be considering when assessing response. For example, different light cycles were used on different voyages.

Different light cycles were used on different voyages to simulate the *in situ* light conditions/light:dark cycles for the time of year of sampling. We don't expect this to have affected the response to OA, and it would have been inappropriate to have used the same light cycle on all cruises.

1.9 They should also consider the degree to which the experimental manipulations alter the carbonate system relative to the ambient mean. The magnitude of change upon acid/base addition from the mean state of the carbonate system may be a more important factor than the regional range. For example, a proportionally larger shift in pH or carbonate upon acid/base addition may initiate a greater stress response and so DMS/DMSP production.

Ambient pCO_2 and pH for each sampling station are shown in Table 1. A large range of ambient carbonate chemistry was observed across all waters, with no consistent relationship with location (e.g. pCO_2: NW Euro shelf 287 – 400 uatm, Arctic 289 – 304 uatm, Southern Ocean 272 – 510 uatm). Experimental manipulation of carbonate chemistry was accurately calculated and implemented using the CO_2SYS programme (see Richier et al. 2018 for methods, and Richier et al. 2014 Figure 3 for a comparison of target vs actual pCO_2 following experimental manipulation for NW European shelf experiments). Our experiments provided no apparent evidence of a relationship between the proportional shift in pH/carbonate chemistry and the magnitude of the response as a function of the initial state of the carbonate system, rather, as emphasised, the presence or absence of any observable response broadly correlated with the initial state of the carbonate system (see also Richier et al. 2018), with subsequent response strength then scaling with magnitude of manipulation in those experiments where any response could be observed.

1.10 Lines 431-434: "In the following section, we explore the causes of this apparent resilience in terms of the environmental conditions to which the communities have presumably adapted." It should be noted that the variation in DIC/Alk reflects regional scale variation in single point measurements at each station (Line 362 "...the sampled waters"), and not the DIC/Alk variation at a particular site. Phytoplankton may experience greater or less variation at a single location on a temporal basis, which may be a more important factor determining response. The role of temporal variation in determining response should be discussed.

We agree with the reviewer that variation of carbonate chemistry on a temporal basis may be important in determining the response to experimental OA, although, as discussed in Richier et al. (2018), we also note that it is the range of variability experienced at the cellular level over generation timescales (i.e. days) which is likely to be the most important. However, given the lack of temporal data for each sampling station we acknowledge that we cannot make definitive statements in this regard. We have acknowledged in the text that the amount of variation in carbonate chemistry experienced by plankton communities at a single location on a temporal basis should also be considered:

L495 onwards: "Although it might be expected that carbonate system variability on the level 'experienced' by the cells, i.e. ~daily cellular level variability, might be the most important factor driving sensitivity (Flynn et al. 2012; Richier et al. 2018), our data represent only a snapshot (4 – 6 weeks) of a year, and thus do not contain information on the range in variability over seasonal cycles. For comparison with Arctic stations, Hagens and Middelburg (2016) report a seasonal pH variability of up to 0.25 units from a single site in the open ocean surface waters in the Iceland Sea, whilst Kapsenberg et al. (2015) report an annual variability of 0.3 – 0.4 units in the McMurdo Sound, Antarctica. This implies that both polar open ocean and coastal/sea ice locations experience equally large variations in carbonate chemistry over seasonal cycles. In open ocean waters this is driven by enhanced drawdown of DIC and CO_2 during the productive spring and summer months, countered by lower productivity and strong mixing in the winter (Hagens and Middelburg 2016). In coastal and sea-ice affected regions, seasonal pH variability may be enhanced further by tidal exchanges, and by dilution of T_C/T_A caused by sea-ice melt (Kapsenberg et al. 2015)."

1.11 Lines 451-457. The examples cited to support the authors contention that variability induces plasticity are from coastal waters and under ice, where greater variability would be expected. Would the variability be equally as large at the open ocean stations in this study?

The carbonate chemistry/pH variability may be as large in the open waters of the polar oceans, as in coastal/sea-ice waters. The open waters of the polar oceans experience high levels of productivity during the spring and summer, and given that these waters are less well-buffered with respect to CO_2 uptake, this will lead to a greater range of both short-term cellular scale variability (Flynn et al. 2012; Richier et al. 2018) and seasonal carbonate chemistry characteristics in these waters (Sabine et al. 2004, Orr et al. 2005). Hagen and Middleburg (2016) assess the factors that control seasonal pH variability in surface waters, using a station in the open waters of the Iceland Sea as one of their examples. At this site, pH is shown to vary by up to 0.25 pH units over a seasonal cycle. This is due to a strong drawdown of DIC and pCO_2 during the productive spring months, and a rise in DIC and pCO_2 in winter as a result of reduced productivity and strong

mixing. This is of a similar magnitude to seasonal pH variability at a coastal, sea-ice dominated site in the Antarctic (McMurdo Sound, Kapsenberg et al. 2015) where pH variability of 0.3 – 0.4 units is observed – and this range is among the greatest observed in the ocean. Therefore, it is reasonable to assume that the seasonal variability at open ocean polar stations may be of a similar order to the variability observed in coastal/sea-ice stations.

The text has been edited and extended to take into account the information detailed above as for point 1.10.

Line 462-463 The authors mention the mean state here. Although the inclusion of the Tynan et al (2016) data is useful, this regional variation gives no indication of the local spatial & temporal variation that phytoplankton would experience at each station. The argument would be stronger if the responses were compared with mean local values for the carbonate system (from Tynan et al) for each station, which will to some extent, integrate temporal & spatial variability, rather than using just the values for the water sampled for the experiment (which I assume is what was done).

The DIC/Alk values used for comparison of the response magnitude are representative of the "mean local values" for each station, and are also "the water sampled for the experiment". We are unclear what the reviewer would like to see here. Table 1 shows the pCO_2 and pH for each sampling station and the DIC/Alk was derived from the other measurements of the carbonate system made from these same samples. We do not believe these values integrate temporal and spatial variability as they are discrete measurements made on the water sampled for the incubations. Furthermore community composition is also transient, such that the phytoplankton that were present in the sampled water would not necessarily have been present all the time/everywhere. As emphasised above, our overall hypothesis is that cellular level variability is likely to be the most significant driver of local adaption of communities and hence sensitivity to manipulative forcing (Flynn et al. 2012; Richier et al. 2018).

<u>Technical corrections</u>

1.12 Line 55 chlorophyll-a maxima IN SURFACE WATERS

Text changed accordingly, now reads "…and have been linked to chlorophyll *a* maxima <u>in surface waters</u> and the presence of aerosols formed from DMS oxidation products such as methanesulfonate (MSA)".

1.13 Line 87 Sentence is a bit clunky

Sentence has been re-worded and now reads: "OA is occurring at a rate not seen on Earth for 300 Ma, and so the potential effects on marine organisms, communities and ecosystems could be wide-ranging and severe".

1.14 Line 130-133 Shorten sentence

The sentence has been shortened and now reads: "Nevertheless, our approach can provide insight into the physiological response and level of acclimation to future OA of a variety of polar surface ocean communities adapted to different in situ carbonate chemistry environments (Stillman and Paganini 2015), alongside the implications this may have for DMS production".

1.15 Line 145 – Clarify that the Hopkins & Archer (2014) is from the NW European Shelf

Text changed accordingly, now reads: "Data are combined with the results from an earlier study on board the RRS Discovery (D366) described in Hopkins & Archer (2014) performed in the temperate waters of the NW European shelf".

1.16 Line 256 – What does E1E4/E5 refer to?

We assume the reviewer is referring to Line 265. E1-E4/E5 describes the experiment identifiers for the polar cruises - but incorrectly - so thank you for pointing this out. The text now refers the reader to Table 1 for station identifiers.

1.17 Line 315 "Initial DMSP concentrations were higher AT THE SOUTHERN OCEAN STATIONS than for Arctic stations..."

Text changed accordingly, now reads: "Initial DMSP concentrations were higher at the Southern Ocean stations than for Arctic stations,…"

1.18 Line 317 "Net increases in DMSP occurred throughout, EXCEPT AT SOUTH GEORGIA..."

Text changed accordingly, now reads: "Net increases in DMSP occurred throughout, except at South Georgia, and were on the order of between <10 nmol L^{-1} - >30 nmol L^{-1} over the course of the incubations".

1.19 Line 320 "the final time point at South Georgia (144 h) when a significantly LOWER DMSP with increasing CO2 was observed"

Text changed accordingly, now reads: "Concentrations were not generally *p*CO_2-treatment dependent with the exception of the final time point at *South Georgia* (144 h) when a significantly lower DMSP with increasing CO_2 was observed…"

1.20 Line 350. As the results from the 4 unpublished NW European microcosm experiments are not presented in this paper, they should be identified as unpublished in Table 1

As stated in the text that the reviewer refers to, the data from the 4 previously unpublished NW European microcosm experiments are included in the meta-analysis in this paper (Figures 6, 7, 8), so it is reasonable to identify them as "*This study*" in Table 1.

Line 365; Table 1 legend should identify that the polar stations are the two JR voyages excluding Station NS & IB.

Text added to Table 1 legend: "All polar stations were sampled for JR271 and JR274, with the exception of NS and IB".

1. Response to Anonymous Referee #2

General comments

2.1 Hopkins et al. Present a large dataset on DMS(P) production by phytoplankton in short term OA experiments from the Arctic, the Southern Ocean and the North Atlantic. This is an interesting

and important dataset. I especially acknowledge the importance to publish 'negative results', i.e. absence of significant effects of experimental treatments, which is often neglected in OA research but should receive a lot more attention.

We thank the reviewer for their positive view of our work, and also agree that it is important to publish 'negative results' to give the full picture on the effects of ocean acidification.

2.2 I find the hypothesis that then environmental history of organisms will determine their sensitivity to environmental change very convincing. Currently, the data (or its presentation) is not really suited to convincingly convey this message though. This does not mean that the hypothesis should not be mentioned, but it should be clearly marked as a hypothesis rather than a finding.

We have changed "suggest" to "hypothesise" on L24.

We have changed "Our findings support the notion that,.." to "This supports our hypothesis that…" on L524.

We have changed "However, results from our study indicate that the DMS response to OA…", to "However, we hypothesise that the DMS response to OA…" on L657.

2.3 Furthermore, I would argue that the significant OA effects observed in the two cited coastal communities really question the validity of this hypothesis, as the degree of carbonate chemistry variability is much more pronounced in coastal vs. open ocean compared to temperate vs. polar. Therefore, your conclusions need to be more specific to the current study, and not towards polar systems in general.

We would argue that pH variability over seasonal time scales in the open ocean polar waters is comparable to coastal/sea ice dominated areas. Additionally, as noted elsewhere in our responses, we hypothesise that cellular scale variations are likely to be the most relevant (Flynn et al. 2012; Richier et al. 2018).

See response to reviewer #1, point 1.13.

2.4 One of my general methodological concerns that need to be addressed in the discussion is the fact that especially in short-term experiments, 50% variation in the experiment duration can have a huge impact on the outcome, especially if the phytoplankton initially show a lag phase as often observed in such experiments with natural communities. It makes a huge difference if OA effects are compared after 48h or 4d or 7d. While after 2 days, physiology most likely is not fully acclimated to the treatment conditions yet, 4d or 7d duration most likely show acclimated responses but potentially also reflect shifts in the composition of the communities. Also the differences in the number of hours at T1 and T2 should be accounted for by always referring to the number of hours rather than the time point throughout the manuscript.

For the NW European shelf cruise and the Arctic Cruise, all experiments were 96 h in duration, with samples taken at 0 h, 48 h and 96 h. As we failed to observe strong responses within experiments performed in the Arctic, incubation times were increased for a subset of experiments on the Southern Ocean cruise. This was to investigate whether the lack of strong response in Artic waters was related to slower microbial metabolism in the low temperature waters. To address

this, three experiments of increasingly longer duration from 96 h to 144 h to 168 h were performed (Weddell Sea, South Georgia, South Sandwich), and with the inclusion of a higher target pCO2 of 2000 µatm. However, the magnitude of the response in both biological and DMS-related variables were found to be independent of the experimental duration. See also Richier et al. (2018). This is evidenced by the effect of CO_2 treatment on net growth rates. Net growth rates may be expected to be 2- to 3 –fold higher in the warmest compared to the coldest waters (Eppley, 1972), and indeed maximum Chl a-normalised photosynthesis rates were 2- to 3-fold higher in polar waters – but the response to experimentally-induced OA in polar waters remained insignificant despite the length of incubation being up to twice the duration of temperate experiments, and up to 4x longer, than the 48 h time point where strong responses were typically observed in temperate waters.

The following has been added at L210:

"The magnitude of response was not related to incubation times, and expected differences in net growth rates (2- to 3-fold higher in temperate compared to polar waters (Eppley 1972)) did not account for the differences in response magnitude despite the increased incubation time in polar waters (see Richier et al. (2018) for detailed discussion)".

2.5 It should also be included into the discussion that the significant impacts that Hussherr et al (2017) observed were measured over a much larger pCO2 range (up to 3000 µatm).

We have added to the text at L606:

"It should be noted that this response was seen over a range of pCO_2 from 500 - 3000 µatm, far beyond the levels used in the present study".

We do not feel this warrants any further discussion, as using a gradient of CO_2 treatment levels is an accepted and useful technique in ocean acidification experiments. It allows the use of regression statistics for assessment of possible CO_2 effects, and increases the chances of detecting any threshold level for CO_2/pH sensitive processes (Riebesell et al., 2013, Biogeosciences, 10, 1835–1847).

2.6 One major problem with this dataset is that the experimental carbonate chemistry was not well controlled. For example, at the 1000µatm pCO2 level, T2 pCO2 levels vary between approx. 400 and 1000µatm (Table S2). Therefore, the data should be represented using the real carbonate chemistry instead of the assigned values. I understand that this implies replotting and reanalysing most of the data, but currently the levels that are tested against each other are actually not separated when it comes to carbonate chemistry.

Although we acknowledge that our approach of allowing the carbonate system to vary as a result of biological activity necessitated that some drift occurred following initial imposed conditions, we would argue that our plotting/presentation of the data remains appropriate – and would not agree that this is a "major problem". In Figures 3, 4, 5, 6, 7 and 8 the legend clearly states the CO_2 values shown are "nominal", meaning they are representative of the initial target CO_2 treatments used for each experiment. This is simpler and clearer than showing a complicated range of values for each individual plot and experiment. However, we do concede that we could have presented

the actual treatment levels somewhere in the original submission, and we have now done so, with the addition of a table to the supplementary information (see below).

For the NW European shelf cruise, a comprehensive comparison of the accuracy and precision of the carbonate chemistry manipulation method, as well as an analysis of 'actual' vs 'target' pCO_2 values and the variability in the former over the experimental duration has already been presented in Figure 3, Richier et al. (2014), demonstrating that the achieved pCO_2 levels were well matched to target values at T_0 for E01 – E05, whilst acknowledging that differences in pCO_2 between target and initial values were more pronounced in the higher-pCO_2 treatments, a reflection of the lower buffer capacity of the carbonate system at higher pCO_2.

We have added the following to the methods section (L181 onwards):

"Full details of the carbonate chemistry manipulations can be found in Richier et al. (2014) and Richier et al. (2018). Broadly, achieved pCO_2 levels were well-matched to target values at T_0, although differences in pCO_2 between target and initial values were greater in the higher pCO_2 treatments, due to lowered carbonate system buffer capacity at higher pCO_2. For all 18 experiments, actual attained pCO_2 values were on average around 89% ± 12% (± 1 SD) of target values. The attained pCO_2 values are presented in Table S1 on the Supplementary Information. For simplicity, experimental data is presented against its target ('nominal') pCO_2 treatment throughout the paper".

And we have added this table to the Supplementary Information:

Table S1. Summary of pCO_2 (µatm) and pH_T (total scale) measured immediately following carbonate chemistry manipulation of experimental bioassays (Time point 0, T_0).

		pCO_2 (µatm) at T_0					pH_T at T_0				
Cruise ID	Expt ID	ambient	550 (nominal)	750 (nominal)	1000 (nominal)	2000 (nominal)	ambient	550 (nominal)	750 (nominal)	1000 (nominal)	2000 (nominal)
D366	E01	342.3	564.1	746.4	969.6		8.1	7.9	7.8	7.7	
	E02	n.d.	533.4	n.d.	862.7		n.d.	7.9	n.d.	7.8	
	E02b	n.d.	n.d.	n.d.	n.d.		n.d.	n.d.	n.d.	n.d.	
	E03	345.4	531.2	673.9	877.8		8.1	7.9	7.9	7.8	
	E04	395.4	533.4	691.4	936.6		8.1	7.9	7.8	7.7	
	E04b	n.d.	n.d.	n.d.	n.d.		n.d.	n.d.	n.d.	n.d.	
	E05	374.7	528.9	730.5	917.5		8.1	7.9	7.8	7.7	
	E05b	n.d.	n.d.	n.d.	n.d.		n.d.	n.d.	n.d.	n.d.	
	E06	n.d.	n.d.	n.d.	n.d.		n.d.	n.d.	n.d.	n.d.	
JR271	NS	286.5	524.7	n.d.	620.1		8.2	7.9	n.d.	7.9	
	IB	280.4	434.3	583.3	673.1		8.2	8.0	7.9	7.9	
	GG	326.8	565.2	741.8	1012.2		8.1	7.9	7.8	7.7	
	GI	312.2	583.9	789.3	948.2		8.1	7.9	7.7	7.7	
	BS	310.6	535.1	649.1	683.6		8.1	7.9	7.9	7.8	
JR274	DP	287.0		598.2			8.2		7.9		
	WS	275.1		533.8			8.2		7.9		
	SG	342.6		n.d.	823.4	1410.4	8.1		n.d.	7.7	7.5
	SS	283.8		n.d.	773.2	1557.5	8.2		n.d.	7.8	7.5

For figure 6, 7 and 8, the data is plotted against the carbonate chemistry (C_T/A_T) of the <u>sampled waters (i.e. measurements made before carbonate chemistry manipulation)</u>, not from the incubations, so again the data does not require replotting.

2.7 In conclusion, I get the impression that the authors really try to tell a story that does not fit their data. I think that the hypothesis (more variable carbonate chemistry causes organisms to be less

sensitive) presented here makes a lot of sense, but for various reasons the data set is not suited to prove or disprove it.

We hope that, given the changes we have now made to the manuscript and the clarification we have provided on all the reviewers points, alongside the related analysis within Richier et al. (2018), that the reviewer will now be satisfied that the dataset appropriately addresses the hypothesis posed.

Specific comments

2.8 Title and throughout: To my knowledge, the term "resilience" refers to the ability of a system to return to the initial state after disturbance. Therefore, I do not think that the experimental setup and the response pattern (or its absence) in your study allows for statements on resilience. I suggest to use "insensitivity", "resistance" or something along these lines instead.

L481: "resilience" replaced with "insensitivity"

L507: "resilience" replaced with "the ability to resist"

L589: now reads: "The results of our microcosm experiments suggest insensitivity of *de novo* DMSP production and net DMS production in the microbial communities of the polar open oceans to short term changes in carbonate chemistry".

2.9 L22-27: As you refer to the other studies conducted in the Arctic, you also need to include their results in your statement, or be more specific that you only refer to the presented dataset and not the polar evidence in general.

We have altered the text at L23-28 to take the reviewers comment into account, and emphasise we are specifically referring to the presented dataset:

"Based on our findings, we hypothesise that the differences in DMS response between temperate and polar waters reflect the natural variability in carbonate chemistry to which the respective communities of each region may already be adapted. This implies that future temperate oceans could be more sensitive to OA resulting in a change in DMS emissions to the atmosphere, whilst perhaps surprisingly DMS emissions from the polar oceans may remain relatively unchanged".

2.10 L24-31: In the discussion, you do not refer to "geographical" or "regional" differences but compare temperate vs. polar systems. I would try to be more consistent here.

L23 – 25 now reads: "Based on our findings, we hypothesise that the differences in DMS response between temperate and polar waters reflect the natural variability in carbonate chemistry to which the respective communities of each region may already be adapted".

At L32 we use "geographically distinct regions" and "regionally distinct" in reference to the temperate vs polar waters, which we believe is appropriate.

2.11 Introduction: The introduction is quite long, especially DMS(P) biogeochemistry is described in a lot of detail, even though most of this is not referred to in the discussion. I would suggest to

shorten it. If your discussion does not focus at all on biogeochemistry, do you really need all this detail here?

We disagree with the reviewer. It is important to set the scene, to convince the reader of the importance of DMS, and justify why we are interested in the response of DMS to OA, both in general biogeochemical terms, and specifically with regard to the polar regions of this study.

2.12 L92-95: This is correct, but one shouldn't forget that it is the coastal areas that are the most productive and therefore important ones. In my opinion you do not even have to somehow restrict the importance of these two previous studies, your study is a valuable contribution even though two other ones exist.

We have re-worded the sentence at L96-97 at the reviewer's recommendation:

"However, these two single studies provide limited information on the wider response of the open Arctic or Southern Oceans".

2.13 L118: Here and in a few other instances you refer to your incubations as being "identical", but in the methods you state that the day length was adapted to the respective in situ conditions. Therefore, I would not use the term "identical".

L20: "identical" has been replaced with "similar"

L120: "identical" changed to "near-identical"

L411: "identical" has been deleted

2.14 L119-120: I think the differences in nutrients and incubation temperatures play a big role in understanding the results, so they need to be shown in one of the tables. Referring to a paper under review is not sufficient for such important information. Generally, the authors should provide all relevant information (at least in the supplement) if the other manuscript is not publically available yet.

As indicated above, the paper of Richier et al. (2018) has now been accepted for publication.

In situ temperatures at the time of sample collection are already shown in Table 1. Incubation temperatures were maintained (±1°C) at the in situ value.

Methods text has been adjusted so now reads:

L190: "Bottles were incubated inside a custom-designed temperature- and light-controlled shipping container, set to match (±<1°C) the *in situ* water temperature at the time of water collection (shown in Table 1) (see Richier et al. 2018)".

The nutrients and incubation temperatures did not play a role in understanding the results. We refer to Richier et al. (2018) for more detailed discussion of this, and have added the following to the current manuscript:

L440: "Across all experiments, the response of net total community Chl *a* and net growth rates of small phytoplankton (<10 µm) scaled with pCO_2 treatment, and strongly correlated with in situ carbonate chemistry, whilst no relationships were found with any of the other wide range of initial physical, chemical or biological variables (Richier et al. 2018). Overall, the observed differences in regional response to carbonate chemistry manipulation could not be attributed to any other measured factor that varied between temperate and polar waters. These include ambient nutrient concentrations, which varied considerably but had no influence on the response, and initial community structure, which was not a significant predictor of the response (Richier et al. 2018)".

2.15 L122-125, L130: While I do agree that differences in environmental variability most likely have an impact on the adaptive capacity of communities, you cannot estimate this adaptive capacity in short-term incubation experiments that run for several days only.

See response 1.3 to reviewer #1. Text has been altered accordingly.

2.16 L229-231: I am wondering if it wouldn't make more sense to normalize DMSP concentrations to biomass? This is especially the case if you want to test for "stress-induced algal processes" (L135-136) rather than biomass-dependent effects.

We feel this is not necessary, as we present specific rates of DMSP synthesis. In vivo DMSP synthesis is closely associated with photosynthesis within the cell, so determination of the rate of this process gives an indication of the effects of stress-induced algal processes on DMSP production. This is a much more useful parameter than biomass-normalised DMSP standing stocks, as the DMSP pool is the net results of various and varying processes (see Stefels et al. 2009), with variable contributions to DMSP production by different groups of phytoplankton.

2.17 L252-259: I do not think that you can infer growth rates from the Chla measurements, given that there was probably strong photoacclimatory processes happening in response to the change in light fields (naturally varying to constantly high). You do not really need these rates for your story, so I suggest to omit this parameter all together, i.e. also from results and discussion.

At the reviewer's suggestion we have removed relative growth rates from the paper.

2.18 L278: The results from the Atlantic experiments are used a lot in the discussion, they should therefore also be included in the results (and methods), especially but not exclusively the previously unpublished ones.

We have now described the methods and results from the 6 previously unpublished experiments in temperate waters.

Some minor adjustments have been made to the methods text to account for the temperate experiments, but the reader is generally referred to the related studies for the full details (Hopkins and Archer 2014, Richier et al. 2014, Richier et al. 2018):

L151: "Additionally, four previously unpublished experiments from D366 are also included (E02b, E04b, E05b, E06) as well as two temperate experiments from JR271 (NS and IB) (see Table 1)".

L197: "For Southern Ocean and all temperate water stations, an 18:6 light: dark cycle was used".

L202: "Experiments were generally run for ≥4 days (15 out of 18 experiments), with initial sampling proceeded by two further time points. For three temperate experiments (E02b, E04b, E05b, see Table 1) a shorter 2 day incubation was performed, with a single sampling point at the end. For E06 (see Table 1) high time frequency sampling was performed (0, 1, 4, 14, 24, 48, 72, 96 h) although only the data at 48 h and 96 h is considered in this analysis".

Figure 2 now includes depth profile data from all 18 sampling stations, and the results text now includes full description of the data for all 18 stations:

L301: "At temperate sampling stations, sea surface temperatures ranged from 10.7°C for *Iceland Basin*, to 15.3°C for *Bay of Biscay*, with surface salinity in the range 34.1 – 35.2, with the exception of station E05b which had a relatively low salinity of 30.5 (Figure 2 and Table 1)".

L312: "Chl *a* concentrations in temperate waters ranged from 0.3 μg L^{-1} for two North Sea stations (*E05* and *North Sea*) up to 3.5 μg L^{-1} for *Irish Sea* (Figure 2 and Table 1). Chl a was also variable in polar waters, exceeding 4 μg L^{-1} at *South Sandwich* and 2 μg L^{-1} at *Greenland Ice-edge*, whilst the remaining stations ranged from 0.2 μg L^{-1} (*Weddell Sea*) to 1.5 μg L^{-1} (Figure 2)".

L318: "In temperate waters, maximum DMS concentrations were generally seen in near surface measurements, ranging from 1.0 nM for *E04* to 21.1 nM for *E06*, with rapidly decreasing concentrations with depth (Figure 2 G). DMSP also generally peaked in the near surface waters, ranging from 12.0 nM for E04 to 72.5 nM for *E06*, but the maximum overall DMSP concentration of 89.8 nM was observed at ~20 m for *E05b* (Figure 2 H). Surface DMS concentrations in polar waters were generally lower than temperate waters, ranging from 1 – 3 nmol L^{-1}, with the exception of *South Sandwich* where concentrations of ~12 nmol L^{-1} were observed (Figure 2 G)".

The DMS and DMSP results from the 6 previously unpublished temperate microcosm experiments are now shown in the Supplementary Information, in Table S4 (E02b, E04b, E05b, E06 from D366) and Figure S2 (*North Sea* and *Iceland Basin* from JR271), and described in brief in the results section:

L355: "Results from the previously unpublished experiments from temperate waters are in strong agreement with the five experiments presented in Hopkins and Archer (2014), with consistently decreased DMS concentrations and enhanced DMSP under elevated CO_2. The data is presented in the Supplementary Information, Table S4 and Figure S2, and included in the meta-analysis in section 4.1 of this paper".

Table S4. DMS and DMSPt response (mean ± SD, n = 3) to high CO_2 treatments during previously unpublished small-scale experiments from the NW European shelf cruise D366. For details of stations, see Table 1 in the main paper.

	0 h ambient	**48 h ambient**	**48 h 550 µatm**	**48 h 750 µatm**
DMS (nM)				
E02b	2.4 ± 0.3	2.1 ± 0.6		2.7 ± 0.6
E04b		6.4 ± 1.4		14.7 ± 8.1
E05b		3.3 ± 0.1		4.5 ± 0.6
E06	18.7 ± 0.5	18.1	24.2	25.2
DMSPt (nM)				
E02b		49.5 ± 2.0		26.4 ± 2.9
E04b		68.2 ± 10.3		36.8 ± 7.5
E05b		48.7 ± 11.2		37.4 ± 4.8
E06	76.7 ± 5.7	114.6	98.43	108.5

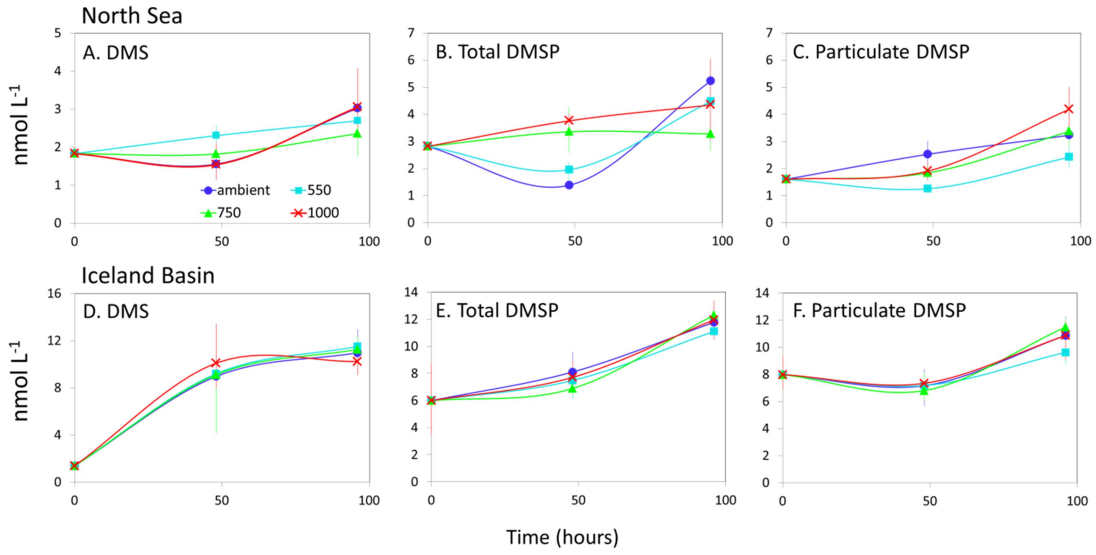

Figure S2. DMS, total DMSP and particulate DMSP concentrations (nmol L^{-1}) during experimental microcosms performed in temperate waters at stations *North Sea* and *Iceland Basin* from cruise JR271. Data shown is mean of triplicate incubations, and error bars show standard error on the mean. Locations of water collection for microcosms are given in Table 1.

2.19 L284-287: Methods are missing for the nanoflagellate and bacteria abundances data,

The following has been added to the methods section (L278):

"2.8 Community composition

Composition of small phytoplankton and bacteria community composition was assessed by flow cytometry. For details of methodology, see Richier et al. (2014)".

2.20 L291: Methods for irradiance measurements are missing

Text has been added to the methods at L162:

"At each station, pre-dawn vertical profiles of temperature, salinity, oxygen, fluorescence, turbidity and irradiance were used to choose and characterise the depth of experimental water collection".

2.21 L314: This is important information that really helps your line of argument, I would therefore put stronger emphasis on this in the discussion.

It is slightly unclear but does the reviewer refer to the statement: "Significant differences ceased to be detectable by the end of the incubations (168 h)"?

The section referred to describes the DMS response within the *South Sandwich* experiment which showed a significant CO_2 treatment effect after 96 h of incubation, which then ceased to be detectable by the end of the experiment (172 h). This single time point measurement out of the 7 polar experiments is an exception to the general overall trend of no DMS response – it was also accompanied by almost identical mean DMSP concentrations under all CO_2 treatments (Figure 4 G). Therefore it is difficult to gauge the significance of this result. Once this data is combined into the meta-analysis, it is clear this DMS response at *South Sandwich* is negligible compared to the magnitude of responses we saw in temperate waters (Figure 6 A).

2.22 L328-335: This comparison of standing stocks is highly dependent on the time of sampling. You therefore need to include information about and discussion on the timing of sampling relative to bloom phenology. I.e. if the Arctic and Southern Ocean samples were taken in (macro and/or micro) nutrient depleted waters after a bloom, can you really make such general statements on polar vs. temperate waters? Was the temperate sampling also conducted in similar phases of biomass dynamics? If not, you have a problematic bias towards low productivity in the polar samples that needs to be taken into account.

Coherent responses to OA occurred regardless of initial conditions, in terms of both the general biological response, and the DMS/P response. See Richier et al. (2018) for an assessment of the observed responses in comparison to a range of initial environmental. Importantly, differences in net phytoplankton growth rates as a function of pCO_2 treatment showed no correlation with any of the other wide range of initial physical, chemical and biological variables tested, including nutrient concentrations. Initial community structure was not an important factor in determining responses to pCO2 treatment (Richier et al. 2018). Thus although it is likely that we sampled waters that were at different stages of bloom phenology, this did not appear have an influence on our findings. Indeed, we note that a wide range of initial chlorophyll standing stocks was sampled on both high latitude cruises (Table 1). Overall, the most important factor influencing the biological, and DMS/P response to elevated pCO_2, was the carbonate chemistry characteristics of the sampled waters. Thus, our findings suggest that both the organism and ecosystem level response to OA is related to variability in the mean state of the carbonate chemistry system, alongside the associated natural range of variability in carbonate chemistry experienced by organisms (Flynn et al. 2012; Richier et al. 2018). Both these factors are likely linked to regional variability in the buffering capacity of ocean waters.

The lower rates of DMSP synthesis in polar waters compared to temperate waters is not necessarily due to lower levels of productivity but rather 'slower metabolic processes' as we state in this section. We compare our results to 'non-bloom conditions' from the Archer et al. (2013) paper because although higher rates were observed during this study (10 – 15 nmol L^{-1} d^{-1}), they occurred following artificial addition of inorganic nutrients to the mesocosms, which is not comparable to open ocean rates measured during the OA microcosm experiments.

We addressed the issue of 'slower metabolic rates' in colder polar waters in response 1.7 to reviewer #1.

2.23 L340-342: This in a strong indication for the importance of other drivers (nutrients, species composition, ...). You need to show these and check whether there are significant effects here.

Whilst we agree it would be interesting to attempt to unravel what is driving any differences in DMSP production in polar waters, we feel that it is outside of the scope of this paper.

2.24 L360ff: I really like this way of presenting the data. You should, however, also show the same plot with pCO2 instead of TA/DIC for comparison because I do not agree with you that this ratio gives a full overview of the in situ carbonate chemistry.

We feel it would not be a useful exercise to replot all the data against pCO_2, based on single discrete measurements. The relevance of this value is unclear, as it can be so variable in space and time.

We use DIC/Alk as it is the simplest way of representing the buffer capacity of the sampled waters. We could also have plotted against the Revelle factor of the sampled waters, and the relationship would have looked almost identical, as the Revelle factor is indeed a function of DIC/Alk, and quantified the ocean's sensitivity to an increase in CO_2. Therefore we believe that DIC/Alk is the simplest and more appropriate way of visualising our data in terms of its geographical location.

2.25 L372 and throughout the entire manuscript: Report the time points in days or hours instead of T1, T2 etc. because this is not consistently the same time point as well as for better readability and consistency throughout the text.

Reporting the time points in hours throughout the manuscript would make the results text clunky and confusing to read, as there is some variability. Therefore, we will keep the T_1/T_2 notation, but refer to the times broadly and refer the reader to Table 1 which outlines all the specific sampling times.

Added at L201: "(T_1, T_2, see Table 1 for specific times for each experiment)"..,

L410 now reads: "…was minimal at all sampling points…"

L413 now reads: "…particularly at T_1 (48 – 96 h)…"

Figure 7 legend text has been altered: "T_1 = 48 h, T2 = 96 h, except for *Weddell Sea* and *South Georgia* (72 h, 144 h)".

Figure 8 legend text has been altered: "T_1 = 48 h, T_2 = 96 h, except for *Weddell Sea* and *South Georgia* (72 h, 144 h) and *South Sandwich* (96 h, 168 h)".

2.26 L377-282: This strongly suggests that, due to temperature-driven differences in metabolic rates and their effects on how fast the communities can acclimate to changed conditions, the experiments emerge out of measurement noise at different times.

If the differences were driven by temperature, rates would be expected to be higher in warmer temperate waters and lower in cold polar waters – but this is not the case. Also see response 1.6 to reviewer 1.

To make our point more clearly, the text now reads (L413):

***"De novo* DMSP synthesis and DMSP production rates show a similar relationship with C_T/A_T (Fig. 7 A and B), with a significant suppression of DMSP production rates in temperate waters compared to polar waters (Fig. 7B, Kruskal-Wallis One Way ANOVA H = 8.711, df = 1, p = 0.003). Although a similar trend was seen for *de novo* DMSP synthesis, the difference between temperate and polar waters was not statistically significant (Fig. 7A)".**

Therefore, this suppression of rates in temperate waters is likely related to the relative decreases in net growth (Chl a accumulations, phytoplankton cell counts, community biomass) seen in temperate waters (see also Richier et al. 2018).

2.27 Discussion: A discussion of stress vs. acclimated response is missing

The following text has been added to the discussion at L631:

"Our results imply that the phytoplankton communities of the temperate microcosms initially responded to the rapid increase in pCO_2 via a stress-induced response, resulting in large and significant increases in DMS concentrations occurring over the shortest timescales (2 days), with a lessening of the treatment effect with an increase in incubation time (Hopkins and Archer 2014). Within non-nutrient amended treatments such a reduction in response with time may also have been driven by nutrient exhaustion, which could have lead the system to a similar state across all CO_2 treatments, although we note that carbonate chemistry manipulation induced responses were also similar within nutrient amended treatments (Richier et al. 2014, 2018). The dominance of short response timescales in well-buffered temperate waters may also indicate rapid acclimation of the phytoplankton populations following the initial stress response, which forced the small-sized phytoplankton beyond their range of acclimative tolerance and lead to increased DMS (Richier et al. 2018, Hopkins and Archer 2014).

This supports the hypothesis that populations from higher latitude, less well-buffered waters, already possess a certain degree of physiological tolerance to variations in carbonate chemistry environment. Although initial community size structure was not a significant predictor of the response to high CO_2, it is possible that a combination of both community composition and the natural range in variability in carbonate chemistry – as a function of buffer capacity – may influence the DMS/P response to OA over a range of timescales (Richier et al. 2018)".

2.28 L399: Everything until here reads more like results than like a discussion section. Please consider rearranging.

The section the reviewer refers to describes the meta-analysis of all the data from the 3 cruises – we believe this can be considered suitable content for the discussion and have left it unchanged. It should be viewed as a synthesis rather than a simple description of results.

2.29 L410-412: The authors seem to imply that CO2 sensitivity is only occurring in form of negative effects, even though there are many studies that show beneficial effects of increased substrate availability for photosynthesis, which is particularly true for picoeukaryotes (e.g. Schulz et al. 2017). Please take this aspect into account.

The section the reviewer refers to does not imply this. Rather, here we provide an explanation for our observations.

2.30 L436-439: I do not agree that your data really shows this: Figure 9 indicates the Arctic Ocean carbonate chemistry to be actually more similar to the Atlantic than to the Southern Ocean.

We disagree. The data show that the variability in carbonate chemistry in the polar oceans is much larger than in temperate waters – as described in the text the reviewer refers to.

2.31 L444-448: Such a comparison only makes sense if the same geographical and temporal ranges, and phases of biomass cycle (pre-bloom/ bloom/post-bloom, before/after winter convection etc.) were covered in the different study areas. Please clarify if this was the case.

See response 1.12 to reviewer #1 with regard to accounting for the possible variability in pH over seasonal scales.

2.32 L451-455: In the Southern Ocean, several studies have shown strong OA-effects on species composition (e.g. Tortell et al. 2008, Feng et al. 2010, Hoppe et al. 2013, Trimborn et al. 2017). L455-457: Similarly, you are missing previous work done in the Arctic (Coello-Camba et al. 2014, Holding et al. 2015, Thoisen et al. 2015, Hoppe et al. 2017a,b) that need to be considered.

We have re-worded the suggested line so that it no longer implies that this is all the available data.

L512:"A number of previous studies in polar waters have reported similar findings". However, we believe it would over-complicate this part of the discussion, and disrupt the flow, if we were to bring in the other suggested references at this point in the paper.

L519: We have added in two recently published study which provides further substantiation for our hypothesis, "Subarctic phytoplankton populations demonstrated a high level of resilience to OA in short term experiments, suggesting a high level of physiological plasticity that was attributed to the prevailing strong gradients in pCO_2 levels experienced in the sample region (Hoppe et al. 2017). Furthermore, a recent study describing ten CO_2 manipulation experiments in Arctic waters found that primary production was largely insensitive to OA over a large range of light and temperature levels (Hoppe et al. 2018). This supports our hypothesis that, relative to temperate communities, polar microbial communities may have a high capacity to compensate for

environmental variability (Hoppe et al. 2018), and are thus already adapted to, and are able to tolerate, large variations in carbonate chemistry".

Therefore, we have made reference to these studies later in the paper at the end of section 4.4, where we already provide evidence that the DMS response is likely to be variable over temporal and spatial scales:

L611: "Furthermore, a number of other studies from both the Arctic e.g. (Coello-Camba et al. 2014; Holding et al. 2015; Thoisen et al. 2015) and the Southern Ocean e.g. (Tortell et al. 2008; Hoppe et al. 2013; Trimborn et al. 2017) suggest that polar phytoplankton communities can demonstrate sensitivity to OA, in contrast to our findings. This emphasises the need to gain a more detailed understanding of both the spatial and seasonal variability in the polar phytoplankton community and associated DMS response to changing ocean acidity".

2.33 L460: n=3 is not "highly" replicated

"Highly" has been omitted.

2.34 L469: Why are you comparing your data in detail with Archer et al. (2013) but not Hussherr et al. (2017)?

We use this section to highlight the differences between experimental approaches, as it is useful for the reader to understand why we might see such different results between microcosm experiments and mesocosm experiments. We have altered the text to make this point come across more clearly to the reader:

"Experimental data clearly provide useful information on the potential future DMS response to OA, but these data become most powerful when incorporated in Earth System Models (ESM) to facilitate predictions of future climate. To date, two modelling studies have used ESM to assess the potential climate feedback resulting from the DMS sensitivity to OA (Six et al. 2013; Schwinger et al. 2017), and both have used results from mesocosm experiments. However, the DMS responses to OA within our short term microcosm experiments contrast with the results of most previous mesocosm experiments, and of particular relevance to this study, an earlier Arctic mesocosm experiment (Archer et al. 2013). Whilst no response in DMS concentrations to OA was generally seen in the microcosm experiments discussed here, a significant decrease in DMS with increasing levels of CO_2 in the earlier mesocosm study was seen. Therefore, it is useful to consider how the differences in experimental design between microcosms and mesocosms may result in contrasting DMS responses to OA".

As Hussherr et al. also used a microcosm approach, we include a short comparison at the end of this section (L594 onwards) to emphasise that discrepancies can also occur even when using similar experimental techniques.

2.35 L475: I would rather refer to the most common not the maximum duration.

"maximum of…" has been omitted.

2.36 L482-488: Is this difference really due to different sensitivities, or differences in biological rates, that lead to the fact that small physiological changes are detectable at different time points?

It is unlikely that biological rates will vary significantly between mesocosms and microcosms, as each experimental system should be a reasonable representation of the natural system which was sampled. Although there is not total certainty that it is due to differences in sensitivities, this is a hypothesis we put forward to help explain the differences in response to OA that we observe between these experimental approaches.

2.37 L515-521: You first imply that the short duration of the experiments would render changes in species composition rather unlikely, but then you report one case where you indeed observed changes. I would say that this indicates that the timescales in general would have allowed for changes in composition also in the other experiments.

Text has been altered and now reads (L589): "We did not generally see any broad-scale CO_2-effects on community structure in polar waters. This can be demonstrated by a lack of significant differences in the mean ratio of >10 µm Chl *a* to total Chl *a* (*>10 µm : total*) between CO_2 treatments, implying there were no broad changes in community composition (Table 2). *South Sandwich* was an exception to this, where large and significant increases in the mean ratio of *>10 µm : total* were observed at 750 µatm and 2000 µatm CO_2 relative to ambient CO_2 (ANOVA, $F = 207.144$, $p<0.001$, $df = 3$), demonstrated at even the short timescale of the microcosm experiments, it is possible for some changes to community composition to occur".

2.38 L543-550: I agree that it is an interesting finding that coastal DMS production seems to be more sensitive to OA than that from the open ocean. This finding does, however, really hint against the proposed mechanisms of insensitivity, because coastal systems are a lot more variable in carbonate chemistry compared to the open ocean (e.g. Thoisen et al. 2015). Thus, the interpretation of and conclusions from the dataset have to be reassessed.

Given the reviewers comments on this issue, we believe that the comparison between 'coastal' and 'open ocean' waters complicates this part of our discussion, so we have removed mention of this comparison. We instead discuss the possibility that there is likely to be regional variability in the response of DMS to OA. The key point is that the DMS response to OA in polar regions is complex and likely to be influenced by a number of temporal and spatial factors. The main users of our data are climate modellers, and we wish to emphasise that when trying to model the future flux of DMS, it is important to take this variability into account. The section now reads (L623):

"Our findings contrast with two previous studies performed in Arctic waters (Archer et al. 2013, Hussherr et al. 2017) which showed significant decreases in DMS in response to OA. These discrepancies may be driven by differences in the sensitivity of microbial communities to changing carbonate chemistry between different areas, or by variability in the response to OA depending on the time of year, nutrient availability, and ambient levels of growth and productivity. This serves to highlight the complex spatial and temporal variability in DMS response to OA which warrants further investigation to improve model predictions".

2.39 Figures 3, 4, 5, 7, 8, S3: Given the lack of control in carbonate chemistry in many experiments (Table S2), this representation is misleading. The data needs to be presented accounted for the real carbonate chemistry in the incubations.

We have addressed this above in point 2.6.

Technical Corrections

2.40 L11: I suggest replacing "we increase" by "to increase"

Text changed accordingly.

2.41 L12: I suggest referring explicitly to climate change instead of environmental change. Otherwise, the step to OA is kind of abrupt.

Done.

2.42 L28: Do you really mean "region may vary in response to OA" or rather "region may vary in their response to OA"?

Text changed accordingly, now reads: "By demonstrating that DMS emissions from geographically distinct regions may vary in their response to OA,…"

2.43 L190: replace "made" by "taken"

Done.

2.44 L207: omit "all" as in the caption of figure 5 you state that these data are not available for two of the stations.

Text changed accordingly, now reads: "*De novo* DMSP synthesis and gross production rates were determined for all microcosm experiments, except *Barents Sea* and *South Sandwich*,…".

2.45 L237-238: According to the Journal style, it would be A_T and C_T for total alkalinity and total dissolved inorganic carbon, respectively

We have changed T_A to A_T and DIC to C_T throughout.

2.46 L372: Omit "identical" as irradiances and temperatures were not the same

Done.

2.47 L497-500: Something does not see correct in this sentence, please rephrase

This sentence has been rephrased (L571): "Moreover, the coastal Arctic mesocosms were enriched with nutrients after 10 days, affording relief from nutrient limitation and allowing differences between pCO_2 treatments to be exposed, including a strong DMS(P) response".

2.48 L532: Insert "low and" between "periods of" and "stable productivity"

Done.

2.49 L539: "is insensitive to OA during multiple short term microcosm" instead of "is resilient to OA during multiple, highly replicated short term microcosm"

Done.

2.50 L542: add additional references mentioned above

We have instead removed reference to Davidson et al. (2016) as this was incorrectly cited here, and only refer to the results from our own study (which was the intention).

2.51 L559: Replace "results from our study indicate" by a more honest "we hypothesise" or something similar.

Done.

2.52 Table 1: Add macro nutrient (at least NO3) levels and incubation temperatures (will be more variable than in situ). Also "Comment" should read "Reference". Shouldn't "Sample depth" read "Sampling depth"?

The temperate of the incubation container was maintained at the in situ sampling temperature (±<1°C) (see methods in Richier et al. 2014, Biogeosciences, doi:10.5194/bg-11-4733-2014). Methods text (L191) has been altered to confirm this:

"Bottles were incubated inside a custom-designed temperature- and light-controlled shipping container, set to match the *in situ* water temperature at the time of water collection (±<1°C, see Richier et al. 2018)".

Nitrate concentrations have been added to the Table as suggested. And other suggested changes have been made.

2.53 All Figures: Please indicate number of replicates and type of error estimate in the caption

Now included in figure captions for Figure 3 and Figure 4:

"Data shown is mean of triplicate incubations, and error bars show standard error on the mean".

2.54 Figure 2: Replace "µE m-2 s-1" by "µmol photons m-2 s-1" or "µmol quanta m-2 s-1" in figure and caption. Also, the panels are so close together that the top and bottom axis descriptions get messy, please move them apart a bit.

Done.

References

Archer, S. D., S. A. Kimmance, J. A. Stephens, F. E. Hopkins, R. G. J. Bellerby, K. G. Schulz, J. Piontek and A. Engel (2013). "Contrasting responses of DMS and DMSP to ocean acidification in Arctic waters." Biogeosciences **10**(3): 1893-1908.

Coello-Camba, A., S. Agustí, J. Holding, J. M. Arrieta and C. M. Duarte (2014). "Interactive effect of temperature and CO2 increase in Arctic phytoplankton." Frontiers in Marine Science **1**: 49.

Eppley, R. W. (1972). "Temperature and phytoplankton growth in the sea." Fish. bull **70**(4): 1063-1085.

Flynn, K. J., J. C. Blackford, M. E. Baird, J. A. Raven, D. R. Clark, J. Beardall, C. Brownlee, H. Fabian and G. L. Wheeler (2012). "Changes in pH at the exterior surface of plankton with ocean acidification." Nature Climate Change **2**(7): 510-513.

Hagens, M. and J. J. Middelburg (2016). "Attributing seasonal pH variability in surface ocean waters to governing factors." Geophysical Research Letters **43**(24): 12,528-512,537.

Holding, J. M., C. M. Duarte, M. Sanz-Martin, E. Mesa, J. M. Arrieta, M. Chierici, I. E. Hendriks, L. S. Garcia-Corral, A. Regaudie-de-Gioux, A. Delgado, M. Reigstad, P. Wassmann and S. Agusti (2015). "Temperature dependence of CO2-enhanced primary production in the European Arctic Ocean." Nature Clim. Change **advance online publication**.

Hoppe, C. J., N. Schuback, D. M. Semeniuk, M. T. Maldonado and B. Rost (2017). "Functional Redundancy Facilitates Resilience of Subarctic Phytoplankton Assemblages toward Ocean Acidification and High Irradiance." Frontiers in Marine Science **4**: 229.

Hoppe, C. J. M., C. S. Hassler, C. D. Payne, P. D. Tortell, B. Rost and S. Trimborn (2013). "Iron Limitation Modulates Ocean Acidification Effects on Southern Ocean Phytoplankton Communities." PLOS ONE **8**(11): e79890.

Hoppe, C. J. M., K. K. E. Wolf, N. Schuback, P. D. Tortell and B. Rost (2018). "Compensation of ocean acidification effects in Arctic phytoplankton assemblages." Nature Climate Change **8**(6): 529-533.

Kapsenberg, L., A. L. Kelley, E. C. Shaw, T. R. Martz and G. E. Hofmann (2015). "Near-shore Antarctic pH variability has implications for the design of ocean acidification experiments." Scientific Reports **5**: 9638.

Richier, S., E. P. Achterberg, M. P. Humphreys, A. J. Poulton, D. J. Suggett, T. Tyrrell and C. M. Moore (2018). "Geographical CO_2 sensitivity of phytoplankton correlates with ocean buffer capacity." Global Change Biology.

Schwinger, J., J. Tjiputra, N. Goris, K. D. Six, A. Kirkevåg, Ø. Seland, C. Heinze and T. Ilyina (2017). "Amplification of global warming through pH-dependence of DMS-production simulated with a fully coupled Earth system model (under review in Biogeosciences, doi: 10.5194/bg-2017-33)." Biogeosciences.

Six, K. D., S. Kloster, T. Ilyina, S. D. Archer, K. Zhang and E. Maier-Reimer (2013). "Global warming amplified by reduced sulphur fluxes as a result of ocean acidification." Nature Climate Change **3**(11): 975.

Stillman, J. H. and A. W. Paganini (2015). "Biochemical adaptation to ocean acidification." Journal of Experimental Biology **218**(12): 1946-1955.

Thoisen, C., K. Riisgaard, N. Lundholm, T. G. Nielsen and P. J. Hansen (2015). "Effect of acidification on an Arctic phytoplankton community from Disko Bay, West Greenland." Marine Ecology Progress Series **520**: 21-34.

Tortell, P. D., C. D. Payne, Y. Li, S. Trimborn, B. Rost, W. O. Smith, C. Riesselman, R. B. Dunbar, P. Sedwick and G. R. DiTullio (2008). "CO_2 sensitivity of Southern Ocean phytoplankton." Geophysical Research Letters **35**.

Trimborn, S., T. Brenneis, C. J. M. Hoppe, L. M. Laglera, L. Norman, J. Santos-Echeandía, C. Völkner, D. Wolf-Gladrow and C. S. Hassler (2017). "Iron sources alter the response of Southern Ocean phytoplankton to ocean acidification." Marine Ecology Progress Series **578**: 35-50.