Dear Alison,

Thank you for your decision regarding our preprint PCIEcology #687.

We are happy to see that there are only a few minor comments remaining.

You will see that we are somewhat disagreeing with some of your suggestions. We have provided below extensive and in our opinion clear justifications in each case, and still made changes to the manuscript that should address the spirit of the comments, if not following the letter.

We hope this version will be well-received nonetheless.

Best regards,

Maxime Dahirel, on behalf of the authors

COMMENT 1

"I disagree with the decision not to include the analyses including shell size. The discussion is not too long and so there is space to add this information. A difference <2 in AICc between two different models means that they are equivalent. This means that the model with the size*rural describes the data as well as the model without the interaction. Another option would be to re-run the analyses correcting for shell size as suggested by Reviewer 1 and present this in the results. Even though you found no overall effect of urbanization on snail size, there could still be an interaction with different size snails in different urban environments having differing levels of trapped nematodes."

ANSWER TO COMMENT 1

First, we would like to reiterate that we had no *a priori* documented mechanism or hypothesis justifying a potential size*urbanization interaction effect on prevalence, hence why we did not include one.

Second, the point we made in the Replies of the previous revision was not just that the models with and without the size interaction were within <2 AICc points of each other, but that **the interaction was non-significant anyway if added**. (To quote "That model had a slightly worse AICc than the equivalent one with no interaction (323.1 vs 322), the size*urbanization interaction was not significant, and in any case the effect of categorical urbanization remained significant, and quantitatively near-identical.") So the above suggestion that there could be *"an interaction with different size snails in different urban environments having differing levels of trapped nematodes"* has already been answered, by the negative.

But it was indeed only answered in the Replies; you are making the point that this should go into the study itself.

The solution we propose is a compromise between this requirement and the fact that, again, this interaction has no *a priori* mechanism/hypothesis behind it, is a post-hoc data exploration, and should thus not in our opinion be given the same importance as our main models (see also our answer to COMMENT #2).

➢ We have left the main model selection, Methods and Results as is regarding this comment, but modified the Discussion and Supplementary Materials. In the paragraph where we mention that urbanization has no effect on size and so there is no indirect causal path from urbanization to prevalence through size, we have added the following sentence:

"In addition and as a post-hoc exploration, we re-ran our model set adding size × urbanization interactions, and found no significant interaction, and no evidence that the urbanization effect changed in response (Supplementary Material S4)." (**Line 283**)

The Supplementary S4 in question is essentially paraphrasing the main text Results, but updated for the modified model set containing size*urbanization interactions.

This confirms that the outcome for the urbanization effect is, indeed, the same.

COMMENT 2

"Please can you also add a sentence to the methods stating why you don't include all the interaction terms. It would make sense to include all interactions in the maximal model and then simplify based on those which are not significant. If the maximal model does not converge then why not just include all 2-way interactions."

ANSWER TO COMMENT 2

We strongly disagree with this suggestion to add "all" interactions, as it goes against recommended practices (e.g. the recommendations in Bolker et al., 2009 to "include only important interactions" and "restrict the model a priori to a feasible level of complexity"). In detail:

First, this would involve stepwise parameter selection and/or data dredging, which is strongly recommended against by many authorities as potentially leading to spurious results (e.g. Whittingham et al., 2006, and citations therein), especially given the high number of interaction parameters to include (even if we limit ourselves to all 2-way interactions as suggested; there are 6 variables in the models with continuous urbanization; adding all 2-way interactions would add $(6*(6-1))/2 = 15$ more variables, and even more than that in the model with categorical urbanization).

Second, we disagree with the specific statement that "*It would make sense to include all interactions in the maximal model*". We had no a priori reason to suspect most or any of these interactions, and some of these interactions do not make biological sense: for instance regarding shell banding patterns, there cannot be band fusion when there are 0 bands.

Third, as we mention in the last paragraph of the Discussion, our sample size is not that high (especially keeping in mind that our individuals are not independent and are nested in populations). Interactions are harder to detect than main effects (https://statmodeling.stat.columbia.edu/2018/03/15/need16/, formally published in Gelman et al., 2020, section 16.4). Even if all these interactions were justified a priori, we likely would not have any power to detect some if added one by one, let alone all together at the same time. As an example and demonstration of this, we can see in the new Supplementary S4 that adding the size*urbanization interaction suggested in COMMENT #1 makes the main effect of size disappear in the best model, even though the interaction itself is not different from 0. A model selection predicated on removing non-significant effects from the best model, as suggested, could here have led to removal of the size effect, depending on the procedure.

Finally, even if none of the above were an issue, this would still be mixing AIC based model selection (to decide which of the 5 urban metrics to choose), and stepwise selection (to remove superfluous interactions/parameters once an urban metric is chosen). This creates a "garden of forking paths" issue where there is absolutely no guarantee that the "absolute best" model is a subset of the maximal model with the urban variable having the best AIC. As the total possible set of models given our

variables is not fully nested (because of the different urban metrics), this is not a problem that can be solved by alternative approaches such as all-subsets model averaging and variable weighing.

➢ We have added to the Methods the following sentence:

"Given our sample size, we did not include interactions between our explanatory variables, especially as we had no a priori hypotheses regarding these (but see Discussion)." (**Line 164**),

where the "(but see Discussion)" refers to the size interaction added post-hoc and mentioned in the answer to COMMENT #1

COMMENT 3

"Another point which may be interesting to explore is that Figure 3 shows that there seems to be 2 groupings among urban snails; some sites having very high nematode prevalences and the others low. Is there anything particular about the sites with high prevalence? You could re-run statistical models for only the urban group to see any of the explanatory variables are significant when only considering this group."

ANSWER TO COMMENT 3

With respect to any of the variables considered in this study, there is nothing special about these sites; they are not especially more or less built-up than the others (we considered this already during data exploration). We note that the highly urban populations with prevalence = 1 have small sample sizes, (n = 1, 3, 3 and 8 respectively), so these high prevalences could plausibly be drawn by random chance alone from a population with a much lower true probability.

➢ We added a paragraph to the Discussion acknowledging these high-prevalence points, and to suggest that future studies with more data or a different sampling focus might reveal interesting differences within cities, in addition to the differences between urban and non-urban we show here (**paragraph starting Line 269**):

*"Interestingly, nematode encapsulation prevalence was seemingly more variable between Urban Centre populations than between populations in the other urban categories (**Fig. 3**). While this may merely be due to some populations having low sample size, this suggests that there may be within-city variability in snail-nematode interactions that are not explained by built-up levels or population density. Cities are indeed highly heterogeneous environments, and within-city variation in vegetation, mediated in part by neighborhood-level socio-economic differences, may shape biodiversity, including species interactions (e.g. Martin et al., 2024). Our ability to detect within-city patterns is here hampered by our small number of sites per Urban Centre; future studies focused on this within-city heterogeneity may uncover more on the fine-scale drivers of snail responses to parasites."*

Regarding the suggestion of re-running models on a sub-dataset containing only highly urban sites, this would be equivalent to adding "urbanisation * other variables" interactions, so our response to the previous comment applies here too.

COMMENT 4

"For information, there is only the old version of the Supplementary materials on bioRxiv and the data set made available does not include the urban metrics"

ANSWER TO COMMENT 4

This is not true, the files were correct and complete in the previous revision, and remain so in this new version.

The Supplementary Materials in bioRxiv for the previous revision (v2 on bioRxiv) were the correct revised ones (as evidenced by the mentions of conditional and marginal R² in S2 and the presence of the S3, which were added specifically in response to reviewers _ https://www.biorxiv.org/content/10.1101/2024.03.07.583959v2.supplementary-material).

The data set made available does actually include the urban metrics, and already did so in the original version (v1.0 on GitHub); they are present not in an isolated csv, but within the geopackage containing the sites coordinates (data/HELICITY22_sites_BELambert72_wGHSL.gpkg, called line 33 of the analysis code analysis.qmd [https://github.com/mdahirel/HELICITY-2022_shell-nematodes/blob/v1.0/R/analysis.qmd], and documented starting line 55 of the same file)

The above remains true of the files for the new revised versions (v3 on bioRxiv and v1.1 on Github) (see https://www.biorxiv.org/content/10.1101/2024.03.07.583959v3.supplementary-material and https://github.com/mdahirel/HELICITY-2022_shell-nematodes/blob/v1.1/R/analysis.qmd).

REFERENCES

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White J-SS (2009) Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution*, **24**, 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Gelman A, Hill J, Vehtari A (2020) *Regression and other stories*. Cambridge University Press, Cambridge. https://doi.org/10.1017/9781139161879

Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP (2006) Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, **75**, 1182–1189. https://doi.org/10.1111/j.1365-2656.2006.01141.x