

Dear Dr Anna Cohuet,

Thanks for your decision on our manuscript formerly now entitled “Pre- and post-oviposition behavioural strategies to protect eggs against extreme winter cold in an insect with maternal care”.

We are pleased to see that you and the two referees are enthusiastic about our work. We have extensively revised the manuscript to follow all suggestions and respond to all comments. Among the most important changes, we have (1) edited the abstract, introduction, results and discussion to clarify that the main goal of our study was to test the occurrence of the two behavioural strategies in the European earwig and not to shed light on the drivers of inter-population variation in the expression of these behaviours, (2) rewrote the methods to better explain the goal of the successive experimental steps, add information about the biology of the European earwig and clarify why we used two devices with different temperature gradients, (3) added and modified figures to improve the readability of our results, and (4) added another analysis to improve the robustness of our conclusion (this new result does not change our conclusions). Finally, we have carefully answered the numerous questions and followed all other suggestions that were raised throughout the manuscript.

Overall, we would like to warmly thank you and the two reviewers for taking the time to review our manuscript and for providing such detailed, important and insightful contributions. We are confident that our manuscript is now of better quality and can reach a wider audience.

You can find a point-by-point reply to the comments of each reviewer below and a version of our manuscript with changes highlighted in yellow attached to this submission. We have also added a conflict of interest section in the manuscript, as well as added the corrected R script, data files and new readme file in the Zenodo repository (<https://doi.org/10.5281/zenodo.6338078>).

Sincerely,

Joël Meunier on the behalf of all co-authors

Reviews

Reviewed by Wolf U. Blanckenhorn, 22 Dec 2021 13:35

Review of PCI Zoology MS#120

This MS reports a single-species (??) investigation extending the study of thermal preferences of animals to the tending of laid eggs by earwig females during winter. Two eastern Canadian populations of the European earwig – is the invasiveness of this species an issue worth mentioning here?? – with slightly differing thermal regimes (Fig. 1) are being compared, an absolutely minimal comparison in terms of geographic variation that constitutes not much more than a simple population replicate. (In addition, there seems to be a possibility of cryptic species, which complicates the comparison even further and better should not be discussed here, else there is no population replication whatsoever.)

In total, the study reports interesting and apparently novel data (which does not become completely clear from the Introduction) on the egg-tending behaviour of earwig mothers that are worth publishing in a zoological, entomological, behavioral, or thermal

journal (such as J. thermal Biology). The MS is overall well written, though I have added a few (text) edits and comments here and there (using Acrobat).

Thank you for the very detailed review, the many text changes and the overall positive evaluation of our work. We have edited the abstract and introduction to clarify the novelty of our study (see details below) and corrected the entire manuscript following the changes suggested in the pdf attached to the review. The comments raised in the pdf are reported below, along with our response.

Comments in the pdf

L121 - Give species name and group here, or above

We have edited the text accordingly: “The experiment involved a total of 60 *F. auricularia* females field sampled in Harvey station (67°00'52.2"W, 45°38'23.6"N; New Brunswick, Canada; HNB) and St John's (47°34'42.6"N, 52°44'37.0"W; Newfoundland and Labrador, Canada; SJNL) in September 2020 and then maintained in plastic containers until October 2020. These two populations were selected as independent random units on the basis that they were 2500 km apart by land (1000 km by sea) and had comparable climatic conditions (Fig 2). Subsequent genetic analyses revealed that HNB individuals belonged to the *F. auricularia* genetic clade 'A' and SJNL females to the *F. auricularia* genetic clade 'B' (see details in the supplementary material and discussion; Wirth et al. 1998, González-Miguéns et al. 2020).” (L130-138)

L131 - Supply photograph of thermal set-up as first figure of paper. The temperature relationships depicted in Figures Sx can remain in the supplement.

We have followed this suggestion and added a schema of the thermal setups as Figure 1.

L165 - What was the purpose of this treatment?

We have edited the methods to clarify the purpose of different thermal setups at the beginning of each paragraph: “To determine the temperature range explored by each female before egg production and the temperature chosen to deposit eggs, we transferred the 36 aluminium rails containing each female on thermoelectric plates (AHP-1200CPV- Thermoelectric Cooling America Corporation. 4049W Schubert Avenue. Chicago IL. USA) with temperatures linearly ranging from 0°C to 20°C (Fig 1A and S1).” (L152-156); “We then tested whether mothers moved their eggs depending on experimental changes in environmental temperature and/or egg age during the 15 weeks following oviposition. Three days after each female has laid its eggs, we transferred the mother and all its eggs to a new (shorter) aluminium rail (1.8 x 1.8 x 66 cm), which was deposited into a climate cabinet (Fig 1B) providing three non-linear ranges of temperature: warm (0.9°C to 7.0°C), intermediate (-3.6°C to 2.7°C) or cold (-4.5°C to 1.6°C) (Fig 1B and S1).” (L171-176); and “We finally tested whether the temperatures at which eggs were maintained during the 15 weeks following oviposition affect egg development and hatching rate. Fifteen weeks after oviposition, we transferred the 60 (short) rails containing the mothers and their eggs (i.e. 5 rails per thermal range and population) to thermoelectric plates with temperatures linearly ranging from 0°C to 19°C (and not 20°C due to the shorter rails).” (L193-197).

L176 - Confusing sentence.

Sorry for that. We have simplified the sentence: “The recordings occurred every hour during the entire experiment on either four aluminium rails evenly distributed over the thermoelectric plates or six aluminium rails evenly distributed among the trails and thermal constraints in the climate cabinet. There were no edge effects on temperatures.” (L207-211).

L182 - We believe you. This can go in the supplement.

We have followed this suggestion and moved the genetic part to the supplementary file, together with the very brief results of our genetic analysis.

L216 - This statistical treatment sound overly complicated. Is this because you had (too) low sample sizes?

Sorry for the confusion. This part explains how we managed to interpret the triple-interaction in our statistical models. We have edited the sentence to make it simpler: “To interpret the significant interaction between the fixed explanatory factors, we divided the data set according to temperature range (i.e. in three subsets), in which we ran another series of LMEs in which we used week, population and the interaction as fixed explanatory factors, and the female ID as a random factor. When the interaction between weeks and population were significant, we conducted post hoc pairwise comparisons between the temperature of the location where the eggs were initially transferred (i.e. 5.2, 1.18 and 0.11°C) and the temperature of the location where the eggs were observed each week using a series of one sample Exact Mann-Whitney Rank Sum tests.” (L224-232)

L231 - Give temperatures here.

We have added this information (L207-209): “The warmest temperature at which females were observed was higher in SJNL than HNB females (median = 19.6 and 12.7°C, respectively; Fig 3B; $W = 2$, $P < 0.001$), whereas the coldest location was the same for SJNL and HNB females (median = 5.7 and 4.0°C, respectively; Fig 3C; $W = 135.5$, $P = 0.404$).” (L247-251)

L258 - delete or move to supplement (as above)

As suggested, we have moved it to the supplement.

L558 - Unclear what is meant by 'thermal constraints' here.

We have changed “thermal constraint” with “temperature range” (L673 & L693)

Specific comments:

1) I would de-emphasize the (presumably) latitudinal comparison of merely two populations (or even cryptic species) in lieu of a mere population replicate. Else the study leaves much to be desired.

We agree with this suggestion. While the comparison between populations was a detail of our experiment, the way we wrote our article has inadvertently made it appear as one of its main objectives. Sorry for that and the associated misunderstandings. We have modified the entire manuscript (abstract, introduction, methods, results and discussion) to clarify that we used two random populations because life-history traits often vary between populations in this

species, that the main objective of our study was to highlight the existence of pre- and post-oviposition strategies to limit exposure to extremely cold temperatures and to present the inter-population difference as a secondary result that requires future studies to better understand its origin.

2) A figure of the thermal apparatus should be part of the MS (as Fig. 1).

We have followed this suggestion and added a schema of the thermal setups as **Figure 1**.

3) Why are EU earwigs studies in Canada?

The European earwig can be found worldwide. It exhibits a broad native range extending across Europe, Asia, and northern Africa from which it expanded to Australia, New Zealand, East Africa, East Indies, Tasmania, and North America. We have added this information in the text: "In this complex of cryptic species (Wirth et al. 1998; González-Miguéns et al. 2020) that can be found on almost all continents (Lamb and Wellington 1975; Guillet, Josselin, et al. 2000; Quarrell et al. 2018; Hill et al. 2019), females usually ..." (**L83-86**)

4) There are some unclarities left regarding the Methods, especially regarding the transfers (see my comments in the MS).

Sorry for that. We have rewritten the entire method section to make it overall clearer, particularly regarding the purpose of the transfers between thermal setups.

5) The genetic analyses in Methods and Results really don't fit here and are not necessary (cf. comment 1 above). Remove.

We have followed this suggestion and moved the genetic analyses to the supplementary file.

6) The statistics used are also not super-clear. In particular, I can only hope that the most fitting error-distributions were used in the models. Or were all simple non-parametric analyses after all?

We have edited the statistical part to clarify that we used a combination of non-parametric and parametric analyses and to emphasize that we checked that the most fitting error-distributions were used in the models (**L215-244**).

7) Figures 2 & 3 could be combined (underneath), as the panels are identical.

We have followed this suggestion and combined figures 2 and 3 in a new figure 3.

Reviewed by Nicolas Sauvion, 27 Jan 2022 15:37

I was very interested in this study on earwigs. It's a biological model that I didn't know about, so I would have liked the authors to give a little more detail on the biology of these insects. These details are sometimes lacking to fully understand the observations made. I will clarify these points later. Overall, I find this article clear, very well written, well structured (although I also make some suggestions below). The problematic is well posed, but it seems to me that the authors do not emphasise the biological question enough (what

kind of strategies earwigs adopt to minimise the risks to their eggs during winter). Yet, it would be more interesting to present the study from this angle (as it is very original!!) rather than 'simply' presenting another study on the effect of temperature on the biology of an insect. This is only my point of view!

Thank you very much for the positive assessment and very careful revision of our manuscript. We are convinced that the detailed suggestions have greatly improved the quality of our text. Thanks a lot! Moreover, we agree with this suggestion and have changed the title accordingly, as well as edited the introduction, results and discussion to emphasize the biological question about the kind of strategies earwigs adopt to minimise the risks to their eggs during winter.

Overall, I have no major substantive comments to make: I note that the authors are very familiar with their biological model and the concepts of evolutionary biology on which they argue, especially in the discussion. On the other hand, I have many detailed remarks to make. In particular, I would like to insist on the formalism to be respected for the names of the species, even if I know that the editors are pushing for concision. I will be happy to proofread a revised version of this manuscript

We have added the names of species in the requested format (see below)!

For the first sentence line 32 add in reference :

Meunier J, Körner M and Kramer J (2022) Chapter 17: Parental care. In Reproductive strategies in insects (ed. Omkar & Geetanjali Mishra). CRC Press, Taylor and Francis. pp 337-359 [pdf] - Book chapter <https://doi.org/10.1201/9781003043195>

We have added this reference.

In all rigor and in accordance with the international code of zoological nomenclature (<https://code.iczn.org/> see Article 51. Citation of names of authors ; Article 22. Citation of date) : 'It is strongly recommended that the date of publication (and the authorship; see Article 50) of a name be cited at least once in a work which deals with a taxon. This is particularly important for homonyms and for species-group names not in their original combinations'. I have thus searched for the 'correct' names of the species mentioned in the manuscript. I list them below, with the sources.

• You will notice that the name of the Gerris species mentioned is ambiguous. I am not a specialist of this group. I found references on this question but without being really affirmative on the correct name of species that should be retained [*Gerris paludum insularis* Miyamoto, 1859 ?]

• You will notice that *Hyla versicolor* is a synonym of *Dryophytes versicolor* (LeConte, 1825)

in this case I suggest to write : *Hyla versicolor* (LeConte, 1825), formally *Dryophytes versicolor* (LeConte, 1825)

because *Hyla versicolor* takes the name mentioned in the reference cited by the authors (here Takahashi, 2007)

• the same for the flat rock spider : *Hemicloea major* L. Koch, 1875 synonym of (*Walckenaer*, 1837)

Aquarius paludum insularis ?= *Aquarius paludum insularis* (Motschulsky)?

=> ? *Gerris (macrogerris) insularis* (Motschulsky, 1866) [cf. Damgaard & Cognato 2005 Syst. Entomol. : <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-3113.2005.00302.x>]

=> *Gerris paludum insularis* : Miyamoto (1859 : 118 ; non Motschulsky, 1866).

<https://www.gbif.org/fr/species/2020415/treatments>

see also : <https://research.amnh.org/pbi/library/4039.pdf>

[Anthocharis cardamines \(Linnaeus, 1758\)](https://inpn.mnhn.fr/espece/cd_nom/54451/tab/sources)
https://inpn.mnhn.fr/espece/cd_nom/54451/tab/sources
 Orange-tip (Anglais)
 Culiseta longiareolata (Macquart, 1838) [Culicoidea]
https://inpn.mnhn.fr/espece/cd_nom/225143
 Anopheles punctipennis (Say, 1823)
<https://wrbu.si.edu/vectorspecies/mosquitoes/punctipennis>
 Lechriodus fletcher (Boulenger, 1890)
https://fr.wikipedia.org/wiki/Lechriodus_fletcheri
 Hydroporus incognitus Sharp, 1869
https://inpn.mnhn.fr/espece/cd_nom/9490
 Hydroporus nigrita (Fabricius, 1792)
https://inpn.mnhn.fr/espece/cd_nom/223388
 Lestes macrostigma (Eversmann, 1836)
https://inpn.mnhn.fr/espece/cd_nom/65205
 Hyla versicolor = Hyla versicolor LeConte, 1825 => Dryophytes versicolor (LeConte, 1825)
<https://amphibiansoftheworld.amnh.org/Amphibia/Anura/Hylidae/Hylinae/Dryophytes/Dryophytes-versicolor>
 Phrynocephalus przewalskii Strauch, 1876
 Przewalski's toadhead agama
 The agamid genus Phrynocephalus, known as toad-headed agama
<https://reptile-database.reptarium.cz/species?genus=Phrynocephalus&species=przewalskii>
https://en.wikipedia.org/wiki/Phrynocephalus_przewalskii
 Osmia bicornis (Linnaeus, 1758)
https://inpn.mnhn.fr/espece/cd_nom/816756
 Morebilus plagusius (Walckenaer, 1837)
 Syn. Hemicloea major L. Koch, 1875
<https://wsc.nmbe.ch/species/42887>
 Phyllomorpha laciniata (Villers, 1789)
https://inpn.mnhn.fr/espece/cd_nom/829029
 Abedus (Deinostoma) herberti Hidalgo, 1935
<https://www.gbif.org/species/2007558>
 Forficula auricularia Linnaeus, 1758
https://inpn.mnhn.fr/espece/cd_nom/65991

Thanks a lot for providing us with this very detailed list on names. We greatly value the effort and have added them to the text (L35-128). Note that some examples have been removed to follow the suggestion of reviewer 3.

Line 34. In their introduction the authors mention benefits to (1) egg-laying females, (2) their current eggs, and (3) their future juveniles. This chronological order is biologically logical. Next, the authors expand on point 1 (lines 37-41), point 3 (lines 41-48) and point 2 (lines 48-57). I suggest restructuring the paragraph and detailing point 2 before point 3.

This is an excellent point. We have restructured the paragraph accordingly, i.e. by detailing the benefits to egg-laying females, then to their current eggs and finally to their future juveniles (L32-55)

Line 103. 'shorter exposure to cold speeds up egg hatching' taken out of context this sentence is counter-intuitive. Indeed, one can wonder if it is really the duration of exposure

to cold that affects the speed of hatching, or rather the intensity of the cold. Intuitively, we imagine that the regions with the longest winters are also those where the temperatures are the coldest. Reading the article by Körner et al. 2018, I understand that the observations were made under experimental conditions, and that the minimum temperatures were identical between the two conditions tested (long winter versus short winter). so I do agree with the conclusions and statement of the authors. However, I suggest that the sentence be qualified as follows: 'under experimental conditions, by exposing insects for varying lengths of time to the same minimum temperatures (in this case 5°C), it is observed that shorter exposure...'

We have edited the sentence to clarify that the experiment was conducted under standard conditions, that the minimum temperatures were identical between the two tested conditions and to detail the type of effects (as suggested by another reviewer): "Second, a recent study shows that prolonging egg exposure to cold (5°C) for 15-day during winter delays the hatching date and development of juveniles to adulthood (which typically takes two months in *F. auricularia*), leads to the production of lighter adult females, and shape the basal immunity of these females: it increases their overall phenoloxidase activity and reduces the number of haemocytes in females facing a changing social environment (Körner et al. 2018)." (L99-104)

Line 122 : Explain why these locations were chosen in relation to the biological question; why the authors would expect females from these two populations to behave differently. Quote figure 1 here to explain this choice

These populations were simply used are mere (random) replicates as previous studies often reported inter-population variation in the European earwig. To make it clearer, we have edited the abstract and introduction: "We set up 60 females from two random natural populations (as this species often exhibits population-specific life-history traits and behaviours) under controlled thermal gradients, and then recorded the temperature at which they built their nests, tested whether they moved their eggs after an experimental temperature change, and measured the effects on egg development and hatching rate." (L18-23) and "Because previous studies point out that the European earwig may show population-specific life-history traits and behaviours (Ratz et al. 2016; Tourneur 2017; Tourneur 2018; Tourneur and Meunier 2020), we also tested whether temperature-dependent oviposition site selection and egg transport vary between two (randomly selected) populations sharing comparable climatic conditions." (L114-118).

We had not specifically anticipated any differences between these two populations and were also very surprised to detect species B in one of the two populations. We have added this information in the text: "... our genetic analyses reveal that individuals from the two tested populations belong to different genetic clades: females of Harvey station belong to species 'A' and females of St John's to species 'B' (this was surprising, as this is the first time that species 'B' is found in Canada outside British Columbia)." (L364-366).

Specify gps coordinates : Harvey station Lat. 45° 44.556'N ; Lon. 67° 0.944'W. St John's Lat. 47° 33.691'N, Lon. 52° 42.755'W

We have added the precise GPS coordinates of the two sites: "The experiment involved a total of 60 *F. auricularia* females field sampled in Harvey station

(67°00'52.2"W, 45°38'23.6"N; New Brunswick, Canada; HNB) and St John's (47°34'42.6"N, 52°44'37.0"W; Newfoundland and Labrador, Canada; SJNL) in September 2020 and then maintained in plastic containers until October 2020." (L130-133)

Line 131 : 'insulated with thick foam to ensure complete darkness.' For a naive reader who does not know this biological model, it would be useful to specify here how this device imitates the conditions encountered in natural conditions by the earwig

Sorry for the lack of information about earwig biology. We have rewritten the method section to provide all details about its biology and how our laboratory mimic the conditions encountered in natural conditions by this species (L130-213)

Line 126. Please add that the females were transferred 'individually'. This is important, because otherwise (e.g. batch tests of 6 individuals), one could suspect interaction effects between females, and this should be taken into account in the choice of statistical tests (tests on paired data)

We have edited the text to make clear that females were alone in each rail: "In October 2020, we isolated each female as this is the period when they usually leave their group to search for a future nesting site (Lamb 1976). To this end, we transferred 36 of 60 females (18 from HNB and 18 from SJNL) to the middle of 36 aluminium rails (Fig 1C; 1.8 x 1.8 x 72 cm = height x width x length) covered with a ..." (L145-147)

I assume that you did not use 36 thermal bridges, but N1 bridges on which you placed N2 rails. Are the temperatures perfectly identical laterally on the bridges? Otherwise, there is no risk of edge effect? I think it is necessary to precise this point

We used two bridges and laid 18 aluminium rails on each. The temperature has been measured on four rails per bridge during the entire experiment and there was no edge effect. We have added this information in the method section: "The recordings occurred every hour during the entire experiment on either four aluminium rails evenly distributed over the thermoelectric plates or six aluminium rails evenly distributed among the trails and thermal constraints in the climate cabinet. There were no edge effects on temperatures.." (L207-211)

Line 134 : 'To limit stress on the females due to rail handling, we divided each rail into 12 zones of 60 mm length and defined the distance between a female and the coldest edge as the centre of the zone she was in.' I understand the trick but I have some questions:

- **Have you been confronted with the situation of a female straddling two zones? If so, how did you then estimate the distance?**

This is an excellent point. Thanks for spotting it. We have added this information in the text: "In the very few cases where females were observed between two zones, we arbitrarily assigned females' location to the colder of the two zones." (L163-165)

- **I don't know the behaviour of this insect: was it very mobile? or rather placid? in other words, was it very sensitive to disturbances and therefore very reactive at the time of the counts? and therefore could this behaviour have biased the distance estimates?**

The females were not very sensitive to disturbance and it was therefore very easy to reliably record their location at the time of observation. We have changed the

wording to limit misunderstandings “To facilitate distance measurement, we divided each rail into 12 zones of 6 cm length and defined the distance between a female and the coldest edge as the centre of the zone she was in.” (L161-163) and clarified that “All distance recordings were made very gently so that the females were not disturbed during the observation.” (L202-203)

Lines 132-134. The expressions 'before egg-laying' and 'until they laid eggs' lead me to wonder about one point, because - once again - I am not familiar with this biological model: the choice was made for this study to work on wild populations and not on reared populations. Is egg production relatively well synchronised in time in natural populations, or is there a large variability in individual behaviour, which could bias the results of the study? Females could have responded less well to the temperature stimulus simply because they were not yet in an active period of searching for an egg-laying site. For the biological models of hemipterans that I am familiar with, we know that the periods of sexual reproduction/laying are highly synchronised in time and that the photoperiod plays an essential role in this synchronisation.

We have realised - once again - that our manuscript lacks crucial information on the biology of the European earwig and we apologise for the confusion this has caused. There is both inter- and intra-population variability in the time of egg production in natural populations of the European earwig (see e.g. Tourneur J-C, Meunier J. 2020. Variations in seasonal (not mean) temperatures drive rapid adaptations to novel environments at a continent scale. Ecology. 101(4):e02973. <https://onlinelibrary.wiley.com/doi/abs/10.1002/ecy.2973>). This is also supported by our study, which confirms that HNB females produced their eggs earlier in the season than SJNL females (L254-256; Figure 3F). This variability cannot bias the results of our study because our experimental design standardized our measurement around the natural period of egg production of each female. We have added this information in the method section: “To determine the temperature range explored by each female before egg production and the temperature chosen to deposit eggs, we transferred the 36 aluminium rails containing each female on thermoelectric plates [...]. We then measured the distance between each female and the coldest edge of its rail every day from the time they were placed in the rails until they produced eggs (including the day of oviposition). Because the date of oviposition greatly varies within populations (up to several months between the first and last oviposition; e.g. Tourneur and Meunier 2020), we standardised our measurement to the temperature range explored by each female 15 days before its own oviposition.” (L158-160)

I also have several questions:

- did all the females lay eggs (I understand that they did)?

Yes. All the females we deposited in the rails produced eggs.

- Did they lay eggs almost simultaneously or can the behaviour of finding the laying site be (very) variable? (range of duration? min/max?)

See our detailed answer above. HNB females produced their eggs earlier in the season than SJNL females (Figure 3; L254-256), but the measured behaviours (Temperature range, min and max) were all standardised to the 15 days before laying - irrespective of the date of oviposition.

- how do we know that the females have laid eggs? is the laying site very well identifiable? After reading the rest of the manuscript, I realise that you did not specify that the observations of the females were made during the 15 days prior to egg laying. It is here in this sentence (line 132-134) that it should be specified

We have added a photo to figure 1C to show that the eggs are large and easily observable. We have also added information about the 15 days of measurement (see answer above).

Line 145. ‘Three days after oviposition’, I don't understand what exactly the authors mean. : all the females were left for X days (i.e. same durations for all females) to lay eggs, and three days later they were transferred ? or, as soon as oviposition was observed (i.e. variable duration depending on the females), you waited for 3 days and then transferred them individually ?

We have edited the text to clarify this point: « Three days after each female has laid its eggs, we transferred the mother and all its eggs to a new (shorter) aluminium rail (1.8 x 1.8 x 66 cm), which was deposited into a climate cabinet (Fig 1B) providing ...” (L173-175)

Lines 150-152. You should specify what type of cabinet you used. I found it difficult to understand that you used a cabinet that allows to have such an accurate temperature gradient. I did not know that this type of climatic cabinet existed. I am rather familiar with climatic cabinet in which the temperature can be varied by programming different cycles of variable duration and temperature. The temperatures are surprisingly accurate (100th of a degree, really?). Moreover, it seems to me astonishing that no imprecision is given on these values. From a metrological point of view, there is at least some imprecision on the device that produces/maintains the cold, and on the device that measures the temperature.

Sorry for the confusion and lack of details. We used a homemade climate cabinet made from a freezer to which we connected aluminium rails providing a thermal gradient from the refrigerating part of the freezer to the ambient temperature (this is why we do not have a linear relationship between distance and temperature - unlike thermal bridges). We did not control the temperatures but recorded them to the nearest 0.1°C throughout the experiment using external thermometers directly connected to the rails. The reported temperatures, therefore, reflect the retrospective measurements recorded by the thermometers.

We have added a figure to provide a schema of the climate cabinet (Figure 1B) for which the legend contains details about the apparatus and edited the text to clarify that the temperatures were recorded to the nearest 0.1°C (L205) and to provide standard deviation of the measurements in the two apparatus (Figure S1).

There is no legend for the axes of the graphs in figure S2. I assume that the x-axis refers to the length of the rail.

Sorry for this mistake. We have edited Figure S2 (which is now combined with Figure S1) to add a legend for the axes.

Choose to express lengths in cm or mm but use the same unit throughout (currently mm in the text and cm in figures S1 and S2).

We have standardized the length unit in cm in text and figures.

Why not choose whole numbers such as 1°C to 7°C form warm range, etc ? as for Figure S1.

We have standardized the temperature to one decimal after the comma in the text and figures to clarify the level of accuracy of our measurements.

I suggest to put the 3 curves together in one graph. This would make it easier to visualize the temperature differences tested between the 3 conditions.

This is an excellent suggestion. We have combined the panel of Figure S1 and the 3 panels of Figure S2 into a single panel in a new **Figure S1**. This makes it easier to visualize the temperature differences tested between the 2 thermal setups (bridge and cabinet) and between the 3 conditions in the cabinet.

Conditions B and C do not seem to be very different (less than A and B anyway). Why not have chosen conditions C a little colder, i.e. sharper compared to condition B (technical limits?)

This is correct: sharper contrasts between A, B and C would have been nicer, but it was not possible due to the technical limits of our homemade climate cabinet. As detailed above, we did not fixed temperatures but recorded them throughout the experiment using external thermometers directly connected to the rails. Nevertheless, we believe that such a limited difference between conditions B and C do not reduce the robustness of our conclusions.

Line 153-155. ‘These temperature ranges... during the natural period of egg care’. For readers who are naive about this biological model, specify here the known spawning period for this species. I assume 'autumn' which would explain the choice of populations and the consistency of the results. One locality has a harsher winter (HNB) than the other (SJNL), with a population that seems to have adapted to this climate.

We have edited the text to explain that: “These three thermal gradients encompass the above-ground natural range of temperatures of the two populations during the time females were maintained in our laboratory, i.e. during the natural period of egg care (Gingras and Tourneur 2001; Fig 2).” (**L176-179**).

Line 155-160. I suggest : “To test whether and how mothers transported their clutch throughout egg development, we then measured the distance (in cm) between the centre of the pile of eggs and the coldest edge of the rail once a week during the 15 following weeks. Because rail handling occurred only weekly in this part of the experiment, we measured the distances between the (center of the pile of) eggs and cold edge directly in cm.” Again, no risk of disturbing the females too much during the rail handling?

We have changed the sentence accordingly (**L171-183**). As explained above, the handling was done carefully and there were only very limited risks of disturbing females during the process (**L202-203**)

Lines 163-165. I think this sentence is incorrect. For each population, the authors tested 10 females per modality (warm, intermediate, cold ; lines 151-152). I understand that they took each of these 10 females and transferred them to the thermal bridges used before oviposition (0°C => 20°C). However, the authors probably transferred a random subset of eggs (how many ? variable/fixed number).

We transferred each mother with all its eggs (the use of random was indeed a mistake, sorry for that). We have edited the sentence: “Fifteen weeks after oviposition, we transferred the 60 (short) rails containing the mothers and their eggs (i.e. 5 rails per thermal range and population) to thermoelectric plates with temperatures linearly ranging from 0°C to 19°C (and not 20°C due to the shorter rails).” (L194-197)

Line 166. ‘to record the date of egg hatching’. I guess there were several eggs transferred/observed, and that they did not hatch at the same time, so it would be more correct to write ‘the hatching dates of the eggs’

We have edited the sentence for clarity: “We then checked each female daily to record the date of the first egg hatching, the location of the clutch at hatching (based on the distance between the centre of the clutch and the coldest edge of the rail) and the number of newly hatched juveniles.” (L199-202). Note that we checked the date of the first eggs that hatched for each female, as all the eggs of a clutch typically hatch within one day in this species (with a few exceptions that could hatch up to 3 days later).

Line 173-174. ‘1.5, 22.5, 43.5 173 and 64.5 cm’. these figures appear to be consistent with Figure S2. ‘2, 25.5, 49 and 72 cm’ these figures do not appear to be consistent with Figure S1. the first point is 0, the second point is below 25, the third is close to 45 and the third one seems correct.

Thanks for having spotted the mistake. We have corrected figure 1, which now shows points correctly placed at 2, 25.5, 49 and 72 cm.

Line 175. ‘every hours’ ‘four aluminium rails’

The sentences have been changed.

Line 177. ‘six aluminium rails’. I suppose that the temperatures measured were a little different from one rail to another, from one hour to the next. This remark is in line with the one I made above (lines 150-152) on the imprecision of the measurements. Thus, I understand that the point on the graphs in Figures S1 and S3 are average values and that the ‘precision’ to the 100th is just a choice of the authors to give a value with two decimal places.

This is correct. Nevertheless, we agree that figures S1 and S2 needed to show error bars to clarify that temperature variation occurred in the course of the experiment and to show the number of decimals reflecting the level of precision of the thermometer. We have done these two corrections (error bars and 1 decimal after the comma) in the new Figure S1.

Genetic analyses. This paragraph seems perfect

Thanks. This part has been moved to the supplementary material following the suggestion of reviewer 1.

Statistical analyses

Lines 204-209. Exact Mann-Whitney Rank Sum tests is a non-parametric test. It is a modification of the exact test of Wilcoxon Rank-Sum Exact Test to provide an exact test for (classical) Mann-Whitney test. It would be more correct to call this test ‘Exact Wilcoxon-Mann-Whitney Test’.

We agree with this comment but would like to keep the name “Exact Mann-Whitney Rank Sum tests” in the text. This is because it emphasizes that the data are not paired (which could be otherwise confusing if using Wilcoxon) and it is already commonly used in the literature.

A condition for its application is the independence between the samples (here two populations) to be compared. As the protocol is described, the (individual) data are all independent: each female was tested separately. The test is therefore applicable. In fact, I do not understand the precision in line 204 'correcting for tied observations'. There is no need to correct for the observations (they are not tied) and intrinsically this type of test is not intended to correct for this effect if it were the case. The ‘exact test’ simply calculates an ‘exact p-value’. Thus, I suggest the sentence : ‘We used Exact Wilcoxon-Mann-Whitney Test to test the effect of population on...’

Our observations can be tied, as females from the two populations can be observed at the same temperature (even if they are maintained in different rails), which is an issue for computing exact p-values with non-parametric tests. That is why we needed to specify that the ties have been considered in our analyses.

To be consistent with the logic of the chronological order of presentation of the figures, I suggest restructuring the rest of the sentence: ‘...to test the effect of population on the amplitude of temperatures at which females were observed before oviposition, the warmest and coldest temperatures reached by females before oviposition, location of females at oviposition, the date of oviposition, the number of eggs produced, the number of weeks until egg hatching and the temperature of the area of egg hatching.

This is an excellent suggestion. We have changed the text accordingly (L216-220).

Line 218-219. I think I understand the idea of the multiple comparison test: the principle is to compare the value of one week to the initial temperature, knowing that the position of the nest at that moment is dependent on the position of the nests of the previous weeks. Right ? the test really allows to answer this comparison objective ? To my knowledge, multiple comparison tests take into account all possible pairwise comparisons.

Yes, this is exactly the goal of our analysis and we believe that it is an efficient and robust approach for testing the significance of this pattern. In particular, we conducted a series of pairwise comparisons (initial position versus position of each week) and because this ended up in using the values of the initial positions multiple times, we corrected the p-values of each pairwise test using the FDR method.

Having said that, we realized that there was a mistake in the name of the test we reported for these pairwise comparisons in the 2nd part of the experiment. The initial location of each female was a constant value for each condition (the temperature at which we set them up) and our pairwise tests thus compared this value to the temperatures of the location of the eggs observed each week. We have edited the text to correct the name of the test and clarify the statistical approach: “When the interaction between weeks and population were significant, we conducted post hoc pairwise comparisons between the temperature of the location where the eggs were initially transferred (i.e. 5.2, 1.18 and 0.11°C) and the temperature of the location where the eggs were observed each week using a series of one sample Exact Mann-Whitney Rank Sum tests.” (L228-232).

Line 223-226. I suggest putting the following sentence at the beginning of the 'Statistical analysis' paragraph: 'All the analyses were conducted with the software R v4.1.1 (R Core Team, 2017).' And then specify the package used for each type of analysis. It might even be useful to specify the function used e.g. package car, Boxplot function?

We have provided a detailed list of the packages (with references) we used to conduct our analyses, and we feel that the text would become overloaded and difficult to read if we would also detail the function of each package. We therefore would like to keep the text as such. Nevertheless, we believe that readers interested in this information could get it easily, as our R script and data set are available on a public repository (as mentioned L617-618).

Complete the reference for car (<https://rdrr.io/cran/car/>) J. Fox and S. Weisberg. An R Companion to Applied Regression. Sage, Thousand Oaks CA, 3rd edition, 2019. URL <http://z.umn.edu/carbook>. <https://rdrr.io/cran/car/>

Done

e.g. package exactRankTests, wilcox.exact function?

<https://www.rdocumentation.org/packages/exactRankTests/versions/0.8-34/topics/wilcox.exact>.

I understand that the authors have used the function wilcox.exact, and probably not the argument 'paired' (contrary to what they write line 204), if not we would see the letter V and not the letter W line 228 for example (Fig. 2A; W = 53,...)'

This is correct, we did not use (nor write that we used) paired test.

Line 228. Not sure if it is useful to specify the value of the test statistic (e.g. 53) with a letter describing it (here W for classical Wilcoxon test).

We are not sure to understand this comment. We believe that it is important to provide the letter describing the statistical test, and this is what we did in the manuscript.

Replace 'P' by 'p-value' to be more explicit

As far as we know, "P" and "p-values" are commonly used to report p-values in the scientific literature (it mostly depends on the editorial policy of the journal). We thus gently ask to keep "P" in our manuscript.

Line 227-231. It would be interesting to give the observed temperature differences (differences between median values) (respectively approx. 5°C and 8°C / Fig 2A and 2B)

We have edited the entire paragraph to provide the median values of the different treatments (as the difference in median values is not accurate): "The maximum temperature range was greater for the more arctic SJNL than the HNB females (median = 12.2 and 7.8°C, respectively; Fig 3A; W = 53, P < 0.001). The warmest temperature at which females were observed was higher in SJNL than HNB females (median = 19.6 and 12.7°C, respectively; Fig 3B; W = 2, P < 0.001), whereas the coldest location was the same for SJNL and HNB females (median = 5.7 and 4.0°C, respectively; Fig 3C; W = 135.5, P = 0.404)." (L246-251)

Line 232. Higher? again, give the temperature differences between the two populations: approx. 7°C ? Line 234. Earlier ? approx. 50 days in average ? More eggs ? approx. 23 eggs in average ?

We have edited the paragraph to provide these values: "The temperature of this site was higher in SJNL compared to HNB females (median = 16.1 and 9.2°C, respectively; Fig 3D; $W = 17.5$, $P < 0.001$). Moreover, HNB females produced their eggs earlier in the season (median = 12 and 60 days after first egg production in all females tested, respectively; Fig 3E; $W = 38$, $P < 0.001$) and laid overall more eggs (median = 68 and 46, respectively; Fig 3F; $W = 261$, $P < 0.001$) than SJNL females." (L253-257)

Lines 236-247. On each graph in Figure 4, it would be useful to show a horizontal line at the maximum rail temperature (6.96°C, 2.73°C, 1.61°C). This would allow a better visualization of the females that have reached the end of the rail.

Thanks for the suggestion. We have added a red horizontal line in Figure 4 to show the maximum temperature of the area. We have also indicated the minimum temperature between brackets as the values were much below the data points and we would like to avoid cluttering the figure.

Line 245. Add : '...this move started six weeks earlier in SJNL...'

Done.

Line 246. I don't understand the meaning of this remark 'Interestingly, egg transport was always associated..' Because it is possible that this is not always the case? If the females always progress (without turning back) towards the hottest end of the rail (as the graphs show), this implies that each time they build a new nest. Perhaps this is what the authors meant.

Yes, this is exactly what we meant. We believe that this information is important, as females could move their eggs without building a nest (i.e. just deposit eggs on the sand). We have edited the sentence accordingly: "Interestingly, the mothers did not only move their eggs on the sand, but built new nests each time they moved their eggs." (L269-270)

Line 249. Add : '...was overall significantly higher in HNB compared to SJNL females (Figure 5A; 52% versus 24%; Likelihood ratio $\chi^2 = 8.52$, p-value = 0.004)

Done: "The likelihood to produce juveniles (i.e. that at least one egg hatched) and the egg hatching rate were overall higher for HNB compared to SJNL females (Fig 5A; 53% versus 20%; LR $\chi^2 = 8.52$, $P = 0.004$ and Fig 5C; 30% versus 12%, LR $\chi^2 = 6.20$, $P = 0.013$, respectively), overall higher in females previously maintained under the warmest range of temperatures (Fig 5B; 60% versus 25% and 25%; LR $\chi^2 = 8.09$, $P = 0.018$ and Fig 5D; 36% versus 12% and 15%, LR $\chi^2 = 7.99$, $P = 0.018$, respectively), and not affected by the interaction between these two factors (LR $\chi^2 = 3.49$, $P = 0.175$ and LR $\chi^2 = 0.61$, $P = 0.737$, respectively)." (L272-279)

Line 252. Add : '...overall significantly higher in females previously maintained under the warmest range of temperature compared to the two others ranges (Figure 5B, 60% versus 25%; Likelihood ratio $\chi^2 = 8.09$, $P = 0.018$).

Done: "overall higher in females previously maintained under the warmest range of temperatures (Figure 5B; 60% versus 25% and 25%; Likelihood ratio $\chi^2 = 8.09$, $P = 0.018$), ..." (L254-255)

Line 257-260. I think that the populations were not chosen randomly, and that the results of the genetic analyses were expected. As I wrote earlier, this choice of populations should be made explicit. It would also seem more logical to present the results of the genetic analyses before the results of the experiments themselves.

The populations were selected randomly and the genetic differences were not expected. In addition to the edition of the entire manuscript to clarify this point (see response above), we have edited the text to clarify that “These two populations were selected as independent random units on the basis that they were 2500 km apart by land (1000 km by sea) and had comparable climatic conditions (Fig 2). Subsequent genetic analyses revealed that HNB individuals belonged to the *F. auricularia* genetic clade 'A' and SJNL females to the *F. auricularia* genetic clade 'B' (see details in the supplementary material and discussion; Wirth et al. 1998, González-Miguéns et al. 2020).” (L133-138)

The authors mention a species A and a species B ? does this refer to previous work describing the existence of two species ?

Yes, it does refer to previous works describing the existence of these two subspecies. We mention some of these studies at the beginning of the material and methods: “Subsequent genetic analyses revealed that HNB individuals belonged to the *F. auricularia* genetic clade 'A' and SJNL females to the *F. auricularia* genetic clade 'B' (see details in the supplementary material and discussion; Wirth et al. 1998, González-Miguéns et al. 2020).” (L135-138)

Discussion

The discussion is really very interesting. But reading the lines 321-342, I think that all or part of the information given (existence of two species, hypothesis of local adaptation, etc.) should be taken up again to better contextualise the study in the introduction. This would also help to better understand the generic question (local adaptation versus novel species-specific traits) and make the study described even more interesting. In fact, this is a key question of the study. However, in the presentation of the objectives (lines 109-112), I realise on rereading that this question is somewhat drowned in a series of observations (whether earwig females select an oviposition site etc) which are in fact only the 'how' the authors tried to answer their biological question.

Following the suggestion of reviewer 1, we have reframed the study to clarify that its main goal was not to study population-specific patterns (or species-specific patterns) but to reveal the occurrence of the pre- and post-oviposition behavioural strategies in the European earwig. The fact that the modality of expression of these strategies varies between the two (random) populations raises new questions that are detailed in the discussion (L348-379), but for which our experimental design cannot provide robust conclusions. That is why we have moved all the genetic analyses in the supplementary material and emphasized that the use of two populations was only meant to take into account frequent inter-population variation in life-history traits in this species (e.g. L19-20; L114-118; and L133-135)

To my mind, the biological question is what kind of strategy earwigs adopt to minimise the risks to their eggs during winter. So, this title would seem more catchy to me: ‘Pre- and post-oviposition behavioural strategies in an insect with parental care to protect eggs against extreme winter cold’

This is an excellent suggestion. Thanks. We have changed the title accordingly: “Pre- and post-oviposition behavioural strategies to protect eggs against extreme winter cold in an insect with maternal care” (L1-2)

DATA ACCESSIBILITY. Unfortunately I did not take the time to go through the files provided in the Zenodo archive in detail. On the other hand, I underline this effort to open up the data!

Thanks.

REFERENCES

Please review carefully the list of reference.

There are many corrections to be made

Line 364 : <scp> Forficula auricularia </scp> ,

Line 369 : Hydroporus spp. , species name to be italicised etc

We have done the suggested correction and carefully reviewed the list of references to correct the (numerous) other typos.

In the spirit of data open access and interoperability, I encourage authors to add DOI links or links to download PDFs, where possible

Done.

Reviewed by Ana Rivero, 10 Feb 2022 09:32

This manuscript explores the effect of temperature on the fitness and oviposition site selection behaviour of two different populations of the European earwig. For this purpose, the authors collect adult females from the two different populations in the wild. These females are then provided with 3 (partially overlapping) temperature gradients and their behaviour is quantified.

General appreciation

Although the model is fascinating and the question worth exploring, there are in my opinion several substantial issues that need to be addressed.

The paper has two aims. The first aim is to demonstrate that earwig females select oviposition sites and move eggs around according to the environmental temperature. Their results are interesting in that they clearly show that females move their eggs towards the highest temperature available. However, despite the paper hinging on the potential benefits of egg transport as a strategy for avoiding cold temperatures, the experimental design does not allow a full assessment of the costs of benefits of moving eggs around. Although Figure 5b shows that eggs in the higher temperature gradient do better, I feel that more could have been done to explore this issue further (eg comparing the hatching rate of eggs that are moved around vs eggs that are kept at the highest optimal temperature at each range throughout their development, or comparing gradients with a different distance between the lowest and the highest temperature).

The main goal of our study is indeed to test whether earwig females select oviposition sites and move their eggs according to the environmental temperature.

Unfortunately, our data do not allow comparing the hatching rate of eggs that are moved vs eggs that are kept at the highest optimal temperature, or comparing

gradients with a different distance between the lowest and the highest temperature. This is because the hatching rate is very low overall (Figure 5) and so we have only a few points of comparison between treatments, while we also need to control for the effect of population on overall hatching rate (Figure 5) and egg transport dynamics (Figure 4). There are therefore a fairly large number of parameters for a very small number of points (hatchings) and each of the suggested comparisons is likely to provide results that are poorly robust and highly speculative.

Nevertheless, we believe that our experimental setup (climate cabinet) allowed us to assess the costs and benefits of maintaining eggs under different thermal stresses. It demonstrates that maintaining eggs at temperatures below 2.7°C significantly reduces the hatching rate (we have added a new analysis about hatching rate L271-279). Our data thus shows that eggs that would be maintained under these temperatures are more likely to suffer from a low hatching rate – indicating that egg transport toward warmer temperatures (the behaviour we report in this study) can provide benefits to eggs and mothers. We have clarified this point in the discussion: “When experimentally exposed to temperatures below 5.5°C after oviposition, earwig mothers of both populations transported their eggs to warmer locations. Interestingly, these eggs were less likely to hatch when mothers were experimentally prevented to reach such warmer locations, i.e. when mother and eggs were maintained in the cold and intermediate temperature ranges. These results overall support the hypothesis that egg transport is an adaptive post-oviposition behaviour by which earwig mothers protect eggs against extreme cold and/or adjust the thermal needs of their embryos.” (L314-320).

We have also edited the text to mention that our results “[...]calls for future studies on the physiological costs of egg transport for females at a time when they typically stop their foraging activity (Kölliker 2007)(but see Van Meyel and Meunier 2020), and on the impact of temperature variation during egg development (see Figure 4) on hatching success and offspring quality.” (L324-327)

The second aim of the experiments is to establish differences in the behaviour of the two different earwig populations. The large majority of the paper is dedicated to exploring and interpreting these differences. However, very little context is initially given as to why these two particular populations are compared, and what biological insights may eventually stem from the comparison. Both populations experience very similar mean temperatures in the wild (Figure 1) and no further information about other environmental differences between these two populations is provided. When we get to the discussion, we learn that these two populations correspond in fact to two different, previously described, cryptic species (named Species A and Species B). I am not convinced that any meaningful conclusions can be obtained from the comparison of a single population of each species (or subspecies, the authors use both terms alternatively) as genetic and environmental factors will be confounded. In addition, to what extent these differences may reflect differences in female condition at collection (nutrition, phenology, infection status etc) is not known, or discussed.

We apologize for the lack of clarity about the goal of our study regarding the use of these two populations. These populations were actually used are mere replicates because previous studies reported inter-population variation in several life-history

traits in this species. We have edited the entire manuscript to make its goal clearer (see responses to the comments above).

We had not specifically anticipated any differences between these two populations and were also very surprised to detect species B in one of the two populations. We have added this information in the text: "... our genetic analyses reveal that females of Harvey station belong to the genetic clade 'A' and females of St John's to the genetic clade 'B' (this was surprising, as this is the first time that the clade 'B' is found in Canada outside British Columbia)." (L364-366).

As our study was not designed to shed light on why the pattern of behavioural expression is population dependent (it was designed to test whether the behaviours occur generally in that species), we fully agree that a comparison of these 2 populations is not sufficient to provide meaningful conclusions. This is what we tried not to do, even though our manuscript was not clear enough on this point. Sorry for that. In addition to the edition described above, we have changed the text to discuss that this pattern could be due to local adaptation to environmental conditions, population-specific differences in the phenology of females at the moment of field sampling and/or could reflect traits specific to different genetic clades within the European earwig complex (all discussed in L348-379) and emphasize that "better understanding what drives population-specific dynamics of maternal strategies to protect eggs against cold needs additional studies involving, for instance, several populations of 'A' and 'B' females and/or population transplants" (L376-379)

More background about the species biology would help to understand the logic behind the experimental design. For instance, these females are expected to be virgin at collection and then mated to the males in the lab (there is an unexplained reference to male collection in the m&m)? Or are they expected to be mated in the field? Can females lay unfertilized eggs? How many eggs does an average female lay? Were there any differences in weight between females of the two populations?

The experimental design is complex and I found it quite confusing: 36 females are initially collected, then there's a second part of the experiment (why?) with a further 24 females, which are kept in petridishes instead of aluminum rails. Experiments are climatic cabinets and thermal bridges (why these two different equipments?) and rails of different lengths. All this needs great clarification before the experimental design can be fully appraised.

We are sorry for the lack of clarity in the methods section. We have completely rewritten it (L130-243) to provide detailed information about the biology of the European earwig (in brief: females were expected to be mated at collection but still needed to express social behaviours with males until oviposition, females are unlikely to produce unfertilized eggs, the number of eggs produced is known to vary between populations as in the two studied populations, and we did not weigh females between the two populations), explain why we needed 36+24 females (for technical reasons), and to explain the purpose of the two types of equipment. We also endeavoured to address all the potential misunderstandings about the different goals of our study in this new section.

Below I detail these and other issues in a little bit more detail. I hope the authors find these useful.

Introduction

Q1 - L102-104 – Shorter than what? (or do you mean 'short exposure'?). Fast hatching and accelerated development are not necessarily negative traits, unless they are correlated with eg smaller size and lower lifetime fecundity. Please expand on what is meant by 'alters the immune system' : what is altered and in what way? Is this alteration correlated with an increased risk of infection. This is key as the whole paper hinges on the potential benefits of avoiding cold temperatures.

Sorry for the lack of clarity. We have edited the sentence to clarify that “a recent study shows that prolonging egg exposure to cold (5°C) for 15-day during winter delays the hatching date and development of juveniles to adulthood (which typically takes two months in *F. auricularia*), leads to the production of lighter adult females, and shape the basal immunity of these females: it increases their overall phenoloxidase activity and reduces the number of haemocytes in females facing a changing social environment (Körner et al. 2018). Because these traits are likely to affect negatively and/or positively the fitness of the resulting adults (Koch and Meunier 2014), temperature-dependent oviposition site selection and egg transport during development could be adaptive strategies in *F. auricularia* mothers.” (L99-107)

Q2 - L96 – As the term "population" is used interchangeably with that of species and subspecies across the paper, it is not obvious what this sentence refers to.

We have changed the sentence to remove the term population: “First, the duration of egg development in winter varies from three weeks (e.g. in Southern Europe) to several months (e.g. in North America) (Ratz et al. 2016; Tourneur 2018).” (L95-97). More generally, we have standardized our use of population (and avoided the use of species) in the rest of the text.

Q3 - L111 – Please provide context by explaining whether you expect these behaviours to be population specific. Judging by Figure 1 these two populations seem to be very similar temperaturewise.

We have edited the text accordingly: “Because previous studies point out that the European earwig may show population-specific life-history traits and behaviours (Ratz et al. 2016; Tourneur 2017; Tourneur 2018; Tourneur and Meunier 2020), we also tested whether temperature-dependent oviposition site selection and egg transport vary between two (randomly selected) populations sharing comparable climatic conditions.” (L114-118)

Materials and Methods

Q4 - L122 – Why collect males?

We have edited the text to provide an answer to this question: “The field sampled females were maintained in groups with males from the same population (sampled at the same time as the females) to allow the completion of the gregarious phase of the life-cycle, which lasts several months during which females mate with multiple partners (Sandrin et al. 2015) and express simple social behaviours (Costa 2006; Weiß et al. 2014).” (L138-142)

Q5 - L126 – Please provide background about the species biology. Females are collected from the field and directly used for the experiments. Are these females fertilized? Can unfertilized females lay eggs and would you expect the same type of behaviour in fertilized

and unfertilized females? (presumably the latter should not go into any lengths to protect eggs that will never hatch?)

We have rewritten the the methods section to provide details about the species biology (L128-205). To adress the questions specifically pointed out here: field-sampled females are typically fertilized at this moment of the year, females can lay non-fertilized eggs under laboratory conditions but this has never been reported in nature.

Q6 - L137 – Please clarify what 'second part of the experiment' means and why a second batch of females was collected

We have edited the material and methods and do not refer to any 'second part of the experiment' anymore. We have also explained why the second batch of females was collected: "This measurement of egg transport involved the 36 females used in the measurements of temperature range explored before egg production (see above), and the 24 females (12 from HNB and 12 from SJNL) previously maintained in Petri dishes until oviposition. We have not been able to use these 24 females in the electrothermal plates due to the lack of space in the units." (L184-188)

Q7 - L139 – These females were not placed in a temperature gradient?

These females were not placed in a temperature gradient before oviposition, but they were after oviposition (see Q6)

Q8 - L145-146 – At this point it would be useful to understand why some females are maintained in aluminum rails and some in petri-dishes

See Q6.

Q9 - L160-165 – Sorry I got lost here. What is the difference between a climatic cabinet and a thermal bridge and why are two different equipments used? How is the temperature distributed in a thermal cabinet as opposed to a thermal bridge? What are the practical implications of some rails being shorter than others ?

That was indeed very unclear. Sorry for that. We have rewritten this entire part to clarify these points (L146-199). In a nutshell: thermoelectric plates offer a linear gradient from 0 to 20°C and allow to follow only 36 females at the same time. This was very convenient to follow females before oviposition (a moment in their life-cycle where environmental temperatures can greatly vary) and because their number was limited. Climate cabinets offer three non-linear thermal gradients (0.9°C to 7.0°C), intermediate (-3.6°C to 2.7°C) or cold (-4.5°C to 1.6°C) and allow to follow 60 females at the same time. This was very convenient to follow females in the 15 weeks following oviposition (a moment in their life-cycle where environmental temperature shows reduced variation) and because their number was relatively large. There is no implication of some rails being shorter than others, as length variation is only due to technical limits in the climate cabinet (L191-192) and there is no comparison between short and long trails.

Q10 - L163 – Confused here too: you transfer a 'random subset' of 60 females, but 60 is the total number of females in the experiment (L140). Please clarify

This was a mistake. Sorry for that. We have transferred all the 60 females and edited the text accordingly: "We finally tested whether the temperatures at which

eggs were maintained during the 15 weeks following oviposition affect egg development and hatching rate. Fifteen weeks after oviposition, we transferred the 60 (short) rails containing the mothers and their eggs (i.e. 5 rails per thermal range and population) to thermoelectric plates with temperatures linearly ranging from 0°C to 19°C (and not 20°C due to the shorter rails).” (L193-197).

Q11 - L168 – Why not count the total number of eggs hatched (L208 makes a reference to the number of eggs produced)? How many eggs does a female lay?

This is an excellent point and we have added a new analysis of the hatching rate (L271-279). It provides results that are consistent with the previous analyses on the occurrence of at least one nymph (i.e. we found an effect of the condition and the population). We would like to keep the two analyses in the manuscript, as the distribution of hatching rate appears very bimodal (Figure 5 C and D), with either a relatively high hatching rate or no hatching at all for a female. We believe that providing the two statistical analyses thus offers a more robust and comprehensive approach to our data.

The number of eggs produced by the females has been added to the main text: “Moreover, HNB females [...] laid overall more eggs (median = 68 and 46, respectively; Fig 3F; $W = 261$, $P < 0.001$) than SJNL females” (L254-257).

Q12 - L187 – The manuscript uses the term population, subspecies and species interchangeably. A background to the genetic structuring of this species with reference to Species A and Species B needs to be made in the Introduction.

Following the suggestion of the first reviewer (and because it is difficult to conclude about species difference with only one population per species), we have tuned down the between-species comparison in the manuscript and only discuss it at the end of it. We believe that it is the most parsimonious and robust decision.

Results

Q13 - L257-260 – This is a very short and not surprisingly brief account of the genetic differences found between these two populations. Please provide more detail. Context must be given in the Introduction.

Following the suggestion of the first reviewer, we have moved the results of the genetic analyses to the supplementary material.

Q14 - The results are interpreted as being the result of intrinsic differences between the two "populations". However, these experiments were carried out with females collected from the field, so to what extent the differences observed are the result of differences in e.g female condition, or physiological status at collection, or to differences in the reproductive phenology between the two sites??

We agree that our study (and experimental design) does not allow us to disentangle what explains inter-population variation in the tested behaviours and in particular to conclude about the role of intrinsic differences between the two populations. This was not our aim, and we apologise for the lack of clarity of our manuscript in this regard.

We have rewritten the discussion to clarify this point and discuss the robustness of the different explanations: “Somewhat surprisingly, our study finally reveals that the modality of expression of the reported strategies to protect eggs against severe

winter cold varies between the two studied populations. Although our experimental design does not allow us to conclude robustly on the reasons for this variation, we propose three potential explanations. First, this could be due to local adaptation to environmental conditions. In line with this explanation, previous studies reported that multiple traits can vary between populations of the European earwig, such as the number of clutches produced by females, clutch size, juvenile quality, the timing of egg production, duration of egg development and female body mass (Ratz et al. 2016; Tourneur 2017; Tourneur 2018; Tourneur and Meunier 2020). However, these studies often compared populations with contrasting climatic conditions, which was not the case between St John's and Harvey stations (Fig 2). Second, the reported variation could be due to population-specific differences in the phenology of females at field sampling. However, this is unlikely to explain our results, as the females were sampled late in the breeding season and our setup control for this potential variation by standardising the measurements around the natural day of oviposition for each female. Finally, the behavioural variation reported between the two tested populations could be due to the presence of different genetic clades. In line with this hypothesis, our genetic analyses reveal that females of Harvey station belong to the genetic clade 'A' and females of St John's to the genetic clade 'B' (this was surprising, as this is the first time that the clade 'B' is found in Canada outside British Columbia). The European earwig is a complex of cryptic species, for which genetic divergence and reproductive isolation are well established (Wirth et al. 1998; Guillet, Josselin, et al. 2000; González-Miguéns et al. 2020), but the specificity of their life-history traits remains largely unexplored. To date, the only known species-specific trait refers to their reproduction, with females 'A' producing one clutch and females 'B' producing two clutches (Wirth et al. 1998; Tourneur 2018). However, other studies demonstrate that the number of clutches produced by a female can vary within *F. auricularia* species (Tourneur and Gingras 1992; Ratz et al. 2016) depending on numerous parameters acting during the early life of an earwig female (Meunier et al. 2012; Meunier and Kölliker 2012; Wong and Kölliker 2014). If this third hypothesis is true, our results may thus have shed light on the first behavioural difference between 'A' and 'B' females. Nevertheless, better understanding what drives population-specific dynamics of maternal strategies to protect eggs against cold needs additional studies involving, for instance, several populations of 'A' and 'B' females and/or population transplants." (L348-379)

Q15 - In the introduction it is mentioned that in some "populations" eggs need to be exposed to near-zero temperatures to trigger embryo development. Figure 4 shows a dip towards 0 at around weeks 4-5 (in the intermediate and cold range) – could this be it?

Very well spotted! This pattern could indeed reflect the need for HNB eggs to experience near-zero temperatures to trigger their development. However, we have not mentioned this possibility in the text as it remains a speculative interpretation that would require further data (based on another experimental design) to be confirmed

Q16 – Figure 5 – It would be interesting to add a figure of the nb of weeks until hatching by gradient (cold, intermediate, warm)

Our manuscript already includes several figures and this additional figure will only include a limited number of data points due to the low hatching rate in the cold and intermediate levels. We therefore would like not to add it to our manuscript.

Grammar and style suggestions

The first paragraph of the Introduction is very long and could be shortened to make it more focused.

We have shortened the first paragraph.

L27 – Reword sentence. I would not expect oviposition site selection and egg transport to be mutually exclusive?

Done: “More generally, it also reveals that egg care and/or egg transport do not prevent behavioural thermoregulation via oviposition site selection and highlights the diversity of behaviours that insects can adopt to enhance their tolerance to global climate change” (L30-33)

L42 - habitats (instead of habitat)

Changed.

L76 – predation (instead of predations)

Changed.

L82 – Take off "by holding them in their mouth": it references a very specific form of egg transport while, presumably, the statement may also apply to other species with other transport strategies

Removed.

L87 – until they hatch

Changed.

L89 – ,the application of chemical compounds....and fierce protection....

Changed.

L101 – egg exposure

Changed.

L111 – egg ag

Changed.

L116 – juvenile production

Changed.

L129 – I found the term thermal "bridge" confusing, I looked for the reference in the manufacturer's website and this is described as a thermoelectric plate. Wouldn't it make more sense to call it like that?

This is an excellent suggestion. We have changed “thermal bridge” with “thermoelectric plate” in the entire manuscript.

L138-144 – This should be said earlier (L125)

We have completely rewritten the methods section and now provide this information earlier in the text.

L152 – 'with eggs deposited' replace with new sentence 'Eggs were deposited in the middle of the range corresponding to 5.22o C (warm)...'

Changed: “To experimentally change the temperature at which females and eggs were maintained at oviposition, we deposited each female and its eggs in the middle of the new rail (Fig 1C), i.e. at either 5.2°C (warm), 1.2°C (intermediate) or 0.1°C (cold).” (L179-182)

L215 – interaction

Changed.

L257 – Reword 'each population contained a different member of...' is not what you want to say here.

We have edited the sentence: “The COI analyses revealed that all (6/6) HNB females belonged to the species “Forficula auricularia A” and all (6/6) SJNL females belonged to the species “Forficula auricularia B” (L714-716) which is now in the supplementary material

L242 – week

Changed.

L247 – 'new nest in which the eggs moved' – please reword

We have edited the sentence: “Interestingly, the mothers did not only move their eggs on the sand but built new nests each time they moved their eggs.” (L269-270)

Figure 4 – The * for the p values (comparing each point to the reference point) clutter the figure and are not helpful: the standard error bars already do a pretty good job of showing us which values are significant. Set up reference horizontal lines for min and max temperatures within the thermal range, this will provide a useful visual image of the range of temperatures explored within the range provided.**

We have added a red horizontal line in **Figure 4** to show the maximum temperature of the area and indicated the minimum temperature between brackets to avoid cluttering the figure. We believe that the stars are still helpful, so we have moved them closer to the data for easier reading.

Figure 5 legend – 'Effect of population (A) and thermal constraint (B) on the percentage of females with at least one hatched. Effect of population on the location of the eggs at the time of hatching (C) and on the number of weeks between oviposition and egg hatching (D)'

Changed.

Figure S2 – Please provide axis labels

Changed. Note that we have combined Figure S1 and S2 following the suggestions of the other reviewers.