Dear recommender,

Thank you for the opportunity to submit our revised manuscript.

We took into account all your comments and the comments of the two reviewers (in blue in the following text). We brought a particular attention to clarifying the methods (QST calculation).

The suggested modifications improved the quality of the paper. We thank you and the reviewers for your comments and hope that you will find the manuscript suitable for recommendation. Please find below a detailed response to technical requests, recommender comments, and reviewers comments.

We uploaded a version of the revised manuscript where modifications appear in the track changes format so that they can be visualized easily. We also uploaded a clean Latex version of the revised manuscript using the PCI Evol Biol template.

Looking forward to hearing from you.


Regards,
for the authors,
Benoit Pujol


--------------------------------

**RECOMMENDER'S COMMENTS**

Abstract: the words « potential » or « potentially » appear 6 times in the abstract. It is good to be cautious but surely some of these could be reworded (check lines 21-24 out in particular).
Done, we have deleted and/or replaced 4 of them.

Line 249: I am not sure what "linearized" means here.
Done, changed to "normalized".

Line 263: replace 'last' with 'latter'.
Done

Lines 282-285: this latter analysis seems to be the most comprehensive, and the one you really need to address your main questions. What is the point of the other Mantel tests described above?
We agree that the last Mantel test that we performed (population pairwise $Q_{ST}$ and population pairwise altitude differences while controlling for neutral genetic differentiation) is the most important to address our main question. However, Mantel tests are only correlation tests. We believe that the step by step procedure that we used here strengthen our results. The correlation between Qst and Fst was not significant and the correlation between Qst and altitude was significant. As a result, the results of the last Mantel test appear more reliable.

Lines 312-317: Could you confirm that the difference in the correlation between trait values and altitude between the two taxa is not due differences in the variance in altitude across populations for each taxon? Is variance in altitude the same for both taxa?

The variance in altitude between subspecies was quite consistent. Yes, we can confirm that the variance in altitude was not significantly different between subspecies (anova p.value= 0.4). We added text to clarify this point in the methods (LL 163-165 and supplementary information).

Lines 332-336: do you mean that the results obtained via the examination of overlapping confidence intervals are fully consistent with those obtained by the bootstrapping method? If so, this is not entirely clear from this sentence.
Yes, this is what we meant. We have modified the sentence to clarify this point LL348-352.

Lines 384 and 385: replace two instances of "less" with "fewer".
Done

---------------------------------

**TECHNICAL COMMENTS**
**Additional Comments of the managing board**:
**Mandatory modifications**. As indicated in the 'How does it work?' section and in the code of conduct, please make sure that:
-Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.
Done (data files deposited on ZENODO) 10.5281/zenodo.3821949

-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.
Done ("read me" listing the data and analysis files and ready-to-run-analyses (R protocols) calling for the data files automatically are deposited on ZENODO) 10.5281/zenodo.3821949

-Details on experimental procedures are available to readers in the text or as appendices.
Done

-Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI XXX recommenders."
Done

In order to reach a better referencing and greater visibility of your recommended preprint, we suggest you to do the following modifications:
- **add the following sentence in the acknowledgements**: "Version 3 of this preprint has been peer-reviewed and recommended by Peer Community In Evolutionary Biology (https://doi.org/10.24072/pci.evolbiol.100097)"
Done

• in the Zenodo metadata "Notes" you can indicate: "This article has been peer-reviewed and recommended by Peer Community In Evolutionary Biology

([https://doi.org/10.24072/pci.evolbiol.100097](https://doi.org/10.24072/pci.evolbiol.100097))"
Done

- in the Zenodo metadata "Related article" you can indicate: "Reviewed by PCI Evol Biol 10.24072/pci.evolbiol.100097"

Done

**Optional modifications**. ==> **We advise you to use templates** (word docx template and a latex template) to format your preprint in a PCI style. This is optional. Here is the links of the templates: [https://peercommunityin.org/templates/](https://peercommunityin.org/templates/)
Done

For word template, please be careful to correctly update all text in these templates (doi, authors' names, address, title, date, recommender first name and family name …). Please be careful to choose the badges "Open Code" and "Open Data" only if appropriate (in addition to the "Open Access" and "Open Peer-Review" badges). If some of the reviewers are anonymous, indicate for example "Albert Ayler and two anonymous reviewer". Indicate in the "cite as" box the right version of your preprint. It is version 3.

For Latex template, main.tex and sample.bib should be filled. Please be careful to choose the badges "Open Code" and "Open Data" only if appropriate (in addition to the "Open Access" and "Open Peer-Review" badges). Preambule_xxx.tex should be modified (comment lines 115, 117) to select badges. If some of the reviewers are anonymous, indicate for example "Albert Ayler and two anonymous reviewer". In sample.bib, indicate the right version of your preprint. It is version 3.
Done

--------------------------------

**REVIEWS**

REVIEWED BY SANTIAGO C. GONZALEZ-MARTINEZ, 2020-04-10 10:08
In this second round of revision, I only have a major comment. I am afraid that the equation provided to compute Qst, Vb / (Vb + 2$h2$Vw), is wrong. The correct denominator of this equation is the sum of the variance among populations (Vb) and twice the additive variance within-populations, which in your case (full sibs) is twice the family variance (Vw), thus the equation should read: Vb / (Vb + 2*Va) or Vb / (Vb + 4*Vw) or Vb / (Vb + 2$h2$Vp), with Vp = Vw + Vres. I tried to check out if this is just an erratum by recomputing Qst from the variance components from Table S3 but Qst values do not seem to fit with those in Figure 4 (Qst values themselves are not reported in the manuscript) in any way I compute them. Please, check this out, as it may imply major revisions to the manuscript.
You are right; there was a transcription error in the methods. Qst should have appeared as $Q_{ST}$ = Vb / (Vb + 2 h² (Vw + **Vres**)), as you noted.
Very worried, we screened all our R code and detected an error (the Vres was there but the factor 2 had jumped). We re-ran all the analyses.
It did not change the results much, the values barely changed. The only noticeable change was one $Q_{ST}$ value in *A. m. pseudomajus*, which P-value was not significant by much, that became non-significant. Nothing changed for *A. m. striatum*. In fact, this makes our conclusion on the potential absence of local adaptation in *A. m. pseudomajus* more consistent, which indirectly strengthened the observed contrast with the potential presence of local adaptation linked to

altitude in *A. m. striatum*. This reinforced our conclusion that different mechanisms are at work in the two subspecies.
For the sake of clarity and transparency, we added the values of the QSTs in Table S3. We also provided the code and data with the MS.


Minor comments:
Ln 35. Replace "subspecies" by "subspecies divergence".
Keywords: done

Ln 49-51. Can you be a bit more specific?
Done, we added text to provide ecological context to the citations made in the text (LL 48-58)

Ln 100-102. Not clear to me why you would not expect shorter internode length at high altitudes, as I would expect than to only height but also growth rate would be lower (but I am not an expert on this, so I am probably wrong). In any case I like very much this paragraph setting the expected trait variation with altitude.
In fact, a lower biomass might not be due to a lower growth rate at a higher altitude. We agree that on the one hand, this may be expected because the final biomass or height should be lower at a higher altitude. However on the other hand, plants are also expected to germinate later at a higher altitude, and consequently to have a shorter growth period. These plants might then have a higher growth rate at a higher altitude to reach a lower vegetative biomass required to reproduce in due time.

Ln 111. 'based on trait heritability', a bit weird to express it in this way... perhaps better to say 'based on partition of genetic variance'?
Done

Ln 138. I guess that you estimated genetic differentiation using Fst? Can you add this between parentheses?
Done

Ln 159. It is perhaps not very important for this study, as climatic data is only used to characterise the populations, but worldclim does notably bad in mountain regions (in particular for precipitation) and alternative climate databases are preferable.
We agree that wordclim might lack some precision in mountain regions. It was nevertheless used satisfactorily in previous work to characterize the ecological niche of *A. majus* at the level of populations, across the altitudinal gradient in the Pyrenees, and across the species geographic distribution (Khimoun et al. 2013. *Evolutionary Ecology* 27: 51–64.). We'd rather keep this approach that has been proof tested in previously published work.

Ln 205. I assume that you mean '…the 637 plants that germinated'? (As there should have been 40 x 24 = 960 pots). Please, specify.
Yes indeed, we clarified this point by adding "that germinated" in the text.

Ln 221/230/241. '… plastic container ('block') as a random…'
Done

Ln 231. '…where Vw is the family…'
Done

Ln 243. Please, see my major comment above. Hope it is just an erratum!
See response to major comment above.

I think you would need to report Qst values somewhere in the Results (perhaps adding these values to Table S3?).
Done, Qst values have been reported in Table S3.

I also miss some discussion on these values and the comparison with values for other plants in the literature.
In the results section, we added text to discuss overall Qst –Fst differences found in our study and those found in the literature (L359-366). We are not aware of any other studies using a Qst-Fst approach on this species or any other *Antirrhinum* species.

Ln 323. 'Several traits had similar…
Done

Ln 366-367. I would say that signatures of potential selection were found for only few traits. Do you think this is because low power of the study (CIs for Qst are large, judging by Figure 4) or does it truly reflect the species' biology?
We now mention that "signatures of potential selection were found for only few traits" L392 in the discussion. We added text in the results LL359-366) where we acknowledge the low power of our approach. Significant QST-FST differences were likely very conservative in our study. One must note that our conclusion about *A. m. striatum* and *A. m. pseudomajus* differences is robust to this lack of power.

Ln 389. Should this section not come before the previous one? (As it talks about overall differences in adaptation patterns across subspecies)
We'd rather keep it that way. We adopted this logical structure following the suggestions made during the first round of reviews. We believe that this logical structure of the narrative is well suited as it goes from observed global patterns (altitude signal) to observed detailed patterns (subspecies and populations) to mechanisms underlying detailed patterns (ecological and then population genetic aspects).

Overall, the Discussion section is still a bit reiterative and, towards the end, it puts too much emphasis on adaptive divergence.
We chased the reiterations and removed them as much as we could as to lower the emphasis on adaptive divergence.

This is surprising as adaptive divergence was found for only a handful of traits in this study. I think that a section devoted to why the study did not identify adaptive divergence in more traits (some of them with obvious adaptive value, e.g. SLA) and why altitude did not seem to be a major factor for adaptive divergence in this system would be desirable.
We agree that these results were surprising. We added text to discuss (i) the power of the analysis (see comment before), (ii) whether other studies found similar QST-FST differences (see comment before), (iii) had already discussed expected trait adaptations in response to altitude in the introduction in the previous version (including SLA), and (iv) added text in the discussion to address the lack of altitudinal signal for SLA in both subspecies (LL413-415). We'd rather not speculate on the adaptive value of *A. majus* traits that our statistical analysis did not identify as potential candidate traits involved with adaptation. We'd also rather not

speculate on the reasons "why altitude did not seem to be a major factor for adaptive divergence" because we already discuss our results that altitude is involved to some extent with adaptive divergence. We hope that additional studies will be able to bring further elements to answer these questions.

Figure 2's legend. No need to report the regression fits as they are not significant.
Done

Figure 5 needs updating (these linear correlations were removed from the manuscript, thus should not be reported in this figure). Also, why Fst in the y-axis?
Done, we removed the regression equations, and added text in the legend to specify that these graphs illustrate the relationship formally tested by mantel tests and that the Fst on the y-axis refers to the grey dots.

All in all, this is a very nice paper and hope that authors will be able to deal with the few revisions that are still needed (and in particular to check out that Qst values are indeed properly computed).

--------------------------------

*REVIEWED BY ANONYMOUS REVIEWER, 2020-04-09 13:31*
The manuscript by Marin et al. has changed greatly after the revision, so that many of the previous comments by reviewers do not apply anymore. In my opinion the new version is a clear improvement, based on what I read in the previous reviews. My review below is thus independent from those.

The paper is an interesting contribution to our understanding of plant adaptive variation (or lack thereof) along elevation gradients and of how recently-diverged taxa could be responding differently to similar environmental pressures. The authors use an Fst/Gst approach, based on a common garden and previous field studies, to study patterns of local adaptation of two subspecies of snapdragons. The approach is sensible in cases like this where reciprocal transplants are difficult. The paper is overall clearly written but there are some aspects of the analysis and results that need clarification:

- Values of narrow sense heritabilities are estimated in the greenhouse across all populations within each subspecies, rather than for each population. Please justify why this done, given that all subsequent analyses are based on population comparisons, and that we know that h2 values are population-dependent. The authors mention in line 234 that caution must be taken with h2 values, but it is not clear how they are being cautious in the analyses that follow.
We modified the text LL248-252 to clarify this point. In brief, full sib designs provide less precise heritability estimates than full pedigree based estimates. To increase precision, we traded off the detailed information at the level of populations and produced only one global estimate. We then took into account the uncertainty around $h^2$ by providing its 95% CI (calculated by using a parametric bootstrap method). We were cautious not to interpret trait $h^2$ values in terms of quantified evolutionary potential to respond to selection or to compare them between traits in the discussion.

- The text explains that pairwise relationships were tested with Mantel and partial Mantel tests, but legends of Fig 2 and Fig 5 include regression equations and p values – please clarify

what the analysis was.

We removed the equations produced by linear regression from Figures 2 and 5. These analyses were removed from the manuscript after the previous round of reviews. Figure legends should have been updated accordingly. Plot figures were kept to illustrate graphically informal trends in the data that were formally tested by Mantel tests.

- I would expect that some of the traits included in the analysis are correlated to each other, notably the number of nodes, branches and internode length. The two traits that are found to vary with altitude, nodes and branches, appear to strongly co-vary within species in the figures. I suspect they are developmentally strongly correlated and thus just two sides of the same coin – if this is the case, they should be combined for analysis or the Mantel tests corrected for multiple comparisons.

We added text LL410-412 to acknowledge that these two traits are strongly correlated because axillary buds producing branches are located at nodes. We believe that using a correction for multiple comparisons (eg. sequential Bonferroni correction) is not useful here because only one trait (number of nodes) was considered significant. We flagged another trait on the list as being marginally significant. No other trait was found significant. Using a correction for multiple comparisons here such as the sequential Bonferroni correction would therefore give the exact same result.

Other comments: The analysis is focused on comparing two closely related subspecies that, I believe, differ mostly in their floral traits. Have there been any previous studies comparing vegetative traits between the subspecies in wild populations? If not is this a new contribution to this system? If yes, it would be very nice to include in the Methods to provide some background, and to explain how traits vary in the common garden when growing at the same altitude compared to in the wild.

This is indeed to our knowledge a new contribution to this system. We are not aware of studies having measured vegetative traits of *A. majus* natural populations in the wild or in experimental gardens. We therefore cannot compare. As you might already know, research on *A. majus* conducted to date was mainly focused on floral traits.

Along those lines, the main traits that differ between the two taxa are floral traits, yet the current study is focused on vegetative traits. This is fine and interesting but it requires some justification – why exclude reproductive traits if their comparison could also be interesting in this particular system?

We added text in the methods to clarify this point LL212-215. Following the reviewers' rightful request on the first round of peer reviews, we have split the global data set in two, one for each subspecies. As a consequence, we had not enough statistical power to analyze reproductive traits because less data was available (not all plants that grow flower).

Given that the paper is focused on the differential adaptation patterns between the two subspecies, it could be good to add a bit more about what is known (or unknown) on what keeps the two subspecies separate. I see that there is some explanation on this in the response to the previous round of reviews, so it should be relatively straightforward to include a justification for it here. It seems important to rule out that similar mechanisms could be affecting vegetative traits as well.

We added text LL151-153 to clarify that frequency dependent selection based on flower colour maintains *A. majus* subspecies separate. It would however be speculative "to rule out that similar mechanisms could be affecting vegetative traits" in the absence of a dedicated

study and in the presence of evidence [from our study] for vegetative differentiation between subspecies [that is by definition correlated to flower colour].

Line 161-162: unclear, please be more explicit
We added text L172 to clarify that as a result of sampling populations on different valley and summits, "populations with similar elevation are not geographically closer."

Minor comments:
Please carefully check the manuscript for style and English.
Done

Here are just a few examples:
- repeated use of the Word "potential" in the Abstract
This was modified

word choice:
-"combined" in line 48;
changed to "combined the information from"

- "participates" in line 119;
changed to "contributes to"

- "all traits" in line 340
 removed

- "comforted" in line 369
changed by "support"

- Fig S1: please use the colour codes used in other figures to show the two subspecies
Done

- Line 165: "across the entire spatial distribution" is unclear
removed

- Lien 174: sentence on maternal effects: rephrase for clarity .
Done, rephrased to: "This intermediate generation of plants grown in controlled conditions allowed us to reduce potential maternal environmental effects that could have otherwise affected plants grown from seeds sampled in the wild." LL184-185

- Line 192: some of this information on the traits is already in the Introduction so seems redundant; please explain instead how/when traits were measured.
Done, we now presented how/when traits were measured.

- Line 221: is the "plastic container" the randomized block?
Yes indeed, we added "block" into brackets.

- Line 272: Mantel tests – specify that they were run separately for each subspecies.
Done

- Table 1 legend refers to analysis in JMP?

Every analysis was done in R on this new version of the manuscript, the legend needed to be updated. We corrected this.

- Paragraph starting in line 339 needs revision; there is some repletion and awkward phrasing
Done, rephrased to "Mantel tests showed no relationship between population pairwise $Q_{ST}$ and $F_{ST}$ for most of traits (Table 2). Only population pairwise $Q_{ST}$ for the germination date in *A. m. striatum* was significantly correlated with population pairwise $F_{ST}$." LL368-370


- Please check figure numbers across manuscript
Done