1. The authors have revised their manuscript in accordance to the referee’s comments, but have not adequately responded to all identified problems. The long list of comments below indicates that there are still serious issues. I have especially problems with some of the methods and analyses which are below the standard of such a high ranking journal. In addition, much of the interpretation and conclusions is based on the inclusion of the results from the first days the MCs have been conducted. These are largely influenced by the sudden decrease of pH which is not naturally occurring. The authors have not responded to this issue. In general, analyses should be restricted towards the end of the mesocosms period when animals are acclimated. In my opinion, some of the conclusions are, therefore, faulty. Carry-over effects (size) are also not discussed. In summary, I still cannot recommend publication of the manuscript in Biogeosciences.

REPLY: We would like to take the opportunity to thank the reviewer for constructive comments and suggestions for major improvement of the manuscript quality. Further, we would like to apologize for having to some extent misunderstood part of the comments from the previous reviewer report, and therefore provided unsatisfactory replies. The fact that we did not understand all points raised by the reviewer caused some difficulties to reply in an appropriate manner. We hope we provide a more targeted response this time.

2. Introduction:
The authors have misunderstood the original comment by one referee. Here, it is not very important to know that there is a seasonal variation in pH. Here the referee questions whether it makes sense to conduct the experiments at a time point at which the pH is minimal. Thus, a community acclimated to a high pH (8.15)? is subjected to a sudden decrease of more than 0.6 units. This might cause unwanted treatment effects such as a comparison of an not acclimated population at reduced pH vs acclimated population at ambient conditions. As one of the referees has pointed out, the development of female size in the mesocosms point to such acclimation effects by showing a delayed size increase in the low pH mesocosm compared to ambient mesocosms. This brings cohort development out of tune during the first days and, thus, the developing populations in each of the mesocosms experience different combinations of pH, T and food. Thus, the rates measured cannot be interpreted on the background of pH alone. Under such conditions, maternal effects induced by pH can hardly be analysed, but require that a new cohort/generation has developed. That different pH can be experienced by vertical migration does not change much because intermittent exposure might contrast constant exposure and gives in addition animals a choice. (calculate how long this would take…..The authors have not responded to this important issue raised in the original review.???

REPLY: Thanks for the comment. Whether it made sense to conduct the experiments at the chosen time point is relevant, but was unfortunately due to reasons beyond our control. Thus, the time point was not chosen only for our copepod study, and therefore, we had to follow the schedule of the main campaign. The time of season that was selected for the mesocosm campaign was intended to cover the late spring and early summer period, because of the low productivity during this time, indicating the post-spring bloom period. Further reasons behind the campaign are outlined in detail in the overview paper by Paul et al. (2015). The low productivity also guaranteed fairly stable pCO₂ over time. Taken over the annual cycle, and considering that pH and CO₂ vary a great deal at the
site (due to upwelling as well as eutrophication), the post-bloom period was ‘the best available’ time period to conduct the mesocosm campaign.

Even though there are many advantages working with mesocosm setups, there are also some drawbacks (Riebesell et al., 2010). For example, the one mentioned by the reviewer can become a problem if initial conditions differ even slightly and communities develop differently in time and space due to perceived treatments, and subsequently occur in different phases. However, we do not consider this as a problem in the current study, as the phenomenon is assumed to be a greater issue for small-sized organisms having higher reproductive rates, such as flagellates and bacteria. In addition, initial conditions did not deviate to any larger extent in the present mesocosm study.

pH difference that adult Acartia spp. copepods can experience due to diel vertical migration in the study area can be more than 0.5 units (Almén et al., 2014). The difference is comparable with the pH difference between the control and high fCO$_2$ treatments in this study. In our opinion, this makes the issue of steep pH change at the beginning of the mesocosm study irrelevant. Also, the CO$_2$ additions were done stepwise to reduce the initial shock of the CO$_2$ addition.

We apologise if the reviewer feels that we have not responded to all the issues raised in the original review. The comments were so comprehensive that in some comments it was somewhat challenging to figure out which parts to respond to and what was requested.

3.
Methods:
Estimates of egg production is apparently based on the incubation of bulk 17 females only. This is quite below the recommended number (Laabir et al. 1995, Runge & Roff 2000) for this type of experiments. The volume-copepod ratio is also poor (200 mL are recommended). In general egg production can be variable and in the present work no variability estimates are done (recommended are 5-6 replicates (Runge & Roff 2000) (Page 5, line 26). The authors also argue that they have replication in time. This is not correct because size changes indicate the development of cohorts in the respective mesocosms, so principally animals with different history are compared. This is critical because cohorts developed differently in the various mesocosms (e.g. delayed cohort development in the low pH mesocosms) Thus, the cohorts have also a different history regarding other factors than pH. Therefore, replication in time is doubtful.

REPLY: The methods used for Acartia sp. egg production estimates ((17 females + 3 males) 1.2 L bottle$^{-1}$) have been carefully tested and were found to be appropriate in a number of studies (e.g., Kozlowsky-Suzuki et al. 2003 (20-25 females), Vehmaa et al. 2012 (20 females and 3 males)).

4.
During egg production, the pH increased in several bottles. This is unexpected considering the dim-light conditions and 17 animals added. How is this explained? In egg hatching experiments there was no outgassing, but partly large changes in the pH in both directions. Any explanations? (Page 5, line 30).

REPLY: Thanks for the comment! The units we employed were meant to be air-tight to minimise out-gassing. However, some out-gassing was likely occurring. On the other hand, there are a number of reasons that could have caused slight pH variability. For example, even though incubations were done in dim-light, phytoplankton probably photosynthesized slightly in the units. In addition, the low buffer capacity of the Baltic seawater could also have caused some variability
in the response of the seawater to treatments and handling. We assume that this is what may have happened in the units the reviewer is referring to.

5. Formalin shrinkage, for instance, depends also on temperature at sampling (Kapiris et al. 1997). Thus, one should use length measurements of fixed animals in comparisons only if these kind of effects are known. I assume little is known about similar effects of RNAlater. (page 6, line 7).

REPLY: Thanks for the comment! We picked the copepods to RNAlater in the temperature controlled room where temperature was 11-16°C [11°C (day 10), 15°C (day 17), 16°C (days 24 and 45)]. We are not aware of any similar studies on the effect of RNAlater. One replicate was handled at a time, and as soon as all individuals were in an Eppendorf tube, the sample was rushed to the ultra freezer. Hence even if the temperature at sampling had an effect on length measurements, this would have been minimal and would not affect the treatment difference and the effect would be the same for every mesocosm. It is possible that the size measurements between sampling days would be affected however, we were testing the effect of CO$_2$ and not time in this particular study.

6. The authors argue in their response that ‘contamination’ of eggs in the incubation water was low because females are deep and egg sink out. However, nauplii normally have very shallow distributions and might have biased the counts and according to the description in such cases counts were corrected (Page 6 line 26-32). The sentence in line 26 does not make sense to me; ‘calculated … only’ implies introduced females and eggs were included. So why only?

REPLY: Thanks for the comment! We assume that here might have occurred a slight misunderstanding; we will delete the word “only” and clarify the sentence (page 7, line 2–5). What we did was to count the number of produced eggs and adult females after the egg production incubation, and use them for calculating egg production rates. This enabled us to take into account also the possible extra females when calculating EPR. Consequently, having the egg numbers we subsequently assumed that there cannot be more nauplii at the end of the development incubation than eggs originally put into egg hatching units. Therefore, nauplii arriving from the sampled mesocosm water did not pose a problem for egg hatching and nauplii development results, and consequently for drawn conclusions. Hopefully, this clarified the confusion.

7. To exclude the nauplii >stage 4 is very subjective, especially since delayed hatching/development of nauplii occurs and thus such procedure affects treatments differently. Thus, there are many assumptions concerning numbers of eggs and nauplii inherent in the methods. The authors have not responded to this issue as well that the egg transplant study transfers eggs from the low pH into high pH for some low pH mesocosms but into basically similar pH in the high pH mesocosms. Similar to the initial setup, individuals (this time embryos) from the low pH treatment are subjected to large changes, but not the individuals from the high pH. Thus, maternal effects are tested in the low pH treatments only. (Page 6 line 32).
REPLY: Thanks for the comment! Please check the previous reply (above). More developed nauplii were excluded only if the total number of nauplii exceeded the total number of eggs prior to the development incubation. Hopefully, this issue is less confusing now.

The reviewer is right that this was not a reciprocal transplant experiment and we did not “return” the eggs that were produced in high f/CO₂ conditions into the control conditions. However, that does not mean that maternal effects would not play a role also in the control and lower f/CO₂ treatments, which is seen also in the Fig. 2. The “maternal conditions” cause a better offspring development even though they would be similar to the Baltic conditions. Earlier studies with the same species suggest that the larger the pH difference between the egg production and egg development conditions, the lower is egg hatching success (Vehmaa et al. 2012). Thus, a change in the environment between egg production and egg development might be bad for the egg. In this study, we were aiming to find limits for those differences, thresholds, above which the adaptive maternal effects cannot alleviate the negative effects of acidification on egg hatching and nauplii development. We found out that when egg are produced in high f/CO₂ conditions, they develop better in Baltic water than in their “maternal environment” even though that makes the pH difference between egg production and egg development conditions larger. We concluded that this was because the negative effects of f/CO₂ for egg development were beyond the limits of adaptive maternal effects.

8. Food characterized by TPC < 55 µm, C:N: It is correct that this size fraction includes food particles > 10 µm. However, dynamics of both size fraction do not correlate; e.g., small zooplankton <10 µm can cause a high TPC (as shown in Paul et al. 2015 for all mesocosms; roughly only 25% of TPC is > 10 µm), but do not contribute much to nutrition of zooplankton. The authors are referred to the papers of Dam et al. for the predictive power of such fractionation. Moreover, a filtration pressure of 200 mbar as used in the present study will likely destroy important food organisms in post bloom conditions such as ciliates (Taylor & Lean 1981, Broglio et al. 2003) and underestimate food concentrations. Thus, TPC is a poor predictor in the present case; in addition, size-selective feeding which is common under the described environmental conditions further complicates the straight forward interpretation conducted by the authors. When TPC < 10 µm is available, the authors could have used TPC < 55 µm minus TPC < 10 µm. This all applies to C:N as well (Page 7 line 23, and response to p. 188548, line 8).

REPLY: As already mentioned in the previous response to the referee comments, we are fully aware of the fraction chosen to indicate the quantity of food is not perfect, but it certainly is the best one available. The other options would have been chlorophyll a, TPC (<10 µm) or biomass calculations of different phytoplankton taxa. As mentioned also by the reviewer in the previous comments, Acartia is omnivorous, and Chl a or a selection of phytoplankton taxa would not describe its potential diet in a satisfactory manner. We have here used available data, measured during the mesocosm campaign. Unfortunately, TPC <10 µm was measured only from 4 mesocosms during the second half of the study. Therefore, we cannot use the difference of TPC <55 µm and TPC < 10 µm. We have widened the discussion about the possibility that used food quantity and quality did not fully describe the diet that Acartia sp. was consuming in the mesocosms (Page 14, lines 5–14).

9. Why 20 eggs? It is recommended to use at least 30 eggs (Laabir et al. 1996, Runge & Roff 2000). Thus, more than one egg sample does not meet the standards and day 24 needs to be excluded from the statistical analyses (Response to authors reply p 188546 line 17 in the original
REPLY: 20 eggs was not a target number of eggs; we apologize if our previous reply gave a false impression. We used all the eggs that were produced, and shared them (in two). The number of eggs used in the development incubations are listed in the Table S2. Please notice that Acartia has a low egg production in the Baltic Sea (much lower than in the oceans).

10. To exclude only day 3 from the analyses of length versus environmental condition is not sufficient. The first 10 days of the experiment T was < 11 degrees. Development time increases exponentially with decreasing temperature (e.g., the work on generation time by McLaren et al.) roughly doubling the G-time given by the authors in their response. These estimates are based on food replete conditions, thus a further increase in G-time is expected for the mesocosms. Thus, it is expected that copepods alone take already more than 17 days to mature during the initial half of the experiment (Page 8 line 15). Thus, length measurements of females beyond day 17 largely represent a mix of females with different history. Moreover predators are excluded and the life expectancy of ‘old’ females is likely large (they can be present for weeks), increasing their inclusion in size estimates far beyond the first 3 days. This cannot be ignored, especially since the females were small from the beginning in the low pH mesocosms.

REPLY: Thanks you for the comment! We are aware of that the length measurements represent and originate from a mix of females that matured in the mesocosms, and females that had reached the adult stage already at the start of the experiment. The assumption is that the variation in copepod female size is the same over all mesocosms at the beginning of the study, and that we would start to see treatment effects gradually as old females are replaced by younger ones, and that the proportion of the younger ones is increasing in our samples. We decided to include the data into the statistical model from Day 10 onwards, because then some younger adult females have had time to grow and reach maturity.

We have added error bars to the Fig. 1b. It shows that the size differences were not as large at the beginning of the study as the average size would possibly suggest.

11. Results:
Reproduction: While it is acceptable that the authors do not present the weight specific egg production, the effect of size on body weight needs to be discussed because this can have an effect on EPR. The error bars of length measurements are missing (n=17) (Page 9, line 13; response to 188550, line 6).

REPLY: We have added error bars to the Fig. 1b. EPR was not significantly related to environment variables, so we do not see the reason why we should discuss this issue. Or what does the reviewer mean with the effects of size on body weight?

12. Length analysis: When I understand correctly, these analyses (Table 2) are based on the whole experimental period including day 10 to 45. As outlined above, the females from day 10 to 20 have a history largely outside the mesocosms and should be excluded. Table 1 and 2 are confusing, as Table 1 says that only TPC was used in LMM/GLMM for length (Page 9, line 18;
REPLY: Thanks for this comment! We are afraid that a misunderstanding from the reviewer’s part has occurred here. The days that we included in the LMM/GLMMs are marked with a cross in the table (please see Table 1). TPC (Table 1) was used in all analyses performed and not just for length. Table 2 is a summary of the final (simplified) models. Please see our response above, which explains why we included measurements from Day 10 onwards to the prosome length analysis.

If only data from days 24 and 45 is used to analyze fCO$_2$ effects, as suggested by the reviewer, the negative relationship between prosome length and average fCO$_2$ since start of the experiment is almost statistically significant (linear regression: $R^2 = 0.3167$, $F_{1,10} = 4.635$, $p = 0.057$). At the end of the study the smallest Acartia sp. females were found from the mesocosms with highest fCO$_2$ treatments.

![Graph showing relationship between prosome length and fCO$_2$](image)

13.
Maternal effects: Apparently, all measurements of egg hatching success are included in the analysis (this needs to be described in the MatMeth). The original results (hatching in Baltic Water) are not described. As it looks like much of the relationship to pCO$_2$ is driven by first day results in which stressed animals (MC 3 and 8 just adjusted to low pH) are compared to acclimated animals basically kept in the similar environment of a pH around 8. The conclusion of a threshold is erroneously based on these days. Diatoms (Skeletonema) was present during the initial phase of mesocosms and could contribute to low hatching in combination with pH as well. Maternal effects (animals with at least some history in the respective treatments) could be tested from day 16 onwards, but this would also require estimates of within treatment variability for each measurement (nevertheless, no effect was found for these days!). The assumption that all animals share the same history in GLMM is basically wrong. The females tested are a mix between old and newly maturing, and the population is constantly changing, but potentially not at the same pace (e.g., potential effects of delayed development) (Page 10, line 12).
REPLY: Thanks for the comment. Please see Table 1 that presents the days which are included in each model. We have added a figure showing the egg hatching results in Baltic water (Fig. 1e). Concerning the discussion about the threshold in the data between MC 3 and MC 8 versus the other mesocosms, it forms part of the conclusions that we draw from the data. Partly this speculation as one may call it is based on data available in figures, and it could also be interesting for the reader to learn what we concluded from the experiment. Furthermore, we state that there may be a threshold; we never stated that there is a threshold. It is indeed a fascinating thought that *Skeletonema* would have affected the hatching results. The species was present in all mesocosms on Day 3 (same abundances in control and high fCO$_2$ treatments) but its abundance decreased basically to zero already on Day 6. Thus, the lower egg hatching on Day 3 could have partly been caused by this diatom species. Paul et al. (2015) found no relationship between the percentage contribution of diatoms to chlorophyll $a$ and fCO$_2$ in the first phase of the experiment (Days 1–16). Therefore, the combined effects of *Skeletonema* and pH are not a very likely explanation for the differences in egg hatching success between the mesocosms at the beginning of the experiment. We have extended the food quality discussion (page 14, lines 5–14).

72-hour acclimatisation time have been used for CO$_2$-treatments even higher than the high treatments in this study (Cripps et al., 2014a; 2014b). Also, based on our previous experience with *Acartia* sp., three days is enough for this species to acclimatise to changed CO$_2$-conditions (Vehmaa et al., 2012).

We understand the reviewers point about varying backgrounds of the tested females but we do not agree with this view. If all copepods that are ever used to measure *in situ* egg production / egg hatching rates (e.g., on cruises) would need to be “in the same phase”, we could never do such studies, not in the middle of the season at least when there are both old and young females in the water. That is why we use several female copepods and average out individual variation. The individual variation is likely smaller after a long acclimatization period in the treatment conditions, i.e. after all the used animals have been born and developing in similar conditions. However, in long experiments there are also other factors that start to affect the results, such as wall effects and gradual deviation of the enclosed communities from the natural system (e.g. the lack of higher trophic levels), which create the limits for duration of the study (Riebesell et al., 2010). Thus there are different phases in the experiment that, at least to us, are all interesting. If one would only analyse and present one time point at the end of the study, all this valuable information would be lost. Hopefully, the reviewer can see the point we are addressing here.

14. This conclusion is not logic: A negative effect of pCO$_2$ on the ration mesocosms/Baltic would imply that hatching in maternal environment is smaller than in the Baltic with increasing CO$_2$, which contrasts with the statement given here (Page 10, line 18).

REPLY: Thanks for the comment. We noticed that the sentence is unclear, and it is not totally obvious what we mean. We have now rewritten and clarified it, hopefully, the sentence is now less confusing for the reader (page 11, lines 2–5).

15. Discussion:
The authors have not responded to the referee’s suggestion to adequately include a proper analysis of history. In my opinion the threshold conclusion on egg hatching is based on the first days
results, which basically are inappropriate to analyse ‘adaptive maternal’ effects as the history of females is largely outside the mesocosms. The vanishing of the effects with time argues against this conclusion (page 11, line 2).

REPLY: Thanks for this comment! We have replied above in detail to this question and done our best to clarify our intentions with the conclusion. In addition, we have toned down the conclusion and speculations about the threshold (page 2, lines 4–8; page 11, lines 17–20).

16. When food is limiting, interactive effects of food and temperature exists for length and egg production. Whether this is interesting for the authors or not, it should be included in analyses as T was not constant (page 11, line 2).
Where does the negative correlation of TPC originates from? The authors could include some ideas on the effects of changing phytoplankton communities.
The discussion of transgenerational effects is redundant because the history of the females in the mesocosms is for large periods unknown. Only in the second half of the experiment there is a change that a new generation contributes to the population (this coincides with the disappearance of negative pCO2 effects). However, the population is likely a mix of several generations. (Page 12, line 21)

REPLY: Again, we agree that analysing temperature-food interactions would be highly interesting; however, impossible to do in order to avoid over-parameterization of the model. Nevertheless, hopefully we are able to test these effects in a potential future study.
Paul et al. (2015) observed no effect of CO2 in larger organisms (2–20 µm), which were dominant in the phytoplankton community during the period of higher productivity (Phase I, Days 1–16). Instead, total TPC (Phase III, Days 31–43) and TPC <10µm (Phase II, Days 17–30) were positively correlated with fCO2. These effects became apparent when the phytoplankton community size structure shifted towards dominance of smaller phytoplankton size classes, such as picophytoplankton (<2 µm). It is thus likely that the observed negative effect of TPC on copepod size was because the dominant phytoplankton size class was unavailable for copepods as food. We have added this to the Discussion (page 14, lines 11–14).
On average the history of the females (as well as initial generational composition) across all mesocosms should be the same, hence we were testing the response to CO2 treatment. We tested whether copepods can adjust their offspring to prevailing conditions or not, and with the prevailing conditions we meant the conditions during egg laying. As the environmental conditions can change during the lifetime of a copepod also in the wild, for example due to an upwelling event, we do not see why we could not test transgenerational effects under different CO2 treatments in the mesocosms.

17. The authors have not responded to the referees objections that the smaller size in the lowest pH MC was carried over from the field, and that any delayed development due to the initial large decrease in pH (which is not experienced as such by populations!) changed the conditions under
which new females developed to more unfavourable conditions (low food, higher T; again interactive effects) which have a strong influence on size. Under such conditions mesocosms cannot be compared for pH effects alone (Page 14, line 1).

REPLY: We are sorry if our reply to the question was not satisfying. We have added the error bars to the Fig. 1b. It shows that the size differences were not as large at the beginning of the study as the average size would possibly suggest. We have widened the Discussion about the possible interaction effects (page 15, lines 8–12). We agree that analysing interactions would be highly interesting; however, impossible in the current case in order to avoid over-parametrization of the model. Hopefully, we would be able to test these effects in a future study.

18. The estimates of G-time are erroneous. Dzierzbicka et al. (2009) used an average of very different species (A. clausi, A. longiremis) for their estimates of Acartia spp. in a modelling study (not bifilosa), of which one is a polar copepod (A. longiremis). Underlying these estimates are food replete conditions which do not apply to the mesocosms. The initial T conditions were lower than 17 degrees. The estimates of G-time are thus largely overestimates. Error bars (n=17) should be presented and females are largely a mix of generations rather than separate generations.

REPLY: It is a pity that the reference we used did not appeal to the reviewer. We recognise the limitations of using generational times for a different species under different environmental conditions, however in science one does seldom have perfect data or perfect references to cite. Unfortunately these kind of studies have not yet been done on A. bifilosa. but we considered this as the best available study with the most relevant generation times for Baltic Sea Acartia spp. to relate to A. bifilosa in this study. We have clarified this in the Discussion (page 14, lines 29–31). We have also added error bars to the Fig. 1b.

REFERENCES


Ocean acidification challenges copepod phenotypic plasticity

A. Vehmaa¹,², A.-K. Almén²,³, A. Brutemark¹,²*, A. Paul⁴, U. Riebesell⁴, S. Furuhagen⁴, J. Engström-Öst²

[1]{Tvärminne Zoological Station, University of Helsinki, J.A. Palménin tie 260, FI-10900 Hanko, Finland}
[2]{Novia University of Applied Sciences, Coastal Zone Research Team, Raseborgsvägen 9, FI-10600 Ekenäs, Finland}
[3]{Environmental and Marine Biology, Faculty of Science and Engineering, Åbo Akademi University, Tykistökatu 6, FI-20500 Turku, Finland}
[4]{GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany}
[5]{Department of Environmental Science and Analytical Chemistry, Stockholm University, Svante Arrhenius väg 8, SE-11418, Stockholm, Sweden}

[*]{present address: Calluna AB, Torsgatan 30, SE-11321 Stockholm, Sweden}
[+]{present address: Swedish Chemical Agency, Esplanaden 3A, SE-17267 Sundbyberg, Sweden}

Correspondence to: A. Vehmaa (avehmaa.vehmaa@gmail.helsinki.com)

Abstract

Ocean acidification is challenging phenotypic plasticity of individuals and populations. Calanoid copepods (zooplankton) are shown to be fairly plastic against altered pH conditions, and laboratory studies indicate that transgenerational effects are one mechanism behind this plasticity. We studied phenotypic plasticity of the copepod *Acartia* sp. in the course of a pelagic, large-volume mesocosm study that was conducted to investigate ecosystem and biogeochemical responses to ocean acidification. We measured copepod egg production rate,
egg hatching success, adult female size and adult female antioxidant capacity (ORAC) as a function of acidification ($f$CO$_2$~365−1231 µatm), and as a function of quantity and quality of their diet. We used an egg transplant experiment to reveal if transgenerational effects can alleviate the possible negative effects of ocean acidification on offspring development. We found significant negative effects of ocean acidification on adult female size. In addition, we found signs of a possible threshold of CO$_2$ concentration (~1000 µatm), above which adaptive maternal effects cannot alleviate the negative effects of acidification on egg hatching and nauplii development. We did not find support for the hypothesis that insufficient food quantity (total particulate carbon < 55 μm) or quality (C:N) weakens the transgenerational effects. However, females with high ORAC produced eggs with high hatching success. Overall, these results indicate that *Acartia* sp. could be affected by projected near future CO$_2$ levels.

**Keywords:** *Acartia bifilosa*, climate change, maternal effects, total particulate carbon, C:N, oxidative stress

1 Introduction

Increased concentrations of carbon dioxide (CO$_2$) in the atmosphere is changing the carbon chemistry of the world’s oceans. CO$_2$ dissolves in seawater thereby decreasing ocean pH. Ocean acidification is increasing fast and pH is expected to decrease by a further 0.14−0.43 pH units during the coming century (IPCC, 2007). Acidification can cause various problems to biochemical and physiological processes in aquatic organisms. In addition to affecting calcification of calcareous organisms, maintenance of acid-base equilibrium of body fluids may become more difficult and have consequences for example on protein synthesis, metabolism and volume control (Whiteley, 2011).

In a changing environment, populations can respond in three main ways: through plastic responses of individuals, through genetic changes across generations, or through escaping in space or in time by phenology modifications. Under a rapid change, phenotypic plasticity, i.e., the ability of an individual or a population to alter its physiological state, appearance or behaviour in response to the environment is of major importance (West-Eberhard, 2003).

Theory predicts that higher plasticity evolves in extreme environments, and that spatial heterogeneity and dispersal select for higher plasticity (Chevin et al., 2013). One could therefore hypothesise that organisms inhabiting a variable environment, such as the study area, could be
fairly plastic in their response to ocean acidification because they have to cope with both seasonal and sudden changes in pH (Almén et al., 2014; Lewis et al., 2013).

Proteomic studies suggest that oxidative stress is a common co-stress of temperature and acidification (Tomanek, 2014). Increased production of reactive oxygen species (ROS) may result in increased antioxidant and/or repair costs, and further in reduced investment in reproduction or other functions, such as immune defence. In addition, increased production of ROS may lead to accumulation of oxidative damage and further to acceleration of senescence (Monaghan et al., 2009). There can also be a connection between maternal oxidative balance and offspring quality. In birds, for example, females allocate diverse antioxidants to the eggs that protect the embryo from oxidative stress. This maternal effect has a positive effect on offspring development and growth (Rubolini et al., 2006).

Copepods (zooplankton) are indispensable to the functioning of the whole pelagic ecosystem and contribute significantly to many ecosystem services (Bron et al., 2011). For example, they provide food for early-life stages as well as some adult fishes of many economically important fish species (Steele, 1974; Cushing, 1990).

Previous results suggest that calanoid copepods have high buffering capacity against projected ocean acidification for the year 2100 and beyond (Kurihara and Ishimatsu, 2008; Weydmann et al., 2012; McConville et al., 2013; Vehmaa et al., 2013), meaning that they are able to survive, grow, develop and reproduce in lower pH (Reusch, 2014). However, there are also studies showing negative impacts on moderate CO₂ levels (Fitzer et al., 2012), whereas most of the negative impacts have been discovered for extreme, carbon storage scenarios (Kurihara et al., 2004; Mayor et al., 2007; Weydmann et al., 2012). Many studies have tested only one life-stage, adult females, and have therefore possibly underestimated the effects of ocean acidification on copepods (Cripps et al., 2014a). There are indications that transgenerational effects are one mechanism responsible for the high plasticity of copepod reproduction against altered pH conditions (Vehmaa et al., 2012). This maternal effect is most likely dependent on the condition of the mother and the availability of food and quality of her diet (Vehmaa et al., 2012; Pedersen et al., 2014a). Paternal effects can also influence offspring traits. Exposure of both parents to CO₂ leads to fewer adverse effects on egg production and hatching than exposure of only gravid copepod females (Cripps et al., 2014b). Thor and Dupont (2015) also highlight the importance of testing transgenerational effects. They found significantly lower copepod egg production after two generations when exposed to 900 and 1500 μatm compared
to 400 µatm, but transgenerational effects alleviated the negative CO$_2$ response in 1500 µatm (Thor and Dupont, 2015).

We tested direct and indirect effects of ocean acidification (i.e., via food quantity and quality) on the copepod Acartia sp. egg production (EPR), egg hatching success (EH), female body size (measured as prosome length (PL)), as well as antioxidant capacity (ORAC). The study was conducted in association with the KOSMOS (Kiel Off-Shore Mesocosms for Ocean Simulations) project in the Baltic Sea (Paul et al., 2015). The study was intended to cover the low productivity late spring and early summer period, i.e., the post-spring bloom period when $p$CO$_2$ concentrations are at the annual minimum. Over the annual cycle, $p$CO$_2$ and pH vary substantially at the study site as a result of biological activity and mixing/upwelling of CO$_2$-enriched deep water (Niemi, 1975; Omstedt et al., 2014). There are also strong spatial gradients in seawater $p$CO$_2$/pH, most prominently between the surface layer and the CO$_2$-rich deeper waters (Almén et al., 2014). Thus, the copepods in the study area are likely to experience strong changes in seawater carbonate chemistry, both seasonally and during their diurnal migration.

Total particulate carbon (TPC <55 µm) was used as the measure of food quantity. Food quality was indicated by carbon to nitrogen ratio of the same size fraction of seston (C:N <55 µm) (Elser and Hasset, 1994; Sterner and Hessen, 1994). In addition, in order to separate transgenerational plasticity (i.e., maternal and paternal effects) and the effect of environment on copepod egg hatching and development, we performed an egg-transplant experiment. Half of the produced eggs were allowed to develop in respective mesocosm water and the other half in water collected outside the mesocosm bags.

Due to the high buffering capacity of Acartia sp., we hypothesised that there are no $f$CO$_2$ related differences in egg production rate, egg hatching success and prosome length between the mesocosms. In addition, we hypothesised that copepod eggs hatch and develop better in the same environment in which they are produced, because transgenerational effects can alleviate the negative effects of environmental change. Our third hypothesis stated that low food quantity (TPC) and poor quality (high C:N) will weaken the maternal effect by deteriorating the condition of the mother. Finally, we tested if mothers with higher antioxidant capacity (ORAC) produce better quality offspring (EH) by calculating correlation coefficients between the two variables.
2 Materials and Methods

The study was performed in summer 2012 in the vicinity of Tvärminne Zoological Station on the south-western coast of Finland. Six large mesocosms were moored on site in the beginning of June. To enclose the natural plankton community, the mesocosms were left open with only 3 mm mesh size net covering the top and the bottom during filling. After four days, the net was removed and the top was pulled up 1.5 m above the water surface and closed at the bottom (Riebesell et al., 2013; Paul et al., 2015). pH was ~8 and fCO₂ concentrations in the mesocosms prior to adjustment were 237±9 µatm (average±std of daily measurements from all bags). Four mesocosms were manipulated with CO₂ enriched seawater, during three consecutive days to reach fCO₂ concentrations of 600-1650 µatm (Paul et al., 2015). Two untreated mesocosms were used as controls. The water column was mixed in the beginning of the experiment to avoid salinity stratification. Due to outgassing, CO₂ was also added on day 15 to the upper 7 m of the high CO₂ mesocosms to maintain the treatment levels. No nutrients were added.

2.1 Sampling

Sampling took place once a week during the first four weeks of the experiment, and once more at the end of the whole experiment (days 3, 10, 17, 24 and 45). Mesozooplankton were sampled by taking two hauls with a 300 µm net (17 cm diameter) from 17 m depth and from all mesocosms. The samples were rinsed into containers with 4 l of seawater from respective mesocosm taken from 9 m depth with a water sampler (Limnos, Hydrobios). On the same day, integrated water samples (0-17 m) were collected from all mesocosms and the Baltic Sea directly into 1.2 l Duran bottles that were closed without head space. Water samples were kept in cool bags and zooplankton samples were protected from light until transported to a temperature and light controlled room at Tvärminne Zoological Station within 4 h. The light:dark cycle in the room was 16:8 h and light intensity was 7 µmol photons m⁻² s⁻¹ (LI-COR LI-1000). Temperature followed the in situ temperature [9°C (day 3), 11°C (day 10), 15°C (day 17), 16°C (days 24 and 45)].
2.2 Measurements of egg production, egg hatching success and prosome length

Twenty adult *Acartia* sp. (17 females and 3 males) were picked with pipettes from each sample using stereo microscopes, and gently placed in pre-filled glass bottles with respective mesocosm water. The bottles were closed without head-space, to prevent CO$_2$-outgassing during the incubation. pH was measured from the bottles before closing and right after opening them at the end of the incubation using Ecosense pH10 pH/temperature pen (Table S1). The pen was calibrated with standard buffer solutions (Certipur, Titripac pH 4.00, 7.00, and 10.00) every second day. The bottles were incubated in temperature and light controlled room in conditions described above (Materials and Methods 2.1), and mixed three times a day and their place on the shelf was changed randomly. After the incubation (24.3 ± 2.3 h, average ± std), the copepods and produced eggs were filtered using 250 µm and 30 µm sieves, respectively. The copepods were counted and their viability checked before preserving them in RNA*later* (Sigma). RNA*later* can affect size (Foley et al., 2010), and the effect depends on the number of segments in the animal, i.e., the more segments the larger effect. Shrinkage is ~15% for copepods (Prof. Elena Gorokhova, Stockholm University, pers. comm.). Prosome length of the preserved female copepods was measured using a stereo microscope (Leica MZ12) and ocular micrometer (total magnification 100 ×). As all the measured copepods were adult females, we assume the shrinkage to be in proportion similar for all individuals, which means that our results are quite conservative and comparable between mesocosms.

In the egg transplant experiment, the collected eggs were divided for hatching into two 50 ml petri dishes with different conditions; one dish was filled with respective mesocosm water and the other filled with Baltic water. pH of the water was measured as above before the incubations and right after the petri dishes were opened after the incubation (Table S1). The eggs were counted before the petri dishes were completely filled and sealed without head-space using Parafilm. Egg hatching was followed by counting the number of remaining eggs on the dish through the lid using a stereomicroscope twice a day. When the number of eggs had remained the same on two consecutive counting times, the dishes were opened and the water containing the remaining eggs and hatched nauplii was preserved with acid Lugol’s solution. Therefore the hatching incubation time varied between 63.9 and 137.6 h, depending on incubation temperature. *Acartia* sp. nauplii stages were determined and the number of nauplii and remaining copepod eggs counted using a stereo microscope.
SAs some adults, copepodites, nauplii or eggs could have ended up in the incubation bottles or petri dishes with the unfiltered incubation water. The possible extra adults and their contribution to the egg production rate (EPR, eggs copepod$^{-1}$ d$^{-1}$) were taken into account as EPR was calculated using only the number of eggs and adult *Acartia* sp. females found in the incubation bottles after the 24 h incubation. When estimating the egg hatching success (EH, %), the total number of hatched *Acartia* sp. nauplii and remaining eggs at the end of the hatching incubation was compared with the number of eggs counted before the hatching incubation. If the total number exceeded the egg number prior to hatching, the most developed nauplii (>N4) were considered to be carry-over individuals, and were therefore not considered in the estimation of EH. For estimation of nauplii development, rate the development index (DI) was calculated (Knuckey et al., 2005) accordingly,

$$DI = \frac{\sum_{i=0}^{N_4} N_i \times n_i}{\sum_{i=0}^{N_4} n_i}$$

where $N_i$ is the assigned stage value (0 for eggs, 1 for N1, 2 for N2 and 3 for N3 and N4) and $n_i$ the number of individuals at that stage. We assume all the *Acartia* sp. adults and nauplii to be species *A. bifilosa*. However, because another *Acartia* species, *A. tonsa* occurs in the area in late summer too (Katajisto et al., 1998), we cannot be totally sure that we only had one species in the experiments.

### 2.3 Antioxidant capacity

For antioxidant capacity (ORAC) samples ~25 live female *Acartia* sp. were picked from every zooplankton sample onto a piece of plankton net in the temperature and light controlled room on days 3, 10, 17 and 31. The net containing the copepods was folded and stored in Eppendorf tubes at -80°C. The samples were homogenised in 150 µl Tris-EDTA buffer containing 1% sarcosyl. The antioxidative capacity was assayed as ORAC (Ou et al., 2001). As a source of peroxyl radicals, 2, 2-azobis (2-amidinopropane) dihydrochloride (AAPH) (152.66 mM) was used and fluorescein was used as a fluorescent probe (106 nM). We used trolox (218 µM, Sigma-Aldrich) as a standard and the assay was performed on a 96-well microplate and to each well, 20 µL sample, 30 µL AAPH and 150 µL fluorescein were added. ORAC values were normalized to protein and expressed as mg Trolox eq. mg protein$^{-1}$. Protein concentration was measured with NanoOrange® (Life Technologies).
2.4 C:N and TPC

Samples for TPC and C:N were collected onto GF/F filters (Whatman, nominal pore size 0.7 µm) using gentle vacuum filtration (<200 mbar) and then stored in glass petri dishes at -20°C. GF/F filters and petri dishes were combusted at 450°C for 6 hours before use. Gauze pre-filters were used to separate the size fraction < 55 µm. Filters were not acidified to remove inorganic carbon, therefore total particulate carbon is used. C and N concentrations were determined on an elemental analyser (EuroEA) following Sharp (1974), coupled by a Conflo II to a Finnigan DeltaPlus mass spectrometer and were used to calculate C:N ratios in mol:mol. For further details on sampling and analyses, please refer to Paul et al. (2015).

2.5 Statistics

The effect of acidification and food quantity and quality on Acartia sp. egg production (EPR), prosome length (PL), antioxidant capacity (ORAC) and nauplii development index (DI) was tested using linear mixed effect models (LMM) with restricted likelihood (REML) approximation from the nlme-package (Pinheiro et al., 2014), where EPR, PL or ORAC were used as response variables, fCO2, TPC (<55 µm) and C:N as fixed explanatory variables and repeated measure of the mesocosms over time as a random factor (Table 1). Due to the binomial nature of the data, the effect of fCO2, TPC (<55 µm) and C:N on egg hatching success (EH) was tested with generalized linear mixed model (GLMM) with Laplace likelihood approximation, binomial error structure and logit-link function from the lme4-package (Bates et al., 2014) (Table 1). The average of fCO2, TPC (<55 µm) and C:N measurements from each mesocosm within three days before the zooplankton sampling were used as explanatory variables for EPR, ORAC and EH, because 2–3 days are considered to be an appropriate acclimatisation period for A. bifilosa (Yoon et al., 1998; Koski and Kuosa, 1999). For PL, the average of all fCO2, TPC (<55 µm) and C:N measurements from the start of the mesocosm experiment were used since PL reflects the environmental conditions of the whole lifespan of the animal. In addition, Day 3 was excluded in the LMM testing the PL (Table 1), since three days is too short period for detecting differences in copepod size. Egg–adult generation time for A. bifilosa at 17°C is approximately 16 days of which ~7.5 d taken by nauplii stages and ~8.5 d by copepodite stages (Yoon et al., 1998). Collinearity between all explanatory variables was checked. Temperature was not considered in the models, because it changed similarly in all the bags (Paul et al., 2015). The model simplifications were done manually in backward stepwise manner by removing the non-significant effects and by using Akaike’s information.
criterion (AIC). We report t- or z-statistics (EH) of the retained fixed effects. To separate the
effect of hatching environment from maternal environment, EH and DI were divided with the
 corresponding values measured in the Baltic Sea water. The ratio of Mesocosm EH (or DI) /
Baltic EH (or DI) >1 indicates that eggs hatch or develop better in the maternal conditions
(Mesocosm water), whereas the ratio <1 indicates that eggs hatch or develop better in the Baltic
Sea water. The effect of maternal environment (\(f\)CO\(_2\), TPC (<55 µm) and C:N) on the ratio was
tested with LMM, where the ratio of Mesocosm EH / Baltic EH and Mesocosm DI / Baltic DI
were used as response variables; \(f\)CO\(_2\), TPC (<55 µm) and C:N as fixed explanatory variables;
and repeated measure of the mesocosms over time as a random factor. The model
simplifications were made as above.

To test if maternal antioxidant capacity (ORAC) correlates with egg hatching success,
Spearman rank correlation tests were used. Data from Days 3, 10 and 17 were included in the
test (n = 17, EH result for MC 6 in Day 3 is missing) because those are the days when both
ORAC and EH were measured.

All the statistical analyses were performed using software R 3.0.2 (R Core Team, 2013), and
the significance level was 0.05.

3 Results

3.1 Egg production, prosome length, antioxidant capacity and egg hatching
success

Acartia sp. egg production (EPR) increased in all mesocosms between Day 3 and Day 10, but
decreased after that, reaching very low rates (1-2 eggs copepod\(^{-1}\) d\(^{-1}\)) on Days 24 and 45 (Fig.
1a). Neither food quantity (TPC, <55 µm), food quality (C:N, <55 µm), nor ocean acidification
\(f\)CO\(_2\) had a statistically significant effect on copepod egg production (Table 2), even though
there seemed to be variations in those parameters between the mesocosms (Table 3).

Prosome length (PL) of Acartia sp. females increased during the first week of the study;
however there seemed to be some differences between the mesocosms already on Day 3, which
was not included in the analysis (Fig. 1b). From Day 10 onwards, the smallest A. bifilosa adults
were found in the mesocosm with the highest \(f\)CO\(_2\) concentration (Fig. 1b). \(f\)CO\(_2\), but also TPC
(<55 µm) had a statistically significant negative impact on copepod body size (Table 2).
Antioxidant capacity (ORAC) of the female copepods increased from Day 3 to Day 10 in all mesocosms (Fig. 1c). Interestingly, on Day 3 ORAC was highest in the three mesocosms with highest \( f_{CO_2} \) treatment, whereas on Day 31 the situation was opposite and ORAC was lowest in the three mesocosms with highest \( f_{CO_2} \) (Fig. 1c). Despite this, only TPC (<55 µm) had a statistically significant effect on ORAC; ORAC decreases with increasing TPC (Table 2).

The overall egg hatching success (EH) was high throughout the study; over 80% of the Acartia sp. eggs hatched. As seen for EPR, PL, and ORAC, EH also increased from Day 3 to Day 10 in all mesocosms (Fig. 1d). Variance in the EH between the four samplings was highest in the mesocosms with highest \( f_{CO_2} \), whereas EH varied the least and remained >90% in both control mesocosms (MC1, MC5). In spite of this, only TPC (<55 µm) had a statistically significant negative effect on EH (Table 4). Eggs that were produced in MCs 3, 5, 6 and 7 had fairly similar hatching success in Baltic water, whereas hatching success of eggs that were produced in MCs 1 (control) and 8 (the highest \( f_{CO_2} \)) was alternately either lower or higher than in the other MCs (Fig. 1e).
However, on Days 17 and 24 the fCO$_2$ treatment did not have a clear effect on hatching success. Nevertheless, fCO$_2$ had a statistically significant negative effect on the ratio of EH mesocosm / Baltic, meaning that egg hatching was higher in the Baltic water than in the maternal environment when the maternal environment had a high fCO$_2$ (Table 5). When maternal environment had low fCO$_2$ the situation was vice versa. The level of fCO$_2$ had also a significant negative effect on the DI mesocosm / Baltic ratio (Fig. 2b; Table 5).

3.3 Correlations between antioxidant capacity and offspring quality

Copepod antioxidant capacity (ORAC) was correlated significantly with copepod egg hatching success. The relationship between the two variables is positive and stronger for eggs developing in the mesocosm water ($\rho = 0.75, p < 0.001$) than for eggs developing in the Baltic water ($\rho = 0.62, p = 0.007$) (Fig. 3).

4 Discussion

In this study, conducted in semi-natural mesocosm environments, reproduction of the Acartia sp. copepod showed high phenotypic buffering against acidification, i.e., the species was able to maintain similar egg production rate and also high egg hatching success in all fCO$_2$ conditions. Nevertheless, we found significant negative effect of ocean acidification on adult female size. Even more interestingly, we found signs of a possible threshold of fCO$_2$ concentration (~1000 µatm) for offspring development, above which adaptive maternal effects cannot alleviate the negative effects of acidification on egg hatching and nauplii development (Fig. 2). However, we did not find support for the third hypothesis that poor food quality (higher C:N) would weaken the maternal effect by deteriorating the condition of the mother. Conversely, higher food quantity (TPC <55 µm) correlated negatively with egg hatching success, adult female size and antioxidant capacity, whereas C:N ratio did not correlate with any of the measured variables significantly. Copepods were possibly food limited in all the mesocosms, especially after Day 17 due to a sharp decline in Chl a concentrations and in phytoplankton community size structure (Paul et al., 2015). Dominance of picophytoplankton that is too small to be consumed by copepods could be the reason for the observed negative effects of food quantity, and that may have also masked the food quality effect. Also, after Day 17 egg production rate was so low that it was practically impossible to find differences in egg production between the mesocosms. Finally, we found a positive correlation between maternal antioxidant capacity and egg hatching success,
suggested that the female antioxidant defence might also protect the embryo from oxidative stress.

The fact that *Acartia* sp. egg production and egg hatching were unaffected by high fCO$_2$ but egg transplant experiment revealed that development was slower for nauplii at high CO$_2$ supports the importance of looking beyond egg production and egg hatching, which is also pointed out by Pedersen et al. (2014b). They concluded that the first endogenously feeding nauplii stages of *Calanus finmarchicus* are more sensitive to CO$_2$-induced acidification than eggs or later nauplii stages (Pedersen et al. 2014b). Longer developmental times in high CO$_2$/low pH have been observed in crustaceans, echinoderms and molluscs (Cripps et al., 2014a and references therein). Weydmann et al. (2012) also reported a significant developmental delay for *Calanus glacialis* eggs when exposed to highly acidified conditions. Pedersen et al. (2014a) observed that development of C4 copepodites of *C. finmarchicus* was delayed by 8.9 days in high CO$_2$ treatments in comparison to control condition, when also the previous generation had been exposed to the same conditions.

We expected maternal effects to be most obvious in a high stress situation (high fCO$_2$ treatments), as seen for three-spined sticklebacks in a study testing the effects of global warming (Shama et al., 2014). Instead, egg hatching was higher and nauplii development faster in the maternal environment than in the Baltic water, when the maternal environment had a low fCO$_2$ (low stress). In high fCO$_2$ maternal environment the opposite response was observed, thus indicating that maternal effects are in fact weak and cannot compensate for the higher fCO$_2$ levels that correspond to near-future levels or that the eggs are damaged by the high fCO$_2$. This suggests that *Acartia* sp. and its reproduction are after all somewhat sensitive to ocean acidification. However, the effects were not as clear over the following weeks as in the beginning of the study, which may be due to an overall low egg number and large variation in hatching after Day 17, or due to acclimation of the copepods to the treatment conditions. In addition, the maternal effects seemed to weaken over time. This could be due to weakening condition of the mothers. In the absence of fish predators, zooplankton density, and especially *Bosmina* sp. (cladocerans) increased strongly in the mesocosms (Lischka et al., 2015). Senescence and food limitation were thus plausible problems for copepods, and a likely cause of weakening maternal provisioning. In addition, conditions in the Baltic Sea changed after Day 17 due to an upwelling event, which caused an increase in fCO$_2$ and decrease in pH (Paul et al., 2015). This might have made the Baltic conditions less favourable for copepod egg
development and evened out the differences between high $\text{fCO}_2$ mesocosms and the Baltic conditions.

A few studies have highlighted the importance of testing for transgenerational effects to avoid over- or underestimation of the effects of ocean acidification on copepods. Thor and Dupont (2015) found decreasing egg hatching of *Pseudocalanus acuspes* with increasing $\text{pCO}_2$. In addition, transgenerational effects alleviated the negative effects on egg production and hatching of the second generation when the mothers had been acclimatised to the same treatment. Also, reciprocal transplant experiment showed that the effect was reversible and an expression of phenotypic plasticity (Thor and Dupont, 2015). Contrary to the current study, Pedersen et al. (2014a) found no effect of the $\text{CO}_2$ environment on egg hatching or development of pre-feeding nauplii stages N1 and N2 in their multigenerational study using *C. finmarchicus*. However, the development time of larger nauplii and copepodite stages was increased by $\text{pCO}_2$, although the development delay was not detected in the following generation (Pedersen et al., 2014a). Vehmaa et al. (2012) studied combined effects of ocean acidification and warming, and found indications that negative effects on *Acartia* sp. reproductive success can partly be combated with maternal effects. The used pH treatments (-0.4 from ambient) were at the same level as the low $\text{fCO}_2$-treatments in this study (MC6, MC7), which makes the results of the two studies consistent.

The measurements of female copepod antioxidant capacity were done in order to provide possible additional information of the maternal provisioning on the offspring. A preferable practice in oxidative stress studies is to measure several of the four components consisting of free radical production, antioxidant defences, oxidative damage, and repair mechanisms (Monaghan et al., 2009). In the current study we only have the estimate for the defences, antioxidant capacity (ORAC) measurements, which makes our conclusions slightly more uncertain. However, an earlier study with the same species has indicated that at intermediate stress levels an upregulation of the antioxidant system enhances protection against oxidative damage, but at higher stress, the pro-oxidants may exceed the capacity of the antioxidant system and lead to oxidative damage (Vehmaa et al., 2013). In this study, upregulated antioxidant defence seemed to have a positive effect on offspring quality, as indicated by the positive correlation between female ORAC and egg hatching success. Higher ORAC in the two highest $\text{fCO}_2$ mesocosms in the beginning of the study could be a sign of an upregulated antioxidant system in a sudden stressful situation, whereas the lowest ORAC in the high $\text{fCO}_2$ treatments
at day 31 (Fig. 1c#) could be caused by prolonged stress and exhausted antioxidant defence. The change from positive to negative effect in the course of the study could explain why fCO$_2$ did not show a significant correlation with ORAC, whereas food quantity (TPC <55 µm) did. Ismar et al. (2008) showed that Acartia spp. development can be either slow or altered by certain algal groups causing death before the first copepodite or reproductive stage. A non-optimal diet could explain why higher food quantity would cause smaller adult female size, lower egg hatching success or lower antioxidant capacity. Skeletonema-diatoms had fairly high abundance in the mesocosms during the first days of the experiment when egg hatching success was lowest in every mesocosm, but then declined rapidly. Diatom-dominated phytoplankton composition has been shown to cause low copepod egg hatching success in the field (Miralto et al., 1999). Another quality aspect is the size and shape of the food, which may make it difficult to ingest or assimilate. From day 16 onwards, over 50% of chlorophyll a was in picophytoplankton (<2 µm) (Paul et al., 2015), which is too small for Acartia consumption (Rollwagen Bollens and Penny, 2003) could explain the observed contradictory effects of TPC. It is hard to explain why higher food quantity would otherwise cause smaller adult female size, lower egg hatching success or lower antioxidant capacity, unless it is nutritionally unbalanced or difficult to catch or assimilate. Since we did not study what the copepods preyed upon we can only speculate on diet quantity and quality. Satiated food conditions can strengthen the maternal or transgenerational effects. The transgenerational effects were of minor importance for hatching success in C. finmarchicus when exposed to long term high CO$_2$ and food limited conditions (Pedersen et al., 2014a). Long term stress and food limitation could thus also be the reason for weakening maternal effects in the current study.

We found body size (prosome length) to be negatively affected by high CO$_2$. The result seems to be mostly driven by the mesocosm with the highest fCO$_2$ (MC 8), where the adult Acartia sp. copepods were smallest on all the four sampling times that were included in the analysis (Days 10, 17, 24 and 45) (Fig. 1b). It takes ~8.5 days for a sixth stage nauplius of A. bifilosa to develop through the five copepodite stages and reach adulthood at 17°C (Yoon et al., 1998). According to the Bělehrádek’s temperature function it takes 12–15 days for VI nauplii to reach adulthood at 9–11°C (Bělehrádek, 1935; McLaren, 1966). The constants used in the equation (n=1008, a=8.701) were the same as used in Dzierzbicka-Glowacka et al. (2009) for the Baltic Sea Acartia spp. bifilosa. It is thus possible that the copepods could have developed through several stages causing the differences in prosome length between the treatments on Day 10.
Lowered pH may have increased copepods’ energy requirements and if energy is reallocated towards maintaining homeostasis, their somatic growth can be reduced. Pedersen et al. (2014a) found *C. finmarchicus* body size to be inversely related to $pCO_2$. They also found higher respiration rate under more acidified conditions, and claimed that increased energy expenditure via rising respiration and consecutive decreasing growth and reproduction could lower the energy transfer to higher trophic levels and thus hamper the productivity of the whole ecosystem (Pedersen et al., 2014a). This is especially alarming when considering the projected climate warming, since copepod size is negatively correlated with temperature (Foster et al., 2011). In addition to temperature, food quantity and quality can affect the copepod body size (Hart and Bychek, 2011), and create surprising combined effects with acidification. Garzke et al. (2016) reported an indirect positive effect of $pCO_2$ on copepod body size, which was explained by higher food availability when acidification acted as a fertilizer for phytoplankton. Temperature and food also interact because temperature affects the respiration and metabolism, thus the satisfying diet depends on temperature (Boersma et al., 2016). If high CO$_2$ treatment (MC 8) caused a developmental delay in maturation, as could be interpreted from the prosome length results (Fig. 1b), the maturation would have occurred at different temperature than in other mesocosms and possibly in non-optimal food conditions. Anyway, higher food quantity and quality would be expected to increase copepod size, contrary to our results. It is therefore possible that the used food quantity (TPC <55 µm) and quality estimates (C:N <55 µm) do not fully describe the diet that *Acartia* sp. was consuming in the mesocosms.

Adult copepods have in general shown robustness against acidification (Mayor et al., 2012, McConville et al., 2013), whereas eggs and nauplii appear to be more sensitive (Cripps et al., 2014b; Fitz et al., 2012). In addition, there seems to be notable differences in sensitivity between species. Nauplii production, adult female fatty acid content and antioxidant capacity (ORAC) of *Eurytemora affinis* were not affected by $fCO_2$ in the current mesocosm campaign (Almén et al., 2016). Similarly, Lewis et al. (2013) found differences in ocean acidification sensitivity between the species *Oithona similis* and *Calanus* spp. (*C. glacialis* and *C. hyperboreus*). They argued that *O. similis* is more sensitive to future ocean acidification than *Calanus* spp., because *O. similis* remains in the surface waters whereas *Calanus* spp. migrates vertically, and encounters a lot wider $pCO_2$ ranges daily than *O. similis* (Lewis et al., 2013). The same applies to *Acartia* sp. and *E. affinis* in our study area. Although *Acartia* spp. is exposed to natural variability in pH environment due to daily variations as well as due to staying at greater depths during the day (low pH in deep water), it does not reside as deep as *E. affinis*.
(Almén et al., 2014) and may therefore show higher sensitivity than E. affinis during the current mesocosm campaign (Almén et al., 2016).

The results obtained for Acartia sp. reproduction in the current study seem to contradict the results obtained for the Acartia sp. abundance determined in the mesocosms. Although our results indicate that Acartia sp. reproduction is in fact sensitive to ocean acidification, no fCO2 effect was found for the abundance of this species (Lischka et al., 2015). It is possible that 45 days was not long enough to detect small negative effects of CO2 on copepod size, egg hatching and nauplii development, to be reflected in copepod abundance. In addition, especially in the beginning of the study Acartia eggs in the mesocosms might have ended up in the sediment trap before hatching due to slow development at low temperature, which might have made it difficult to detect differences in Acartia abundance between the mesocosms. On a longer time scale, small acidification induced delays in offspring development could translate into negative effects for the copepod population, and further on energy transfer within the pelagic food web. In addition, warming will probably enhance the sensitivity of the species towards ocean acidification (Vehmaa et al., 2012, 2013).

5 Conclusions

Our results support the idea that it is important to look beyond egg production as hatching and development can be more sensitive to ocean acidification. Parental effects will likely be important in mediating some of the negative effects of ocean acidification. For Acartia sp., the transgenerational (maternal) effects may alleviate negative impacts of ocean acidification but only under exposure to medium levels of CO2. We did not find support for the hypothesis suggesting that poorer food quantity and quality would weaken the maternal effect by deteriorating the condition of the mother, which could be due to the overall food limitation especially during the latter half of the study or the fact that our estimates of food quantity and quality did not describe the diet in a satisfactory manner. Nevertheless, maternal antioxidant defence seems to correlate positively with offspring egg hatching success. Overall, these results indicate that Acartia sp. could in fact be affected by CO2 levels predicted for the year 2100 (IPCC, 2007). However, it is important to remember that this study shows how today’s copepods would react to tomorrow’s world; thus these results do not take into account the possible effects of evolutionary adaptation. Transgenerational effects can buffer short-term
detrimental effects of ocean acidification and thus give time for genetic adaptation and consequently assist persistence of populations under climate change.

Author contributions

A.V. planned the experiment; A.V., A.-K.A., J.E.-Ö., A.B. conducted the laboratory experiment; A.V. performed the statistical analyses; A.P. analysed TPC and C:N; S.F analysed ORAC; U.R. coordinated the whole project; A.V. and A.-K.A. shared responsibility of writing the manuscript with contributions from all co-authors.

Acknowledgements

We thank the KOSMOS team and all of the participants in the mesocosm campaign for their support during the experiment and the Tvärminne Zoological Station for their warm hospitality, support and use of facilities for this experiment. In particular, we would like to thank Andrea Ludwig for co-ordinating the campaign logistics and assistance with CTD operations, Silke Lischka and Bettina Grönlund for assisting with the zooplankton sampling, and the diving team. We also gratefully acknowledge the captain and crew of R/V ALKOR (AL394 and AL397) for their work transporting, deploying and recovering the mesocosms. This collaborative project was funded by BMBF projects BIOACID II (FKZ 03F06550), SOPRAN Phase II (FKZ 03F0611), and MESOAQUA (grant agreement number 228224), Cluster of Excellence ‘The Future Ocean’ (Project CP1141), and Academy of Finland (project nr. 276947).
References


Knuckey, R.M., Semmens, G.L., Mayer, R.J., and Rimmer, M.A.: Development of an optimal microalgal diet for the culture of the calanoid copepod *Acartia sinfinsis*: Effect of algal species


Table 1. The structure of the full LMM or GLMM models that were used to test effects of ocean acidification, food quantity, and food quality on copepod egg production (EPR), egg hatching success (EH), prosome length (PL), antioxidant capacity (ORAC), the ratio of EH mesocosm / EH Baltic, and the ratio of nauplii development index (DI) mesocosm / DI Baltic. The sampling days that were included in each of the models are listed. Repeated measures of same mesocosm bags was used as a random effect in all the models, because copepods that come from the same bags are more alike than copepods from different bags.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Fixed effects</th>
<th>Effect tested</th>
<th>Days included in the model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td><strong>EPR (LMM)</strong></td>
<td>fCO₂</td>
<td>Ocean acidification</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>TPC (&lt;55 µm)</td>
<td>Food quantity</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>C:N (&lt;55 µm)</td>
<td>Food quality</td>
<td>x</td>
</tr>
<tr>
<td><strong>EH (GLMM)</strong></td>
<td>fCO₂</td>
<td>Ocean acidification</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>TPC (&lt;55 µm)</td>
<td>Food quantity</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>C:N (&lt;55 µm)</td>
<td>Food quality</td>
<td>x</td>
</tr>
<tr>
<td><strong>PL (LMM)</strong></td>
<td>fCO₂</td>
<td>Ocean acidification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TPC (&lt;55 µm)</td>
<td>Food quantity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C:N (&lt;55 µm)</td>
<td>Food quality</td>
<td></td>
</tr>
<tr>
<td><strong>ORAC (LMM)</strong></td>
<td>fCO₂</td>
<td>Ocean acidification</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>TPC (&lt;55 µm)</td>
<td>Food quantity</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>C:N (&lt;55 µm)</td>
<td>Food quality</td>
<td>x</td>
</tr>
<tr>
<td><strong>EH MC/Baltic (LMM)</strong></td>
<td>fCO₂</td>
<td>Ocean acidification</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>TPC (&lt;55 µm)</td>
<td>Food quantity</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>C:N (&lt;55 µm)</td>
<td>Food quality</td>
<td>x</td>
</tr>
<tr>
<td><strong>DI MC/Baltic (LMM)</strong></td>
<td>fCO₂</td>
<td>Ocean acidification</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>TPC (&lt;55 µm)</td>
<td>Food quantity</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>C:N (&lt;55 µm)</td>
<td>Food quality</td>
<td>x</td>
</tr>
</tbody>
</table>
Table 2. T-statistics of the retained fixed effects in the linear mixed effect models testing the effects of TPC (<55µm), C:N and f/CO₂ on egg production rate (EPR), female prosome length (PL) and female antioxidant capacity (ORAC). Repeated measures of same mesocosm bags was used as a random effect in all the models, because copepods that come from the same bags are more alike than copepods from different bags.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Fixed effect</th>
<th>Estimate</th>
<th>DF</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPR</td>
<td>TPC &lt;55 µm</td>
<td>0.21±0.14</td>
<td>23</td>
<td>1.54</td>
<td>0.137</td>
</tr>
<tr>
<td>PL</td>
<td>f/CO₂</td>
<td>-0.0000027±0.000011</td>
<td>16</td>
<td>-2.39</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>TPC &lt;55 µm</td>
<td>-0.0037±0.0017</td>
<td>16</td>
<td>-2.21</td>
<td>0.042</td>
</tr>
<tr>
<td>ORAC</td>
<td>TPC &lt;55 µm</td>
<td>-0.0045±0.0021</td>
<td>22</td>
<td>-2.17</td>
<td>0.041</td>
</tr>
</tbody>
</table>
Table 3. Ranges of $f$CO$_2$, TPC <55 µm, and C:N < 55 µm that were used as explanatory variables in the full LMM and GLMM models. 3-day averages (measured within the latest three days of the sampling day) were used in testing the effects of the explanatory variables on copepod egg production (EPR), antioxidant capacity (ORAC), and egg hatching success (EH), whereas average of all measurements since the start of the experiments until the sampling day were used when testing the effects of the explanatory variables on copepod size (PL). Variations in $f$CO$_2$, TPC <55 µm, and C: <55 µm in the course of the study are presented in Paul et al. (2015).

<table>
<thead>
<tr>
<th></th>
<th>fCO$_2$ (µatm)</th>
<th>TPC&lt;55 µm</th>
<th>C:N &lt;55 µm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-d average</td>
<td>Average since</td>
<td>3-d average</td>
</tr>
<tr>
<td>MC 5</td>
<td>275–481</td>
<td>274–368</td>
<td>15.8–24.5</td>
</tr>
<tr>
<td>MC 6</td>
<td>663–991</td>
<td>683–896</td>
<td>16.5–34.3</td>
</tr>
</tbody>
</table>
Table 4. Z-statistics of the retained fixed effects in the GLMM testing the effect of $f$CO$_2$, TPC (<55 µm) and C:N on egg hatching success (EH). Repeated measures of same mesocosm bags was used as a random effect in the model, because copepods that come from the same bags are more alike than copepods from different bags.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Fixed effect</th>
<th>Estimate</th>
<th>z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EH</td>
<td>$f$CO$_2$</td>
<td>-0.00062±0.00032</td>
<td>1.94</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>TPC &lt;55 µm</td>
<td>-0.09557±0.02505</td>
<td>3.82</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Table 5. T-statistics of the retained fixed effects in the LMMs testing the effect of $f$CO$_2$, TPC (<55 µm) and C:N on ratio of egg hatching success (EH) mesocosm / EH Baltic and nauplii development index (DI) mesocosm / DI Baltic. Ratio >1: higher EH or DI in the mesocosm water (maternal environment) than in the Baltic Sea water, ratio <1: lower EH or DI in the mesocosm water (maternal environment) than in the Baltic Sea water. Repeated measures of same mesocosm bags was used as a random effect in both models, because copepods that come from the same bags are more alike than copepods from different bags.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Fixed effect</th>
<th>Estimate</th>
<th>DF</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EH mesocosm / EH Baltic</td>
<td>$f$CO$_2$</td>
<td>-0.00061±0.00028</td>
<td>16</td>
<td>-2.20</td>
<td>0.043</td>
</tr>
<tr>
<td>DI mesocosm / DI Baltic</td>
<td>$f$CO$_2$</td>
<td>-0.00145±0.00067</td>
<td>16</td>
<td>-2.15</td>
<td>0.047</td>
</tr>
</tbody>
</table>
Figures.

Fig. 1. Development of *Acartia bifilosa* a) egg production, b) prosome length (average ± s.e.), c) antioxidant capacity, and d) egg hatching success in the mesocosms, and e) egg hatching success in Baltic water when eggs are produced in mesocosms and d) antioxidant capacity in the mesocosms in the course of the study. The \( f\text{CO}_2 (\mu\text{atm}) \) values represent the average in Days 1–43 (Paul et al., 2015).

Fig. 2. Development of the ratio of a) egg hatching success (EH) mesocosm / EH Baltic and b) nauplii development index (DI) mesocosm / DI Baltic during the study. Ratio >1: higher EH or DI in the mesocosm water (maternal environment) than in the Baltic Sea water, ratio <1: lower EH or DI in the mesocosm water (maternal environment) than in the Baltic Sea water. Note that because of different development times, the DI values are not comparable between the days. The \( f\text{CO}_2 (\mu\text{atm}) \) values represent the average in Days 1–43 (Paul et al., 2015).

Fig. 3. Correlations of copepod egg hatching success (EH) with maternal antioxidant capacity (ORAC).
Hatching success (%) vs. ORAC (mg trolox eq. mg protein⁻¹)

- Black squares: Mesocosm
- Grey circles: Baltic

The graph shows a positive correlation between ORAC and hatching success, with Mesocosm data points generally showing higher hatching success compared to Baltic.