Associate Editor Evaluations:   
Recommendation: Invite Resubmission   
  
Associate Editor (Comments for the Author (Required)):   
  
The revised version of the manuscript has been reviewed by the same two original experts. They both acknowledge the effort performed by the authors to clarify the substantial issues raised during the first round of revisions. However, they still have several (mostly minor) suggestions to further clarify and improve the clarity of the manuscript. Specifically, reviewer 1 requires further clarification of the statistical analysis and suggests that a careful proof-reading and editing of the text is needed to simplify complex sentences and avoid grammar errors. Reviewer 2 also suggests several issues where further detail or clarification is needed. Both of them agree, and I concur, that these changes should be manageable and feasible to address in a second revision.   
  
  
Referee #1 Evaluations:   
Recommendation: Minor changes needed   
  
Referee #1 (Comments for the Author (Required)):   
  
I salute the effort of the authors for implementing significant changes to their manuscript. Initially, I noticed the substantial work put into the resubmitted version, but I missed the response to my major points and asked for them. I appreciate the authors addressed these concerns, and the responses seem in general sound to me. Overall, the manuscript has improved in terms of clarity and coherence. For example, the new conceptual Figure 2 enhances understanding. The two experiments are now better separated and clarified, and result interpretations appear appropriate now. However, some parts of the text still need improvement in terms of structure and clarity, requiring a careful rereading.   
  
Some other general comments that you should consider are:   
-I would like to discuss further the response regarding the statistical analysis. If I understand correctly, the lack of substantial distance among the three (sub)populations in the southern region led you to include them in a single region, while the others in the northern region were more separated. I also understand that the main question revolves around the origin (and thus region), and that including populations could introduce noise to the model if treated as a fixed factor. However, it seems clear to me that you have a nested design with several populations within one region, and that potential variations among populations should be accounted for in the model, possibly as a random factor. In fact, I don't see how adding this nesting design (either as a nested fixed factor or in the random term) could lead to the multicollinearity or overfitting you mentioned encountering when running your models. If these issues arise, it seems more likely to be related to several traits/variables exhibiting high correlation or an issue with programming, data, or model parameterization (in fact, probably if you included populations and regions separately as fixed terms, it can explain some R errors you received, because this would be not correct). I would recommend checking it, and adding populations within regions as a random factor.   
We included population as a fixed effect in all but one model. For almost all models we had the warnings, “fixed-effect model matrix is rank deficient so dropping 1 column / coefficient” and in some cases, “boundary (singular) fit.” We did not include population for pollen germination at 40 degrees Celsius for Experiment 2, because there were not enough degrees of freedom to estimate coefficients for the region term. We included all statistical information on population in the supplementary information. The results did not change too much by including population.

-Please, highlight in brief in the discussion the limitations associated with including temporal blocks in the experimental design.   
Addressed by adding a sentence explicitly stating the limitations associated with a temporal block.

-Even you decided to start showing the simple effects before explaining significant interactions, be very aware that the effects on simple terms can be misleading if the two-way interaction is significant, potentially leading to incorrect or inaccurate interpretations.   
We acknowledge this, but still think it is useful to discuss the individual variables in isolation. Steve, do you have anything to add?

More specific comments:   
-In the pollen germination experiment, I would account for potential differences among petri dishes in the statistical models (if different replicates, you may add this in the random term), as this can mediate their response.

If we understand this comment correctly, the reviewer wants us to include a term for petri dish in our model to account for pseudoreplication. We did not have multiple petri dishes for each individual and we did not have pollen of multiple individuals on one petri dish. We had one petri dish per individual per temperature. We are using values extrapolated from a temperature performance curve. The response variable for one individual (one data point) is based on a curve fitting the pollen germination or pollen tube growth rate for five petri dishes at 5 temperatures. There is no psuedoreplication here other than among ramets from a genet, which is already in the model.

L6-7, and L10-11: According to your response to one of my previous major concerns, is more that you test for thermotolerance to extreme temperatures, so I would be more clear here.

Addressed by adding the clarification that we are interested in both extreme and moderate increases in temperature.  
L12: Change "asked" to "studied" or related verb.

Addressed by making this change.  
L13-15: The excessive use of "that" (three times in one sentence) makes the sentence difficult to follow.

Addressed by changing “that” to “which” in one case.

As noted previously by reviewer #2 and myself, the text seems to still need more care in several places. I won't go through every instance I noticed it, but I strongly recommend doing so or asking for a blind reading from a colleague.

NEEDS TO BE CHECKED  
L58: Remove "were" from "than plants from" and add an "s" to "seeds." This is another example of needing to go carefully through the text.

Addressed by simplifying the sentence.  
L68: With your design, are you able to find evidence of genetic divergence? I don't think so. Once more, I don't think it's structurally appropriate to include your study in this part of the introduction.

Addressed by deleting the sentence.  
Fig. 2: In the figure caption, explain a bit more this beautiful conceptual diagram, such as the measure of reproductive vs vegetative variables or what the outcomes represent. Also show that Exp. 1 is on the left and Exp. 2 on the right.

Addressed by adding sentences briefly describing the experimental design and predicted outcomes.  
L426-427: Indicate that you implemented six separate models, i.e., one per vegetative trait.

Addressed by adding this explicitly in the data analysis section and the first sentence of the results.  
Table 1: These degrees of freedom seem not clear to me. Did they represent the factor and total degrees of freedom? Please clarify this in the table caption.

Addressed by adding “degrees of freedom (dF: numerator, denominator)” to describe the factor. I also added that degrees of freedom were estimated using Satterthwaite’s method.  
L730: You start writing "Regional differences" and then go to another paragraph. Take care of it.

“Regional differences” is a heading. It is difficult to see when the tracked changes are on.  
L816: I would write "among the maternal and paternal genotypes".

Addressed by changing to “among.”  
L823: Maybe refer in that cases as "extreme" heat or cold.

Addressed by adding “moderate” heat for the Experiment 2 part. For Experiment 1, we tested multiple temperatures (not just the most extreme) and heat in general reduced pollen germination.  
L961: What is the main conclusions you can draw from your experiments?   
Addressed by adding a wrap-up sentence with the main conclusion. Steve, Please check this.  
  
Referee #2 Evaluations:   
Recommendation: Minor changes needed   
  
Referee #2 (Comments for the Author (Required)):   
The overall quality of the manuscript has greatly improved after comments from the reviewers have been addressed. However, I still have a few minor comments to add to the new version of the manuscript that the authors can address at their discretions:   
  
L22-25: In the sense of global warming and rising temperatures, I am not sure if I would include local adaptation in this context as a method to mitigate stressful conditions. But phenotypic plasticity is very relevant.

Addressed by omitting local adaptation.  
L112-114: Needs reference.

Steve, I can’t tell what sentence they are talking about here.  
L138: I was missing the information on the sample size in the beginning of the previous paragraph.

Addressed by adding the number of plants collected in each region to the second paragraph in the methods.   
Fig. 2: Super nice conceptual figure that helps a lot with the understanding of the experimental design! Maybe an arrow could be added from Experiment 1 to Experiment 2, indicating: "Two ramets of all 26 genets..." (L. 216).

We decided to leave the arrow out as the figure is already quite busy and thought that the arrow would need additional explanation.  
L266-271: I would move this paragraph to the beginning of the section.

Addressed by moving paragraph to the beginning of the section.  
L298: I know that I commented on this before, but I still don't think the trait name "style+stigma" is a great trait name. Maybe the anther and style+stigma traits could be renamed to female versus male organs or something along those lines?

Addressed by changing the terms to “Male structure” and “Female structure.”   
L312: Differed significantly between regions?

Addressed by adding the term “significantly.”  
Discussion: It was a very nice introduction that now included some more information about phenotypic plasticity which this manuscript evolves around. However, "phenotypic plasticity" has not been mentioned once in the discussion. I think it might be a good idea to add where relevant to get the keywords in the main text.

Addressed by adding “phenotypic plasticity” in a couple discussion paragraphs.