

Dear Editor,

The manuscript “Calcification and inducible defense response of a calcifying organism could be maintained under hypoxia through phenotypic plasticity” by Leung and Cheung, presents interesting questions about possible eco-physiological adaptations observed on a calcifying polychaete exposed to acute hypoxia, including changes in calcification rates, shell composition and metabolism.

Despite some interesting points, I think that the manuscript, in the present form is not acceptable for publication, because of a substantial lack of detail on the protocols used for the experiments and the analyses, very shallow description of the main results and some over-interpretation of the results. A (very) major revision is therefore suggested.

RESPONSE: We thank this reviewer for reviewing and providing useful suggestions for improving our manuscript. After reading the comments carefully, we found that most of them are related to the methods and can be clarified easily. However, there are some invalid arguments which are explained below. Overall, we are happy to make a major revision and believe that incorporating the good suggestions can greatly improve the quality of our manuscript.

Here listed some of the **major comments**:

1) The materials and methods are too undetailed. Are the specimens at the T0 adult or juveniles? Were they exposed to a day/night light cycle? If so, how did you control algal proliferation during the experiment? If not, are you sure that this does not influence their physiology? Were added algae dead or alive? What basis the 20 ml algal concentration was chosen on? Before organic matter analyses, were the samples washed? If so, how? Also for statistical analyses more details are needed.

RESPONSE: The specimens were adults and exposed to a day/light cycle of 14:10 hrs. Live algae were added, but there was no need to control their growth because the consumption rate of polychaetes is faster than the proliferation rate of algae. That is why we had to add algal suspension on a daily basis to sustain the growth of polychaetes (Ln 81-83). The feeding regime was based on our extensive experience in rearing *Hydroides* spp., which can ensure their normal growth under laboratory conditions. The shells were rinsed with deionized water before organic matter analysis. The above information will be added in the revision and more details about statistical analysis can be provided as suggested below.

2) I do not understand why in the materials and methods the authors say that they used 10 specimens per replicate (x 3 repl.) per treatment, but each observed response in the results is based on only n=3 or n=5 specimens (or replicates??). This is surprising if we consider that most of the performed analyses (e.g. respiration rates, growth rates, clearance rate) are not destructive and that just part of the shell is necessary for the rest of the analyses (e.g. Mg/Ca, organic matter, ACC...). So why did not the author use more than 3 or 5 specimens to do their analyses and calculate averages? Could they add more replicates?

RESPONSE: We adhere to the classic, unambiguous definition of “n” as “number of replicates” (i.e. not “number of individuals”). To clarify the sample size used, we initially had a total of 30 individuals per treatment (Ln 101-102, 105). Considering the mortality following the 3-week exposure, 25 individuals were made into 5 replicates with 5 individuals per replicate for respiratory rate and clearance rate measurements (Ln 149 and 158). Please note that one individual per replicate is unable to give adequate precision and reliability of measurements. Therefore, multiple individuals were pooled and each replicate is represented by an average of the pooled individuals (the units of these variables were expressed as “per individual”). Based on our experience and preliminary test, this design can already give adequate precision of measurement and statistical power. This is also evident from the small standard errors in our results (see Fig. 1 and 4).

3) The experimental design is quite weak. Maybe this impression derives also from the difficulty to understand how many individuals/measures/replicates were performed for each parameter. We lack fundamental information about the ability of these organisms to develop into cultures. All the results are discussed on the basis of comparison to normoxic conditions. How can the authors be sure that a “culturing effect” is not interfering with the results? This is particularly true for shell composition. Why did not the author compare the shells of individuals grown in normoxic conditions to pre-experiment portions of shells (grown into natural conditions) to see if any “culturing effect” is visible?

RESPONSE: Please see the above response for the issue of sampling design. As stated in the Introduction, this study aims to elucidate the effects of hypoxia on calcification and defence response of a calcifying polychaete, which can be experimentally tested under fully controlled laboratory conditions in order to isolate the effects of hypoxia and minimize the number of confounding factors. In contrast, the “culturing effect” cannot answer our research question because it only indicates the effects caused by the difference between laboratory and field conditions, but not the effects of hypoxia. As all the individuals were reared under laboratory conditions, the results were not influenced (or biased) by “culturing effect”.

4) The protocol used for respiration measurements is unusual to me. I cannot understand how the authors did measure the oxygen content of hypoxic waters into a syringe with a relatively thick probe tip and can be sure they avoided oxygenation during the measurement. Also, are they sure that the material the syringe is made of is impermeable to oxygen? Do they have any measure of blanks to estimate the possible gas exchange through the syringe walls during the analysis?

RESPONSE: We have to invalidate this comment. Our DO probe is thin (~2.4 mm in diameter) enough to perfectly insert into the hole of plastic syringe (~2.7 mm in diameter) for DO measurement. The syringes used were ordinary plastic syringes (Terumo Hypodermic Syringe without a needle) with a barrel wall which is definitely thick enough to be impermeable to air. Once the syringe is sealed (Ln 151-152), oxygenation cannot occur under such air-tight conditions. This simple method has been widely used before (e.g. Zhao et al., 2011; Leung et al., 2013).

Leung, Y.S., Shin, P.K.S., Qiu, J.W., Ang, P.O., Chiu, J.M.Y., Thiagarajan, V., and Cheung, S.G.: Physiological and behavioural responses of different life stages of a serpulid polychaete to hypoxia. *Mar. Ecol. Prog. Ser.*, 477, 135–145, 2013.

Zhao, Q., Cheung, S.G., Shin, P.K.S., and Chiu, J.M.Y.: Effects of starvation on the physiology and foraging behaviour of two subtidal nassariid scavengers. *J. Exp. Mar. Biol. Ecol.*, 409, 53–61, 2011.

5) In the experiment the authors consider hypoxic conditions to be reached at ~2.0 mg O₂/L. This concentration is the one which is normally considered to be the upper limit for hypoxia. Considering that the error associated to the probe used for O₂ survey is 0.1 mg/L (source: TauTheta manual) and that the average values the authors report in table S1 are always a slightly above 2.0 mg/L for hypoxic conditions, I wonder why the authors did not test a lower concentration to be sure to never exceed oxygen concentrations corresponding to namely hypoxic conditions. The results obtained should be carefully presented as the response to a case of acute (short time) and slight (upper limit) hypoxia and every over-interpretation or generalization to strong or long-term hypoxia should be avoided.

RESPONSE: There is no solid definition for the upper limit of hypoxia, but we used 2.8 mg O₂/L (~2.0 ml O₂/L) as the upper limit, which is defined in some influential papers and has been widely applied (Wu, 2002; Diaz and Rosenberg, 2008). Therefore, 2.0 mg O₂/L used in our study is definitely low enough to be considered hypoxic according to the conventional standard. We can mention that the results indicate the impacts of short-term hypoxia on calcification.

Diaz, R.J., and Rosenberg, R.: Spreading dead zones and consequences for marine ecosystems. *Science*, 321, 926–929, 2008.

Wu, R.S.S.: Hypoxia: from molecular responses to ecosystem responses. *Mar. Pollut. Bull.*, 45, 35–45, 2002.

6) In lines 96-97 authors say that part of the individuals was left into hypoxic conditions during one week before the experiment (after damage and measures of the tubes). Can you please justify this choice? Does this mean that the T0 for tube sizes represents instead one week under treatment conditions? Did the authors measure the T0 size again after the acclimation and before the start of the experiment?

RESPONSE: This acclimation period is needed to remove the fight-or-flight response after shell breaking. The approximate tube length after shell breaking is provided (Ln 93), but it does not represent the initial tube length. We measured the tube length on Day 1 of the exposure period as the initial tube length (Ln 100).

7) How did the authors get to have stable oxygen conditions during and just after water changes (every 3 days)? Did you measure oxygen into the culture solution before retiring old seawater? Could you please show these data somewhere in supplementary materials?

RESPONSE: Stable DO concentration at hypoxic level can be obtained by aerating the seawater continuously with a mixture of nitrogen and air (Ln 105-107), where the flow rates of these two gases were adjusted to achieve the target DO level (Ln 85-88). Stable equilibrium between gases in seawater can be achieved by this method quickly so that the DO concentration in seawater can be very stable throughout the 3-week exposure period. This method has been widely used for hypoxia study (e.g. Leung et al., 2013; Mukherjee et al., 2013). The DO concentration of seawater was measured daily.

Leung, J.Y.S., Cheung, S.G., Qiu, J.W., Ang, P.O., Chiu, J.M.Y., Thiyagarajan, V., and Shin, P.K.S.: Effect of parental hypoxic exposure on embryonic development of the offspring of two serpulid polychaetes: Implication for transgenerational epigenetic effect. *Mar. Pollut. Bull.*, 74, 149–155, 2013.
Mukherjee, J., Wong, K.K.W., Chandramouli, K.H., Qian, P.Y., Leung, P.T.Y., Wu, R.S.S., and Thiyagarajan, V.: Proteomic response of marine invertebrate larvae to ocean acidification and hypoxia during metamorphosis and calcification. *J. Exp. Biol.*, 213, 4580–4589, 2013.

8) Is edx analysis resolution enough for consistent Mg/Ca measurement in the shell? Can the authors give more information about the accuracy, precision, detection limits etc. of the analysis, please? Or add references if the protocol is routinely used.

RESPONSE: EDX (or EDS) has been extensively used to analyse Mg/Ca in the shell with good results. We can add references in the text (e.g. Ries, 2004; Zhang et al., 2010).

Ries, J.B.: Effect of ambient Mg/Ca ratio on Mg fractionation in calcareous marine invertebrates: A record of the oceanic Mg/Ca ratio over the Phanerozoic. *Geology*, 32, 981–984, 2004.

Zhang, F., Xu, H., Konishi, H., Roden, E.E.: A relationship between d_{104} value and composition in the calcite-disordered dolomite solid-solution series. *Am. Mineral.*, 95, 1630–1656, 2010.

9) Some of the conclusions/discussions are inferential and not supported by the data. For example paragraph from line 249 to 258 should be deleted, in my opinion, as it is not supported by the presented data.

RESPONSE: This paragraph is a general discussion on the mechanism affecting calcification and our results (e.g. seawater carbonate chemistry and shell growth) can support this discussion. This paragraph is important to reshape the concept of calcification and to improve the readership of our manuscript.

10) The hypothesis of relaxed magnesium regulation to explain higher Mg/Ca in the calcite produced under hypoxia is based on benthic foraminifera. These organisms are known to strongly discriminate against magnesium. This does not seem the case for polychaetes, which seem to contain very high concentrations of Mg in the shell. Authors should base their hypotheses on more adapted literature.

RESPONSE: We cannot validate this comment. The Mg/Ca in calcite varies greatly among foraminifera species, where many of them have Mg/Ca greater than 15 mol %, such as *Spiroloculina clara*, *Planispirinella exigua*, *Peneroplis proteus*, *Alveolinella quooii* and *Mychostomina revertens* (Blackmon and Todd, 1959). Please also note that the transport and sequestration of Mg are heavily regulated in eukaryotic cells and the underlying mechanisms are likely conserved among eukaryotic species. Thus, it is reasonable to conjecture that similar mechanism can be shown in our tested calcifying polychaete. In fact, the currently acknowledged regulatory mechanism of Mg in foraminifera is partially inferred from the combined biological knowledge on *Paramecium*, mollusks and cultured rodent cells (see Bentov and Erez, 2006).

Blackmon, P.D., and Todd, R.: Mineralogy of some foraminifera as related to their classification and ecology. *J. Paleontol.*, 33, 1–15, 1959.

Bentov, S. and Erez, J.: Impact of biomineralization processes on the Mg content of foraminiferal shells: A biological perspective. *Geochem. Geophys. Geosyst.*, 7, Q01P08, 2006.

Minor comments/suggestions:

Line 30: use “increase” instead of “augment”

RESPONSE: Suggestion will be adopted.

Line 51: replace “It” by “This”

RESPONSE: Suggestion will be adopted.

Line 53: What does “or other physiological processes via energy trade-off” mean? Can you explain what other processes you’re talking about please?

RESPONSE: Based on the energy budget model, other processes can include growth, reproduction, somatic maintenance, etc. We will revise this sentence for clarity.

Line 58: Please delete “and defense response”.

RESPONSE: “Defense response” has a specific meaning in this study, which refers to the response following non-lethal shell damage.

Materials and methods: Please add titration protocol for alkalinity measures presented in table S1.

RESPONSE: It is unnecessary to add a protocol for alkalinity measurement because it is far too technical and general readers are not interested in how to operate a particular model of titrator. We have never seen research articles, except method papers, describing the operating procedures for a particular model of equipment. If the reviewer feels interested in the titration protocol for this model, here is the user manual: <https://www.manualslib.com/manual/530078/Hanna-Instruments-Hi-84431.html#manual>

Lines 93-97: Please add more details about the procedure used to measure the specimens, associated errors and discuss the potential stress that this manipulation may represent and the effect it could have on the final results

RESPONSE: We can add these details in the revision. However, please note that instant stress due to fight-or-flight response would be induced when breaking the tube. To reduce this stress, the individuals were allowed to rest for a week prior to experimentation (Ln 94) and therefore they were only under stress due to non-lethal shell damage, which is a factor in our experiment.

The order of paragraphs in the Material and methods section is, at present, a bit confusing and should follow a more logical pattern. I would suggest that the paragraph on experimental design should show the setting, the replication, times etc., for all type of analyses. Then a paragraph on procedures for physiological analyses should explain all the used methods for respiration rates, survival rates, feeding rates and shell growth. Then a final paragraph on the shell composition should follow.

RESPONSE: We appreciate reviewer's suggestions and will change the order accordingly.

Lines 99-100: Why is the color of new shell very different from the ancient one? Is it normal? Maybe you should discuss it somewhere. One would think that it is a culturing effect on shell structure...

Figure S1 should be in the main text.

RESPONSE: The colour of new shells should be white due to the presence of calcium carbonate. Yet, the colour will turn slightly yellowish over time because of the biofilm (e.g. bacteria, algae, etc.) growing on the surface. It is very normal and important because the biofilm allows polychaete larvae to settle and form a colony. We can put Figure S1 in the main text.

Lines 103-112 should be part of the "Experimental set up" paragraph

RESPONSE: We will change the order of paragraphs as suggested above.

Line 123: The specimens used for organic matter composition are the same measured for toughness? Please detail this kind of information

RESPONSE: It cannot be, unfortunately, because removing the organic matter in the shell at 550°C can substantially affect shell toughness. We will revise the sentence for clarity.

Paragraph 123-129: 1) for the composite shell powder did you mix calcite from different specimens, isn't it? Did they come from a same replicate for a treatment or even replicates were mixed? How did you prepare the powder exactly? Did you wash the shell before to avoid contamination (from algae given as food for example)? How? Why did you choose Mg/Ca ratio as a parameter to be measured?

RESPONSE: For each composite sample, shells from 3-5 individuals in the same replicate were used (Ln 124) so that we had 3 replicates per treatment for the geochemical properties. The powder was prepared by removing the newly-produced shells, rinsing them with deionized water (to remove the microalgae and other debris), drying them at room temperature and finally grinding them using a mortar and pestle. Many calcifying organisms can change Mg/Ca in their calcitic shells in response to the changing environment (e.g. ocean acidification). The underlying mechanism remains largely unknown, but may be associated with metabolic energy that can be greatly affected by hypoxia. Therefore, we expected that Mg/Ca would change in response to hypoxia.

Paragraph 148-155: How did you calibrate oxygen probes before the analysis?

RESPONSE: Before the analysis, the probe was calibrated by inputting the value of calibration slope in the software. Then, we have to validate the DO concentration using another DO meter (e.g. handheld DO meter), which is calibrated by measuring the DO concentration of oxygen-saturated seawater. We

may need to change the value of calibration slope until the DO concentrations between DO probes are the same. For the operation of the instrument, please refer to the user manual: https://in-situ.com/wp-content/uploads/2015/05/Stable_Optical_Oxygen_System_-SOO-100_Manual.pdf

Line 152: “The air inside the syringe...”: what air? Why is it there during the first measure? This part is not clear to me.

RESPONSE: The air is atmospheric air. For the initial measurement, the individuals were allowed to rest in the syringe for 15 min and the small pocket of air can help buffer the change in DO concentration during this period. We can add this information for clarity.

Line 154: “by gently stirring the FSW inside the syringe”. How did you avoid oxygenation at this step?

RESPONSE: Such gently stirring can only help homogenize the DO concentration in the water body. Oxygenation of water is caused by dissolving atmospheric oxygen in water. However, this process cannot occur because atmospheric air cannot interact with the water inside the syringe, which is under air-tight conditions (Ln 151-152).

Line 156: Please specify the unit used for consumed oxygen. Normally mL or μmol are used. In your figure 4a you use $\mu\text{g O}_2$, which is quite unusual as a unit for oxygen.

RESPONSE: For the unit of consumed oxygen, mg O_2 is frequently used, while $\mu\text{g O}_2$ is just a conversion for the magnitude to avoid many decimal places.

Lines 159-160: Why did you only used one algal species for this experiment? Can you add fundamental details such as if the experiment was performed under light conditions, please? Also you say that 5 replicated bottles were used per treatment. In the first materials and methods section you say you have 3 replicated bottles with 10 specimens for the experiments. I’m lost... Are these different bottles? How many individuals per bottle per treatment do you have then?

Statistical paragraph: Where data transformed before the analyses to homogenize the magnitudes? How many permutations were performed? Which distance parameter was used and why? Can you specify the “aforementioned parameters” of line 168, please?

RESPONSE: One algal species should be used for the feeding experiment to avoid the selective feeding of filter feeders (due to different sizes and/or textures between microalgae), which can complicate the calculation and interpretation. We can add a sentence to state that the experiment was performed under light conditions. We used different bottles for the feeding and shell growth experiments (5 replicate bottles with 5 individuals per bottle for the feeding experiment, Ln 156-158; 3 replicates bottles with 10 individuals per bottle for the shell growth experiment, Ln 102 and 105). For univariate analysis, data are on the same scale (i.e. carrying the same unit), therefore transformation is unnecessary. The number of permutation is 999 and Euclidean distance was used which is commonly applied. We will add these details in the revision. We will also list the parameters (e.g. respiration rate, clearance rate, shell toughness, etc.) in the revision.

Line 172: You say hypoxia slightly hindered, but your statistics say this difference is significant, so maybe this should be emphasized a bit more.

RESPONSE: Please note that “statistically significant” does not necessarily imply “biologically significant”. We can emphasize the statistical significance a bit more, but it is more important to compare the two factors with respect to the magnitude of change, which is of biological importance.

Line 174: Please replace “negligible” with “no”.

RESPONSE: “no” is too absolute. We prefer to be conservative by using “no observable” or “no significant”.

Line 180: You say “but only slightly by non-lethal shell damage”. It looks like a significant difference, visually (fig. 4a). Is it confirmed by statistical analyses? Yes, so you should say it!

RESPONSE: It is statistically significant (Table S2). We can revise this sentence in a more explicit way.

Lines 181-182: specify whether statistical differences are visible and where.

RESPONSE: We can specify this in the revision.

Line 187: what do you mean by “ramifications”?

RESPONSE: It can mean changes in species populations, community structure and ecosystem functioning. We can make the meaning more explicit.

Line 188: “tolerant to hypoxia”, at what temporal scale?

RESPONSE: Short-term scale.

Line 194: I would suggest not to use “unthreatened conditions” as a general for “shell damage”, because hypoxia also is an unthreatened condition, so it can result confusing.

RESPONSE: We appreciate this suggestion, but we have explicitly defined “unthreatened conditions” as “individuals without shell damage” in the methods and mentioned the definition again in this sentence. There should be no confusion.

Line 195: “hypoxia slightly hinders”... again, is it significant or not?

RESPONSE: Significant (see Table S2).

Lines 201-202: redundant concept, already said.

RESPONSE: We can delete this sentence.

Lines 205-206: Please replace “shell growth” with “inorganic components of the shell”. You suggest that under hypoxia the production of organic matter compensate for diminished quality of inorganic components (>ACC). Palmer (1992) on the contrary suggests that organic matter production is costly and that would be the reason why high-organic calcareous microstructures became rare with evolution. How do you explain this incoherence?

RESPONSE: Suggestion for the word choice can be adopted. However, we never suggest that “under hypoxia the production of organic matter compensate for diminished quality of inorganic components (>ACC)”. In this paragraph, we discussed the effects of hypoxia on shell growth and shell strength under unthreatened conditions. As the organic matter content was not affected by hypoxia, we suggest that the polychaetes still allocate similar amount of energy to maintain shell strength at the expense of shell growth under hypoxia.

Line 211-212: Could the overproduction of organic matter be related to higher calcification rates?

RESPONSE: Although we cannot rule out this possibility, we think it is still premature to make this speculation because we found that organic matter content does not necessarily increase with calcification rate (Normoxia vs. Hypoxia without shell damage).

Lines 219-221: I do not understand what you mean

RESPONSE: This sentence means when the individual is under life-threatening conditions and chance of survival becomes very low, it has to prioritize defense response as the last resort to maximize survival rate. We will rephrase this sentence to better illustrate the idea.

Line 224: You say the effect of hypoxia on defense response is not discernible. Although toughness is not affected, I would say that reduced shell growth rates (visible in figure 1) should be taken into account, to discern the effects of hypoxia, as well.

RESPONSE: We will revise this sentence by considering the effect of hypoxia on shell growth.

Line 234: please replace “signifies” with “may suggest”.

RESPONSE: Suggestion will be adopted.

Line 241: At the end of the sentence you should add “.. can generally be maintained at least under slight hypoxia, on a short timescale.”

RESPONSE: Suggestion will be adopted.

Line 248: please delete “and therefore *H. diramphus* prioritized defense response”.

RESPONSE: Suggestion will be adopted.

Line 259: please add “open” between marine and waters (because “costal” are also marine waters!)

RESPONSE: Suggestion will be adopted.

Line 264: when you say that the defense response can be sustained you should specify “on a short timescale”, because your results are based on short-term experiments.

RESPONSE: We appreciate this suggestion and will specify this in the text.

Lines 330-339: Please order the references on a chronological basis.

RESPONSE: Suggestion will be adopted.

Line 338 and 368: add authors to the list

RESPONSE: We will revise them according to the journal’s latest style.

Survival rate figure (S2) should not be in supplementary material, as this is an important result to take into account in the analysis of all the others.

RESPONSE: We will put this figure in the main text.