# Response to Reviewers, PLOS Climate

*Narrowed gene functions and enhanced transposon activity are associated with high tolerance to ocean acidification in a juvenile subarctic crustacean*

Laura H Spencer, William Christopher Long, Ingrid B Spies, Krista M Nichols, Robert J Foy

## Reviewer 1 Comments

**COMMENT: Line 78 – 81: are these studies showing deleterious effects on short term exposure? Is that why results differ to long exposure I Long et al? Please indicate.**
Good catch, it wasn't clear how these previous studies differed from Long et al. 2023. I have edited the text for clarity:

*Laboratory studies on embryos, early zoea, and juveniles indicate that OA negatively affects early red king crab life stages such that development is altered, and growth and survival typically decreases [50–52]. OA also interacts with warming by increasing intermolt duration and decreasing survival in juveniles [53]. … Interestingly, Long et al. [55] recently found that when red king crabs are exposed to OA during multiple stages of early development – from hatch through the early juvenile stage – they are surprisingly tolerant of moderate (pH 7.8) and severe (pH 7.5) acidification, with no change in survival, growth, or development [42].*

**COMMENT: Line 86: remove "Here, we"**
Thank you, that has been removed!

**COMMENT: Line 101-106: rephrase, it comes across as methods.**
We rephrased this sentence, as well as re-organized it within the introduction based on another reviewer's comment:

*We leverage juveniles from Long et al. [55] that were exposed to (and tolerated) three carbonate chemistry treatments from hatching to the first crab stage (C1), thus capturing transcriptional differences among crab that are reared in historically ambient conditions along …*

**COMMENT: Line 141-145 = results**
The paragraph describing how experimental water parameters were maintained and the accompanying table (now Table 2) have been moved to the Results section entitled "Experimental Design."

**COMMENT: Table 1: Is this data pooled for each treatment among the replicates? It would be good to see this data broken down by tank replicate in the supplementary material to see if something untoward may have occurred in a replicate.**

Yes, this table (now referred to as Table 2) shows mean +/- sd of water parameters measured or calculated during the experiment, pooled across all tanks for each treatment. We have added a new table to the supplemental (Supplemental Table 1) which includes the means for each tank.

**COMMENT: Line 138, Supplementary figure S1: you mention elsewhere that it was a three-month exposure trial but you only have two months of time series data in the supplemental material? Sample size? For analyses?**

Good catch! We have updated Supplemental Figure 1 to include the temperature and pH data that were collected daily throughout the larval collection/stocking and exposure period (from April 10th - July 8th). Data for the remaining parameters, which were measured or calculated using weekly samples, begins during week 2 of the exposure period (April 26th). We have clarified this in the Methods section and Table 2 caption (see updated text below). Regarding the comments "Sample size?" and "For analyzes?", I presume you are referring to N used to calculate the water quality parameter summary statistics shown in Table 2. We have added that information.

In text: *"Water samples were taken once a week from each tank beginning the second week of exposure, poisoned with mercuric chloride … "*

Table 2 caption. *"Water parameters during the experiment. Temperature and pH were measured daily; salinity, dissolved inorganic carbon (DIC) and alkalinity were measured weekly beginning on the second week of the exposure period; all other parameters were calculated. Values are mean ± standard deviation, calculated across replicate tanks per treatment (N=430 for pH and temperature, and N=53 for other parameters). See Supplemental Figure 1 for water parameter time-series and Supplemental Table 1 for per-tank means."*

**COMMENT: Line 5: what were the phenotypic results?**

It is not clear which line you are referring to here (Line 5 is the title page), so we have incorporated phenotypic results of Long et al. 2023 to the introduction:

*Interestingly, Long et al. [55] recently found that when red king crabs are exposed to OA during multiple stages of early development – from hatch through the early juvenile stage – they are surprisingly tolerant of moderate (pH 7.8) and severe (pH 7.5) acidification, with no change in survival, growth, or development [42].*

# Reviewer #2 Comments:

## Overarching Comments:

**COMMENT: Connections to other crustacean literature: The focus of the manuscript currently is RKC-centric. While connecting this study with previous RKC studies is important, I think the manuscript would be of interest to a wider audience if the authors discussed their findings in the context of other crustacean OA studies. A handful of studies have examined molecular mechanisms involved in crustacean responses to OA. I've included some suggested references below (one physiology study, two molecular studies) that may be useful for the authors.**

Thank you for providing the additional references. We have added a paragraph that expands on the effects of OA on crustaceans, focusing on studies that use molecular approaches.

**COMMENT: Functional Analyses: It was unclear which enrichment analyses were completed, and which datasets were used for those analyses, when reading the methods section alone. Consider restructuring the methods section and adding subheadings to match the results section for increased clarity. In lines 276-279, the authors state that two analyses were performed for each pairwise contrast. What were these analyses?**

The Functional Analyses section in the Methods has been restructured to more clearly outline the different gene sets examined for enriched processes. We also have edited the section describing the functional analyses of differentially expressed genes:

*Differentially expressed genes. To determine the functions of differentially expressed genes two enrichment analyses were performed for each of the three pairwise treatment contrasts: (1) genes that were upregulated (L2FC > 0.5) and (2) genes that were downregulated (L2FC < -0.5) in response to the more severe OA treatment.*

## Specific Comments:

*Introduction:*

**COMMENT: Line 57: When does the Bering Sea experience lower aragonite saturation, and which part of the crab life cycle does it coincide with?**
We agree that this information was lacking in the manuscript- we have added the following text to the introduction for more context regarding when RKC early life stages may encounter low carbonate concentrations:

*Early life stages, which are thought to be particularly vulnerable to OA, are present as brooded embryos year-round and from hatch in late spring through settlement in fall [48,49] and may encounter low carbonate saturation states that already occur in bottom waters and during seasonal upwelling [13].*

**COMMENT: Line 61: Replace first "]" with ","**
This change has been made.

**COMMENT: Lines 73-75: What have existing mechanistic crustacean OA studies demonstrated?**
We have added a paragraph that expands on the effects of OA on crustaceans, focusing on studies that use molecular approaches:

*The diversity of responses in crustaceans, and all marine invertebrates for that matter, has highlighted the need for a more mechanistic understanding of how organisms function in OA conditions. Targeted gene expression analyses in green shore crab (*Carcinus maenas*) gill tissue reported that genes involved in acid-base regulation are affected by high $pCO_2$ [39,40]. Transcriptome-wide expression analysis identified changes in genes involved in energy metabolism and apoptosis activity in shrimp (*Exopalaemon carinicauda*) [41], and immune functions, energy metabolism, and ion transport in the Chinese mitten crab (*Eriocheir sinensis*) [42,43]. Metabolomic analyses in juvenile Dungeness (*Cancer magister*) and green shore (*Carcinus maenas*) crabs both found amino acid metabolism to respond to acidified conditions, perhaps due to increased buffering needs [44,45]. As with whole-animal metrics, molecular effects of OA can vary by life stage. Metabolite analyses of American lobster larvae (*Homarus americanus*) indicate that large metabolic shifts (fatty acids, amino acids, and citrate cycle) perhaps underlie their buffering capacity and tolerance of OA, whereas metabolic reprogramming is not observed in the more sensitive juvenile stage [46]. Together, molecular assays to date reveal that OA induces changes in energy metabolism, acid-base regulation, immune function, and cellular stress-response processes, with the strongest responses perhaps reflective of physiological reprogramming. Changes are, however, somewhat unpredictable due to species-, stage-, and exposure-specific effects. For that reason, the most informative studies pair molecular assays with survival and growth data preferably from the same individuals.*

**COMMENT: Line 86: Delete "Here, we"**
This change has been made.

**COMMENT: Lines 88-89: Consider replacing "genes and functions" with "molecular mechanisms and pathways" to encompass DEG and TE information.**
Great suggestion! That sentence now uses the terms "molecular mechanisms and pathways".

**COMMENT: Lines 89-106: Consider restructuring this section of the introduction for improved flow and clarity. Move lines 101-106 after lines 87-89 to introduce the experimental design in the previous study. Then, modify lines 89-91 to explicitly define the molecular approaches used by the study/the different ways RNA-Seq data was leveraged.**
The concluding section of the introduction that describes the experimental design and RNA-Seq has been restructured for clarity, succinctness, and to improve flow. Thank you for pointing this out!

**COMMENT: Line 94-96: Consider moving to line 75, as that paragraph discusses previous crustacean OA literature.**
The sentence describing the Stillman et al. study, which is the one previous transcriptomics study in red king crab, has been relocated. We have opted to keep it within the paragraph that introduces red king crab to discuss studies on the species together. We hope that the overall reorganization of the introduction and additional paragraph with more mechanistic background in crustaceans has improved the flow and bolstered the background.

**COMMENT: Lines 100-101: Leveraged the genome to do what?**
After restructuring the last paragraph of the introduction this sentence was no longer needed, so it has been removed.

*Methods:*

**COMMENT: Lines 141-149: Consider moving this information to an Experimental Design section in the results.**
The paragraph describing how experimental water parameters were maintained and the accompanying table have been moved to the Results section entitled "Experimental Design."

**COMMENT: Lines 218-220: Consider modifying Table 2 to include how many samples were used for each analysis (DEG, SNPs, etc.)**
I have edited the caption for the table containing RNASeq sample size (now "Table 1") to clarify that the sample sizes listed were used in all analyses:

*Table 1 RNA-Seq sample size and replication by treatment after the removal of one outlier sample from the Ambient treatment. All libraries listed were used in all analyses (genetic, global patterns, gene co-expression, differential expression, and expression variation).*

**COMMENT: Lines 220-222: Why not use an RDA, or something similar, to understand how genetic variation constrains gene expression variation?**
Our primary goal of the genetic analysis was to determine whether selective mortality occurred during the experiment, resulting in differing genetic composition among OA treatments that could have influenced expression differences. In addition to the analyses included in the paper (parentage, Fst among treatments, genetic composition using PCA, correlations among PC axes that were generated from genetic and expression PCAs), we examined correlation between pairwise genetic distances and expression distances. The RDA approach is a very good suggestion that we hadn't yet used! As part of our revision process, we explored this option by performing an RDA in addition to a distance-based RDA to understand how genetic variation explains expression of genes that responded to acidification. We used counts of all differentially expressed genes as the response variables (n=6,806 DEGS) and PCs 1-5 from the genetic PCA as the explanatory variables. The RDA and dbRDA found that genetic PC axes explained a total of 13.6% and 11.7% of the variance in expression, respectively, but permutation tests of results did not find significance of the RDA/dbRDA models, explanatory variables, or canonical axes. We have opted to not add this additional analysis to the

manuscript as the interpretation is similar to our other correlation-based analyses – gene expression differences among OA treatments are not likely due to genetic composition differences among treatments. We will certainly add this multi-omics data integration method to our repertoire!

**COMMENT: Lines 249-254: Can be hard to follow methods for gene expression variation calculations when DEG analysis is not yet described. Consider incorporating the relevant gene-wise variation analyses into the global gene expression and DEG analysis sections.**
We have reorganized the methods and results sections to describe the coexpression and differential expression analyses before the variation analysis. The methods and results sub-sections also now are in the same order.

**COMMENT: Lines 278-279: Specify L2FC thresholds in the "Differential gene expression analysis" section, as I assume the same thresholds listed here were used with DESeq2.**
The approach used (DESeq2) does not use a minimum $|L2FC|$ to define a differentially expressed gene. However, we did opt to filter for $|L2FC| > 0.5$ prior to interrogating the DEG sets. The following information was added to the Differential gene expression analysis section:

*No minimum log2 fold change (L2FC) was used to identify differentially expressed genes, but DEGs they were filtered for those with $|L2FC| > 0.5$ prior to functional analysis.*

**COMMENT: Lines 288-289: I would argue that searching for transposable elements wouldn't constitute an enrichment analysis, but a separate functional analysis. Consider changing the wording and adding a descriptive subheading to match this.**
There are now subheadings in the "Functional Analyses" methods section that first describes the enrichment analyses on the various gene sets and the transposable element composition analysis.

**COMMENT: Line 291: Restate the gene sets used for clarity.**
The gene sets have been added for clarity:

*The proportion of genes that were transposable elements was estimated for each gene set described above (co-expressed gene modules, differentially up/down-regulated genes in each pairwise contrast, and low-variance genes) by searching within the protein names of annotated genes for the words …*

**COMMENT: Lines 290-294: Why not use RepeatMasker to search for transposable elements in addition to the manual curation already performed?**
Our manual curation approach, which identified genes in our gene sets of interest that were TE's, was performed at the final step of the analysis once we identified DEGs, co-expressed genes, and low-variance genes that were annotated by the Uniprot/Swissprot database. We feel that the comprehensive annotation information pulled from the Uniprot/Swissprot database was sufficient for TE identification. While it would be interesting to also search for transposable elements in the un-annotated genes using RepeatMasker, those genes were not included in any

of the other functional analyses, so we opted to focus on the reduced dataset for all functional analyses.

***Results:***

**COMMENT: Lines 327-329: ...mapped to non-coding regions (22.3%), or were assigned…**
Indeed, the "or" has been added.

**COMMENT: Line 335: The high number of genes that were removed? That had reads mapped to them?**
The following text was added for clarity: *"The high number of genes to which reads mapped reflects the large P. camtschaticus draft genome … "*

**COMMENT: Line 336: Add "...repeat elements characteristic of crustacean genomes" to add context for readers who are unfamiliar with crustacean genomics.**
The suggested text has been added!

**COMMENT: Line 338: Replace "and ranged" with "ranging"**
This change has been made.

**COMMENT: Figure 4: Consider changing the green-purple color scheme to a single-color gradient-based color scheme to facilitate better interpretation, especially for color-blind individuals. Move A/B/C and plot titles above the volcano plot panels so they are easier to read.**
We have moved the A/B/C plot titles above the volcano plot panels as suggested. Heatmaps are challenging to visualize for color blindness! We have changed the gradient to a black/green gradient in hopes that will aid the interpretation for all readers.

**COMMENT: Line 386: Consider "darker shades" instead of "darker colors"**
The text now reads "darker shades"

**COMMENT: Lines 387-388: Are the DEG those with darker shades in the volcano plots? If so, add that information for clarity.**
The figure 4 caption has been edited for clarity:

*Figure 4. Pairwise differential expression among OA treatments for (A) ambient (pH 8.0) vs. moderate OA (pH 7.8), (B) ambient vs. severe OA (pH 7.5), and (C) moderate OA vs. severe OA. Volcano plots (left panels) show expression of all genes. Each point represents a unique gene, with non-black points residing above the dotted line representing differentially expressed genes. -Log10 p-value is along the y-axes, with higher numbers indicating higher significance, and log2 fold change is along the x-axes, with higher absolute values indicating larger differences among treatments.. Differentially expressed genes (p-adj < 0.05) are color-coded to indicate those that are expressed at higher levels in ambient treatment (blue), moderate OA (orange), or severe OA (red). Points with darker shades indicate those that have |Log2FC| >0.5,*

*which were used in functional analyses. Heatmaps (right panels) show expression of differentially expressed genes only (rows=genes) at per-sample resolution (columns=samples), with the green-black gradient indicating the z-score of expression values standardized across samples for each gene, where green and black indicate higher and lower expression, respectively.*

**COMMENT: Lines 414-415: Are these GOterm names or user-defined categories? If user-defined, how were they defined?**

The enriched biological processes are the GO term names. We have tweaked that sentence for clarity: *"Enrichment analysis revealed 26 and 104 biological processes GO terms that were enriched in genes … "*

**COMMENT: Figure 6: Consider adding a pathway diagram that summarizes important information from these figures, as these pathways are a cornerstone of the discussion section.**

Many pathways are involved in the variety of biological processes that were implicated in this study, therefore we are hesitant to include such a complex pathway figure. We have leveraged the hierarchical structure of GO terms to construct GO networks to aid the interpretation of the data; those are now included in the supplementary for reference (Supplemental figures 5-7).

**COMMENT: Lines 448-449: What is meant by a "more active" biological process? Increased expression of these genes?**

Yes, that is exactly what was meant. The text has been edited to make that clear.

**COMMENT: Lines 451-466: Was this a separate enrichment test conducted in DAVID? What was the gene background used for the enrichment?**

The Functional Analyses section in the Methods has been reformatted to more clearly outline the different gene sets examined for enriched processes. It also now specifies the background gene list used in all enrichment analyses: *"all analyzed genes that mapped to the Uniprot/Swissprot database (n=32,435)."*

**COMMENT: Line 467: Was there a statistical test associated with examining differences in TE activity/composition between the different OA treatments?**

We had not included a statistical test in the manuscript! We added the following information to the Methods and Results sections.

Methods: *"Beta-regression and likelihood ratio test assessed whether the TE proportions differed for gene sets that were up-regulated and down-regulated in OA treatments."*

Results: *"A likelihood ratio test examined the proportions of gene sets that were TEs, and found higher TE proportions in upregulated gene sets compared to downregulated gene sets ($\chi^2$ (2, N = 20) = 10.7, p = 0.001, Figure 7)."*

**COMMENT: Figure 7: Did the authors consider making a version of this figure where TE activity/composition is examined in relation to OA treatment?**

We explored a variety of figure options, one of which included a line plot showing the average expression value of each TE by treatment as suggested! The large number of TE genes and widely ranging expression values made the figure too messy for publication. As shown, Figure 7 reports the proportion of each gene module that are TEs, which we prefer as it mirrors how information is discussed in the final section of the Results.

**Discussion:**

**COMMENT: Lines 675-677: Is there another crustacean study that could be cited here instead of the Olympia oyster study?**

Definitely. The Olympia oyster studies have been replaced with a more relevant study in Tanner crab (Long, Swiney, and Foy, 2016; https://doi.org/10.1093/icesjms/fsv251).