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.

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.
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?
3) The experimental design is quite weak. Maybe this impression derives also from the difficulty to understand how many individuals/measures/replicates where 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?
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?
5) In the experiment the authors consider hypoxic conditions to be reached at ~2.0 mg O_2/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 O_2 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.

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?
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?
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.
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.
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.

Minor comments/suggestions:

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

Line 51: replace “It” by “This”

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

Line 58: Please delete “and defense response”.

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

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

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., <u>for all type of analyses</u>. 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.

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.

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

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

Paragraph 123-129: 1) for the composite shell power 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?

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

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

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

Line 156: Please specify the unit used for consumed oxygen. Normally mL or µmol are used. In your figure 4a you use µg O_2, which is quite unusual as a unit for oxygen.

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?

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

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

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!

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

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

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

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

Lines 201-202: redundant concept, already said.

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?

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

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

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.

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

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

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

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

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.

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

Line 338 and 368: add authors to the list

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.