**Reviewer: 1**  
  
**COMMENT:** Firstly, the authors conduct their experiment at a high (stressful) temperature, and give this much emphasis (e.g. the title, and abstract (lines 18-19) and intro lines (line 69-70)). I find this emphasis unwarranted as although the experiments ARE done in an extreme environment, there are no control experiments done at a more benign temperature. This means it is difficult to say that “polyandry helps colonization success in an extreme environment” over the simpler “polyandry helps colonization success” (irrespective of the environment). Certainly, it is interesting that this can occur in a stressful thermal environment (and this should be discussed in relation to work in other systems which tend to use a more benign environment), but the emphasis should be removed.

**RESPONSE:** This is a fair point, and we have reduced the emphasis to climate change and extreme temperature throughout, including a change of title and a reframing of the abstract and introduction. We agree that this simplification is a better representation of our work.

**COMMENT:** Secondly, although the authors talk about single-mated vs doubly-mated females I don’t think matings were actually observed? (Lines 84-91 – ‘opportunities to mate’). Therefore, it is not clear if how many times/males a female has mated. This needs to be clarified in the methods as throughout the authors refer to ‘doubly mated’.

**RESPONSE:** Yes this is correct. We have changed the phrasing in the abstract, which now refers to opportunities to mate, and we have added a clarification in the methods, lines 122-124:

“We note that with our experimental design we cannot be sure that all polyandrous females mated twice, and that this may result in reduced power to distinguish the true effects of polyandry.”

**COMMENT:** This issue also impacts the interpretation of results. For instance, line 167-175: “We found that 11% of monogamous females produced no offspring in the first colonisation generation, while all polyandrous females produced offspring. This is consistent with a situation in which a failure to produce offspring is the result of male infertility”. This is true, but another plausible explanation is that in these cases females did not mate at all (which may be more likely when presented with the same male twice). The discussion of this issue should be modified to include this possibility.

**RESPONSE:** We have now modified the discussion to include this possibility, lines 216-222:

“We found that 11% of monogamous females produced no offspring in the first colonisation generation, while all polyandrous females produced offspring. The low number of extinctions here mean that caution is required when interpreting this result. Nonetheless, these percentages are broadly consistent with a situation in which a failure to produce offspring is the result of a failure to mate, or due to infertile or incompatible males, from which we can expect only 1.28% of random pairs of males in the double-mating treatment to both be infertile or incompatible.”

**COMMENT:** Lines 131-135 and Fig 2 – “we observed a clear and consistent trend for larger adult population sizes in populations founded by polyandrous compared to monogamous females” - Please test this statistically.

**RESPONSE:** This is tested in the following paragraph - we now make this clear.

**COMMENT:** Finally, the authors have produced quite a lot of useful data, however, they have not made it available! Please add the raw data to dryad or a similar repository. In addition, the authors fit several complicated models to their data, but have not provided their R code. Please provide this as a supplemental file or (better) add it to a code archive such as github/zenodo.

**RESPONSE:** Our apologies, the data are available on Github, as is all the code for a fully reproducible analysis. This is now made clear in the manuscript (we had to remove it as JEB has double blind review and the github site identifies the authors)

**COMMENT:** Fig 1: change “gen 2” to “gen N” to better show the experiment continued over many generations.

**RESPONSE:** Done

**Reviewer: 2**  
  
**COMMENT:** I thought the experimental design was good overall, although I was a bit puzzled by the lack of a control temperature treatment. Previous results in this species did not find any difference in extinction risk between mating treatments in a benign environment, so the moivation for using a high rearing temperature in this experiment was to see if there is a larger effect in a stressful environment. By not including a benign temperature treatment it's not possible to do any direct comparisons between the benign and stressful environment, which is a bit of a shame.

**RESPONSE:** See response to reviewer one – the aim was to test for effects of polyandry rather than temperature, and we have changed the emphasis of the manuscript throughout to reflect this

**COMMENT:** I don't see any problems with the statistical approaches that were used. However I don't think any robust conclusions can be drawn about the risk of immediate extinction (discussed on lines 166-171). Although immediate extinction rate (zero offspring production in experimental females) was indeed higher in the singly-mated females, the difference between the treatments is not significant when tested using a chi-squared test (P=0.73), probably due to the low total number of immediate extinctions.

**RESPONSE:** We have now added a caveat to the relevant section of the discussion, lines 216-222:

“We found that 11% of monogamous females produced no offspring in the first colonisation generation, while all polyandrous females produced offspring. The low number of extinctions here mean that caution is required when interpreting this result. Nonetheless, these percentages are broadly consistent with a situation in which a failure to produce offspring is the result of a failure to mate, or due to infertile or incompatible males, from which we can expect only 1.28% of random pairs of males in the double-mating treatment to both be infertile or incompatible.”

**COMMENT:** One thing I was curious about when reading the manuscript is whether anything is known about heterozygosity levels in this population. It would be interesting if expected inbreeding levels could be calculated for each mating treatment.

**RESPONSE:** We agree that this would be an interesting exercise. However, we only know that the strain we used is expected to be outbred. So, while interesting, it would not be possible to calculate expected inbreeding levels.

**COMMENT:** My main criticism with this manuscript, however, is that it's not really clear to me what the novelty of the results is. It seems totally unsurprising that more inbred populations (monandry) should have poorer long-term survival compared to less inbred populations (polyandry). If the main novelty is that the effect is stronger under stressful than benign conditions, then it's a substantial weakness that this hypothesis can't actually explicitly be tested with the dataset at hand. In line with this, I found the discussion to be quite narrowly focused on the study species, with no real discussion of the broader relevance of the results.

**RESPONSE:** We have now revised the emphasis of the manuscript to make it clear that the aim is to study long-term survival in relation to mating regime. While intuitive, this theory has surprisingly little empirical support. We therefore believe that our data and findings will be a useful addition to the evolutionary biology literature.