Please find below a point by point response to each of the reviewers' comments on our manuscript entitled 'High tolerance of protozooplankton to ocean acidification in an Arctic coastal plankton community' authored by Nicole Aberle et al.. We are grateful for the reviewers' comments and suggestions and appreciate the efforts made to improve the ms.

Anonymous Referee #1

Received and published: 20 October 2012

General

This manuscript presents results from a large coastal mesocosm experiment in Svalbard in which pCO2 and pH were manipulated and a phytoplankton bloom triggered by addition of nutrients in order to examine the impact of ocean acidification (OA) on these planktonic communities. As part of a larger study, here the results concerning the protozooplankton are presented. As a remarkable result, the acidification did not seem to have any effect, neither on the biomass of the different groups nor on the taxonomic composition or population dynamics. This can be interpreted as a high tolerance and might be interesting for the readers of Biogeosciences. However, for a more comprehensive understanding of what was going on in the mesocosms, it needs some additional information and explanations. There is some confusion in the terms that are being used but not properly explained (protozooplankton vs. phytoplankton), missing statistics and some additional data that are necessary to judge the results of this study. Therefore, although the paper is well written it requires a thorough revision (outlined in more detail below).

Specific comments

Introduction

p.3, l. 5ff: do you focus in this paper only on herbivorous protists not on bacterivorous? This paragraph gives this impression but then it should explicitly be mentioned.

Yes, we focussed on those species which might have an impact on phytoplankton standing stocks (e.g. Protoperidiunium spp. Gyrodinium spp.). In a new ms version we would clarify this and state that e.g. HNFs were not included in our analysis (but this is done by Corina Brussard et al, see her ms submitted to the same BG special issue).

p.3, l.13ff: I think here also the impact of OA on calcifying planktonic protists (also coccolithophores are protists!) should be mentioned

You are right- but there were no calcifying protists (e.g. Coccolithophores) present. But we might include this aspect in the discussion section on the impacts of OA on protist communities.

Clarification of terms: Different and partly overlapping terms are used in this paper and the introduction should clarify their meanings: what is "protozooplankton" compared to "phytoplankton"? I think a large portion of the phytoplankton are indeed protists! And dinoflagellates can be herbivorous consumers but also primary producers. So how are the terms differentiated throughout this study? Were you differentiating according to their nutritional modes? But how were theses assessed? Can be very confusing!

We agree that the phytoplankton/protozooplankton classification we gave in the first ms version was poor. In a new version we will clarify both terms by giving clear definitions. All dinoflagellates we included into the protozooplankton (mainly Protoperidiunium spp., Katodinium spp., Gyrodinium spp.) are considered to have predominantly heterotrophic feeding modes. Some small chloroplast-bearing dinoflagellates were included into the phytoplankton since they were considered as predominately autotrophic. We are aware of the problem of putting dinoflagellates into categories but since most dinoflagellates have to date been observed to actively capture prey items (even those which were formerly considered as exclusively autotrophic), categorizing most dinoflagellates into a heterotrophic or mixotrophic feeding mode seems accurate.

Table A1 suggests that only ciliates and dinoflagellates were assessed as "protozooplankton". In this case the title and the terms throughout the paper should be changed, e.g. to "herbivorous ciliates and dinoflagellates" or "microzooplankton" but not "protozopoplankton" in general as this would also include heterotrophic nanoflagellates (which are not considered here) and probably also some of the phytoplankton components.

We will replace the term 'protozooplankton' by 'microzooplankton' in order to avoid any confusions.

Methods

p.4, pCO2 levels: please provide not only the average values but also a measure of the temporal variability of the pCO2 concentrations and the pH in different mesocosms (e.g., in the supplement)

We will provide a table with the pCO2 concentrations over the whole duration of the experiment when encouraged by the BG editor to submit a new version of the ms.
p.5, protozooplankton-phytoplankton enumeration: see my comment above: what is a "phytoplankton", what a "protzooplankton" (e.g., an autotrophic dinoflagellate). Further, I do not think that the lugol-fixed samples are appropriate to distinguish auto- and heterotrophic taxa. Was this ignored?
As already mentioned above, most dinoflagellates were considered as hetero- or mixotrophic. Only some small chloroplast-bearing dinoflagellates were included into the phytoplankton since they were considered as predominately autotrophic.
Results
It would be nice to have not only biomasses but at some point also cell numbers of the major protist groups reported here, as only those numbers can be directly compared to data found in other systems. I suggest to extend for example Table A1 for this (or provide an additional table).
In a new ms version we will be happy to include data on cell numbers by providing an additional table.
Fig. 2 is awful to read and I suggest to take out the phytoplankton lines and put them in extra plots.
We will take out the phytoplankton lines and provide an additional figure just showing the phytoplankton development.
Same with Fig. 3, here at least the lines should be made in colour in order to be able to differentiate the different treatments.
We will colour the lines to differentiate between treatments and make the graph more easy to understand.
Statistics: In the methods it is stated that a regression analysis was used to test for OA effects on protozooplankton. I have not seen any results of those tests! Has this been forgotten entirely? I think, both the composition and biomass as well as the temporal dynamics can be tested with this experimental design.
We tested for significant correlations between the different pCO2 treatments and microzooplankton diversity, peak biomass magnitudes, total biomass and cumulative biomass. For all these analyses we did not find any positive nor negative relationship with pCO2. We apologize for not giving details on the regression analyses and for not providing any results data. In a new ms version we will include a table on the regression results and describe clearly the statistical analyses conducted and the outcome.
Discussion
I think much more focus should be put on species-specific effects of OA. Even if this study could not demonstrate effects of OA on the protist communities present here, this does not mean that nowhere any effects might be expected. Try to gather all relevant data from the literature in order to discuss potential different vulnerabilities of different protist groups.
The idea of including more details on species-specific effects of OA e.g. from laboratory experiments dealing with only few species and not entire communities will be taken into account when handing in a revised version to BG. We tried to include some of the experimental data on specific ciliate and dinoflagellate species already (Pedersen et al. 2003, Pedersen & Hansen 2003) but will extend the discussion section focussing more strongly on species-specific impacts of OA.
The discussion makes it clear that other factors, mainly trophic interactions, might have a more profound shaping effect on PZP compared to pCO2. This is an important aspect but the discussion of food quality effects (p.12) seems to me too speculative and not supported by data shown here and therefore should be omitted or strongly shortened.
As suggested by Rev. 1 (and also Rev. 2) we will shorten the discussion section on food quality aspects and omit any speculations and hypothesis based on trophic cascades and food quality-related aspects.

Anonymous Referee #2
Received and published: 4 December 2012
Review of Aberle et al. 'High tolerance of protozooplankton to ocean acidification: : :' submitted to BGD
The work described in the manuscript focus on potential effects on ocean acidification on heterotrophic dinoflagellates and ciliates. Emphasis is given to estimates of direct effects on community/diversity and their implications in stoichiometry/ phenology/carrying capacity. The topic is of great research interest, especially as the response is addressed in a more complex mesocosm environment. The authors found little evidence for any major effects and conclude that protozooplankton is relatively

insensitive to acidification. Although I can follow this interpretation, I find the data basis quite poor. Addressing the raised questions requires a frequent sampling
and thorough estimate of sampling variability, which is not given here. Often the interpretation
is based on only 1-2 sampling points during one of the 3 bloom phases and
conclusions of 'increasing'/'decreasing' or 'peak biomass' are made without knowledge
of sampling variability. This is highly critical considering the low abundance of protists
and its inherent counting variability. I have my doubts if any quantitative conclusions
can be made at all.

We are aware of the problem that the mesocosm set-up in our experiment had some shortcomings since replicates of the different pCO2 treatment were missing. Therefore, we understand the concerns raised by Rev. 2 in parts. However, the mesocosm set-up was a decision of the whole experimental team and it was agreed upon by the majority of partners that the experiment should span a broad range of pCO2 concentrations rather than e.g. 3 concentrations but with replicate mesocosms. This is a shortcoming of the experiment we had to deal with. The samples obtained for PZP enumeration were taken from a depth integrating water sampler (0-12 m). After a gentle but thorough mixing, 100 ml of fixed sample was settled in a 100 ml Utermöhl chamber and the whole area of the bottom plate was counted. This is a common procedure for quantitative protist enumeration and the counting technique is considered as reliable especially if the whole bottom plate is counted. For the most abundant species between 200-600 cells of protists were counted per bottom plate which gives us ± 8-14% deviation (95% confidence interval) from our actual cell counts (following Lund et al. 1958). In fact, this is quite a good estimate for PZP cell numbers thus providing good quantitative counting data. The same quantitative counting procedure applies for phytoplankton. Therefore, we do not agree to the comment of Rev. 2 doubting the quality and reliability of our counting procedure. Despite the reviewers' assumption, PZB abundance cannot be considered as fairly low. At most sampling dates PZB biomass was around 20-40 µg C l^{-1} (especially during the first half of the experiment) which can be considered as moderate to high biomass for temperate and arctic waters. When encouraged by the BG editor to submit a new ms version we will provide an additional table showing cell count data for PZP as it was also recommended by Rev. 1. This will help to get a better idea on total PZP abundance and biomass in the mesocosms. Contrasting to the recommendation of Rev. 2, we did not take sub-samples for PZP enumeration per mesocosm since sub-samples from a single mesocosm are considered as pseudo-replication. Any means or standard deviations calculated from pseudo-replicates are considered as statistically incorrect. To take only one sample per mesocosm is actually the common practice when analysing plankton data from mesocosm experiments (see e.g. Sommer et al. 2005, Lewandowska & Sommer, 2010).

Statistics regarding quantitative implications are completely lacking.

We are not quite sure which kind of statistical analyses Rev. 2 had in mind. Since no replicate mesocosms were available, the only statistical tests we could run were regression analyses. We tested for significant correlations between the different pCO2 treatments and microzooplankton diversity, peak biomass magnitudes, total biomass and cumulative biomass. For all these analyses we did not find any positive nor negative relationship with pCO2. We apologize for not giving details on the regression analyses in the first version of the ms. But we will do so if encouraged to submit a revised version of the ms (additional table on the regression results and a clear description of the statistical test conducted).

Apart from this principal problem I have with the manuscript, there are many details to
be improved (in case the manuscript is accepted):
The introduction is biased towards negative impacts of OA on organisms, despite a
large body of literature suggesting at least the opposite (e.g., Hansen papers). The
two hypotheses read nice, but it remains unclear how the authors can find separate
answers to them. Stoichiometric changes are not addressed at all later (hypothesis 2).

We agree with Rev. 2 that we focussed more on negative effects of OA rather than taking positive effects into account. In a revised version we will include literature showing positive responses of OA in order to reflect balanced views in the intro. In accordance with Rev. 1+2 we will rewrite our objectives and reconsider our hypotheses in order to avoid any speculations on food quality aspects/stoichiometry since a data base for these aspects is not given.

The material and methods lack some considerable detail about sampling and analysis.

We are not sure which M&M details Rev. 2 was missing. What we will rewrite is the section on statistical analyses (as mentioned before) since this part was definitely not well written and lacking a lot of details. Most of the experimental details are given in the manuscripts by Riebesell et al. and

Schulz et al. (this BG special issue) which contain detailed description of the mesocosm design, the deployment logistics, methodology of CO_2-enrichment, abiotic factors and the maintenance of the mesocosms throughout the duration of the experiment.

The discussion lacks critical evaluation of the sampling scheme and lacking estimates of variability.
In a new ms version we will add 1-2 critical sentences on the sampling scheme and the mesocosm design.
In addition, the discussion on food web aspects (top-down, food quality) is very speculative as no data is presented that would support conclusions.
As mentioned above, we will omit speculations on food quality aspects in the introduction and the discussion section which cannot be proved by actual data.

In addition, the authors also need to discuss the new data in the context of interactions between food size, food abundance and food quality which is presently not the case and – as I feel it – beyond the focus of the manuscript.
The section on indirect effects of a high pCO2 (discussion section 4.2) will be shortened considerably in order to make the ms more concise and less speculative. When given the possibility to submit a revised version we are happy to provide more details on food sizes and abundances to stress the tight link between phytoplankton and microzooplankton biomass.
Besides this, important literature on trophic upgrading and sources of fatty acids in the food web are ignored.
In a new version we will thoroughly add some of the literature which is now missing.

Detailed comments:
Introduction:
P13033 line 1: This is not logic: Open ocean plankton communities are expected to be more vulnerable due to larger variations in pH in coastal areas, so why you argue then that 'therefore, one of the central questions: : :was whether arctic coastal plankton is vulnerable: : :'. What is the rationale behind the investigation when coastal plankton is likely to experience large variation in pH? In addition, Hinga's review (2002) emphasizes the importance of high pH in causing the large variability in coastal areas.
We agree that this section dealing with the different sensitivity of coastal vs. open ocean communities is not coherent. We will rewrite this section in order to make it more concise emphasizing e.g. on the role of a high pH on microzooplankton communities.
P13033 line 5: Protozooplankton is used synonymously with heterotrophic protists and protozoa, which is confusing, similarly to the reduction of protozooplankton to ciliates and heterotrophic dinoflagellates. What about mixotrophic protists, which are part of the PZP? Are they excluded from the analysis?
As also pointed out by Rev. 1, we agree that the phytoplankton/protozooplankton classification we made was not well described. In a new version we will clarify both terms by giving clear definitions. Most dinoflagellates we included into the protozooplankton (mainly Protoperidiunium spp., Katodinium spp., Gyrodinium spp.) since they are considered as predominantly heterotrophic. Some small chloroplast-bearing dinoflagellates were included into the phytoplankton since they were considered as predominately autotrophic. Putting dinoflagellates into special categories is not easy. But to date most dinoflagelles have been observed to be capable of ingesting prey items and purely autotrophic species are rare. Mixotrophic species were included in the analysis (e.g. Myrionecta rubra, Laboea strobila) and are listed in the species list (Table A1).

P13033 line 13: The description of potential effects on plankton is considerably biased towards negative effects. Apart from the cited literature, there is quite some literature available that clearly points out that many groups may not be affected. These should be cited and a more balanced description of potential implications for OA effects should be given.
As already mentioned before, we are grateful for the reviewers comment that our ms does not provide a balanced description on positive AND negative effects of OA. In a revised version we will include literature showing positive responses of OA in order to reflect balanced views and by considering the effects of OA on different components of the plankton.

P13033 line 20: The two hypotheses appear clear on the first look. However, one wonders whether and how these hypotheses can be separately addressed and analyzed in a complex mesocosm experiment. Furthermore, is there any evidence that changes in phytoplankton stoichiometry can alter phenology (in addition, no data on stoichiometry

is presented here)? The knowledge or evidence underlying the second hypothesis is
not adequately described in the introduction.

We will rewrite our objectives and reconsider our hypotheses. We agree that any speculations on food quality/stoichiometry should be avoided since no proof for food quality mediated changes in phytoplankton and a transfer to higher trophic levels can be given. The whole introduction section will be reworked accordingly.

In addition, large protists can be subject to considerable predation by larger zooplankton, the potential interaction of altered prey community and mesozooplankton grazing pressure needs to be addressed in such a complex set-up. Furthermore, changes in stoichiometry are not addressed later at all.

We tried to consider mesozooplankton grazing on microzooplankton using our mesocosms set-up by using mesozooplankton data from Niehoff et al. (this BG special issue). In the discussion section (page 13042, L.1-159) we stress that during the first phase of the experiment the mesozooplankton was dominated by herbivorous mesozoplankton. During phase 2+3 mesozooplankton grazing on microzooplankton played a role and e.g. Calanus copepodites were considered to suppress micrzooplankton biomass considerably.

P13034 line 1: Apparently, the experiment was started after the spring bloom. Although
this is not in the primary focus of the experiment, have the authors considered a potential
bias by the fact that the pelagic community already went through a spring cycle
associated with a pH shifting from low to high? A considerable decrease in pH due to
manipulation after the bloom is uncommon and might influence the protist composition.

We are aware of the problem that the spring community already went through a spring cycle. This might have affected the protist community and the initial phyto- and zooplankton compositions might not represent the same community as usually common in early spring. This might explain the relatively high initial dinoflagellate biomass at the beginning of the mesocosms experiment. However, the development of the microzooplankton community throughout our experiment reflected well a typical spring community development of microzooplankton as usually found in arctic or temperate waters.

Methods:

P13034 line 14: How where the mesocosms filled, by pumping water into them? This
is important to know (potential disruption of delicate organisms).

Details on the mesocosms set-up and the filling procedure are given by Riebesell et al. (this issue). We are aware that the mesocosms need to be filled gently in order to avoid any destruction of delicate organisms. The microzooplankton community sampled from the mesocosm reflected well the natural microzooplankton community and thus we are confident that the community inside the mesocosms represented a typical late spring community.

P13035 line 4: If I understood correctly, the authors want not only to study potential
changes in the community composition, but also in phenology. How can both be addressed
when samples were taken only once per week? Ciliate blooms can be highly
ephemeral (Moritz et al. 2006, Fileman & Leakey 2005,: : :), so part of the development
is missed.

We agree that sampling only once a week is a shortcoming of the present study. By using this sampling procedure we might have missed the actual biomass maxima of microzooplankton in each mesocosm and our present peak magnitude estimates might not be exact. However, we had to make compromises since microzooplankton was originally not included in the EPOCA project and no project money for microzooplankton enumeration was available. Microzooplankton counting is a time-consuming task and thus the number of samples analysed was restricted by manpower.

P13035 line 4-25: The sampling lacks many details. At which depths water samples
were taken? Volume per depth? Total sample volume? Since only heterotrophs are
addressed in the study, how were mixotrophs separated? Is this possible at all? Some
of the taxa included in Table 1 can be mixotrophic (Laboea strobolia for instance)! Any
real subsamples? The analysis lacks details on counting procedures and numbers of
counts per taxonomic group. Literature is missing in the references.

These issues were already answered above by giving details on microzooplankton categorization, counting procedures etc.

P13036 line 4: Again, details are missing: where Chla samples also taken from integrated
water samples?

Details are given in Schulz et al. (this issue)

P13036 line 10: The conducted test doesn’t appear appropriate for the purpose. Samples
were taken only once per week. Any small-term shift in the timing of development
of the protists (phenology) smaller than 6 days is not recognized; in addition, how
are maximum concentrations are defined (compare results)? The grazing pressure by

mesozooplankton might be different between mesocosms, potentially influencing composition, abundance and timing as well. Because of these interactions, one cannot simply compare samples taken at a time point x, but need to compare the dynamic development in the different mesocosms. The regression analysis needs to be specified.

In agreement with Rev. 1+2 we will include more details on the regression analysis. To account for the different dynamics in each mesocosm we compared the peak biomass as well as cumulative biomass. An additional option would be to compare cumulative results which is then independent of distinct dates, but focuses more on rates. This might provide a more detailed analysis and allow a better comparison between mesocosms. We will include such a cumulative analysis when submitting a new ms version to BG.

Results:

P13036 line 15: It would be helpful to see the pH/CO2 results as this is central to the understanding of any effects. From the material and methods, it is not clear whether the pH was manipulated only at the beginning of repeatedly. In Figure 1, PZP biomass bars should be confined to the day of sampling instead of covering several days; symbols don't fit with legend. Arrows could indicate timing of events (e.g., nutrient addition).

All these details are given by Riebesell et al. and Schulz et al. If wished by the Editor we can include these details into the present ms.

P13037 line 1: Is this trend significant? Is there an opposite trend in the second bloom?

We will provide details on this in a new ms version.

P13037 line 9: This is a bit contradictory: Dinoflagellates dominated the 3rd bloom, with higher levels at higher CO2. In contrast, the Chla was higher at lower CO2 in this phase. What makes the difference?

We don't fully understand this question. Why should the dinoflagellates and Chl. a respond in the same way to the different CO2 levels?

P13037 line 13: The counts for different Mesocosms reveal a 3 fold difference in the t0 values for PZP, directly after filling. This can influence the response to manipulation. In addition, I miss estimates of within mesocosm variation. This puts statements as 'decreased ' or 'increased' into doubt. Table 1 provides a detailed list: does it contain all taxonomic groups, or what about unidentifiable groups such as ciliates? The biomass is partly low, and one wonders on how many cell counts the biomass estimate is based? Which test has been used for correlation analysis?

These issues were already answered before (section 'General comments').

P13038 line 13: The description of 'response in biomass' and 'succession' is based on the 'analysis of maximal two consecutive samples without estimates of within mesocosm variability. Conclusions as 'strongest positive biomass response 'and 'peak biomasses' are doubtful with regard to the sampling frequency. How do the authors know when peak biomass was achieved?

See comments above.

Discussion:

I miss a general discussion on the limitation of the sampling scheme to identify responses of heterotrophic dinoflagellates and ciliates to OA. Quite important appears also that the experiment was apparently started following the spring bloom, a phase in which the pH of the system has likely been increased due to biological activity.

P13039 line 22: When the heterotrophic dinoflagellates and ciliates are undoubtedly sandwiched between primary producers and predators, should the dynamics of these organisms not be evaluated in response to both pressures? I assume abundant zooplankton was around, which consume dinoflagellates and ciliates.

The proper approach to address these issues would be a modelling approach. It is considered to perform such a modelling analysis containing all phytoplankton, micro- and mesozooplankton in a second step. At present, no modelling on the data was conducted but it will be a challenge for the near future.

P13040 line 3: Before judging 'that no direct effects on PZP composition and diversity were observed' authors should analyze if the set-up was suitable to answer this question with regard to sampling frequency and analysis. Still unclear to me is the actual values of pH during the experiment, have they been drifting or were pH levels kept constant? This of course has implications for the interpretation.

As already mentioned before all details on the sampling procedure, filling process and abiotic parameters are given by Riebesell et al. and Schulz et al.

P13041 line 1: The conclusion on lacking effects on carrying capacity and phenology are invalid considering the sampling frequency. How do the authors define maximum biomass and what are the appropriate temporal scales for analyzing phenology?

For mesozooplankton the appropriate scales would be 1 sampling per week (see Greve et al.). However, we are aware that microzooplankton reproduces more rapidly when compared to mesozooplankton and thus a higher sampling frequency would be optimal. But the number of samples counted was highly restricted by manpower and thus a higher enumeration frequency of microzooplankton was not feasible.

P13041 line 12: The statement of a strong decline contrasts with the biomass described for the first two sampling events in Fig 3, which were the dominant species. In addition, lacking estimates of the biomass variation within each sampling date make the conclusion doubtful. Many of the groups had a low biomass, and considering their large size, thus a low abundance, suggesting a potentially high variability is inherent in the methodology.

We commented on this issue above (section 'General comments'.

P13042 line 1: Instead of speculating about potential top down control, authors might consider to calculate the potential pressure from the abundance of zooplankton which is available. (Niehoff et al.)

This will be addressed in a second step when a modelling approach containing all phytoplankton, micro- and mesozooplankton will be conducted.

P13042 line 18: The original literature for trophic upgrading should be cited. For both top-down control and trophic upgrading no data is presented and the discussion here is very speculative. Why should food quality of autotrophs decrease as inferred here? Considering the composition of the protists, increasing zooplankton egg production might simply results from the changing size distribution. I find this very speculative here, as no data is presented.

As mentioned above the section on food quality aspects and trophic upgrading will be shortened considerably in order to make the ms more concise and less speculative.