Answer to the reviewers

Round 2

Hugo Gruson
Marianne Elias
Juan L. Parra
Christine Andraud
Serge Berthier
Claire Doutrelant
Doris Gomez

Created 2019-10-07

We thank again the recommender and the two anonymous reviewers for their comments which helped us improve the quality and the clarity of our manuscript!

Reviewer 1

> I have carefully read the revised version of the manuscript "Distribution of iridescent colours in hummingbird communities results from the interplay between selection for camouflage and communication". I provided comments on an earlier version, and although I see notable improvements, I still have some serious concerns about the approach taken and the conclusions that are drawn from the results. I hope the authors find these useful.

Thank you for your comments. Please see our detailed answers and changes below.

> Firstly, measurements from one single individual per species seriously reduce the confidence one has on the degree to which the color measures will be representative of the species. Furthermore, it is not clear from the Methods whether the selected individual was from the communities which are being compared, as there is no way to discard - at least minor - geographic variation in coloration along the specie's range.

This is a very good remark. We indeed did not mention in the manuscript that whenever possible (most of the time), we picked specimens that were collected in Ecuador (88% of the cases), or when not available neighbouring countries, such as Colombia or North Peru (11% of the cases). So this possible source of bias has been dealt with as much as possible in our sampling design.

> Comparative analyses across a number of species rely on precise and representative estimates of species' average trait values, and this becomes seriously undermined when one individual per species is measured. The author's argument that they selected an individual whose coloration was representative of other specimens available in the collection is undermined by another argument, in the ms itself, in favour of the use of spectrophotometry for color measurement as being superior to human visual estimates. Skeptical readers would be allayed if estimates of the amount of within vs among species variation were provided.

We have now added the comment in our previous response in the manuscript itself, which should help alleviate doubts of skeptical readers. To reiterate it here: "Previous studies show that even low sampling per species (as low as one individual/species in some cases) can accurately capture colour characteristics of the species (Dalrymple et al. 2015). Additionally, preliminary analyses on an independent dataset of 834 points across 18 hummingbird species, with up to 5 individuals measured by species, showed that intraspecific coefficient of variation (standard deviation divided by the average) of hue is very low (1.69%) but could be higher for brightness (23.18%) (detailed values for each species in table S3). When comparing intra- to interspecific variation, intraspecific however always remains negligible compared to interspecific variation (intraclass coefficient reported in table S3)."

Additionally, while we may indeed underestimate the total amount of variation, this should have little influence on our results because the indices we use in this study do not depend on total variation, but compare intra- and intercommunity variation, which should be impacted roughly in the same fashion by our possible undersampling.

> " Secondly, I am concerned about the null models used to determine over dispersion vs clustering. The null models represent community compositions based on random sampling from the whole set of species included in the study. This is fine. However, when looking at the trait distribution along the phylogeny it is clear that there are quite notable differences in the degree of color variation among the patches that were measured. The back, for example, is almost always green(ish) to green-brown, with the odd cream coloured or purple species (two in the whole set), while the throat coloration (as is well known for hummingbirds) spans a much wider gamut of colors. What this means is that at least phenotypically, it is much more difficult (not to say virtually impossible) for over-dispersion to be detected for the back as there is virtually no variation in coloration in that patch in the set of studied species. On the other hand, it is similarly difficult to detect clumping for the coloration of the throat as there is a very large amount of variation. In other words, given the large differences in phenotypic variation between patches, the null models appear to "stack the deck" in favour of a particular hypothesis. "

We have carefully considered your remark and we firmly believe this is not an issue, and that we are not "stacking the deck" in favour of our predictions. The indices we use (τ_{ST}) compare intra- vs intercommunity variation. The total amount of variation has no influence on the value of this index. We have also edited the manuscript to clarify the fact that null models are built using colour data from the total pool of hummingbirds, *on a given patch*. The following cartoons show how it is possible to get either clustering (lefthand side of the figure) overdispersion (righthand side of the figure) with both low (low total amoung of variation) and high total variation (high total amount of variation).

> " Thirdly, assuming we trust the results, these provide information about a pattern: clustering or over-dispersion of species in a community, based on phylogeny or for a given patch-coloration. However, the authors make a leap-of-faith to extrapolate a process from the observed patterns, that clustering suggests that similar coloration results from selection to avoid predators, while over-dispersion occurs in patches that play a role in species recognition. These are but two of potentially many different explanations for the observed patterns. Revell et al. (2008) have warned against interpreting evolutionary processes responsible for observed presence or absence of phylogenetic signal in a given trait, and their arguments apply as well for cases such as this one. The results provide information on a pattern, the evolutionary process that has given rise to said pattern cannot be discerned, at least not with the data that is available to the authors in the present manuscript. For example, no evidence is provided for the role of different patches in camouflage from predators or intraspecific recognition. "

You are correct that we cannot know for sure which evolutionary processes caused the patterns we are observing now. To address this point, we have revised the phrasing in several places of our discussion to clarify that selection for camouflage or for species recognition should produce patterns that we observe. This combined with the theoretical background layed out in the introduction make them likely candidates but we cannot rule out other explanations with 100% certainty.

Additionally, as advised by reviewer 2, we now mention in the main text that the back and the wing are respectively green and brown in almost every species, which corresponds to what we would expect as a result of selection for camouflage.

Reviewer 2

> " I have one minor further note about the implementation of one of my suggestions. The conclusion that character displacement acts on certain patches while other patches are under selection for crypsis is well supported by the addition of supplementary figure 4, which shows the hue of the 8 main patches, and on line 325, you say "co-occurring hummingbird species tend to display the same hues on dorsal patches." It is clear from supplementary figure 4 that back and wing colors are conserved as brown and green specifically across the phylogeny, but you do not explicitly say so in the text of the manuscript and
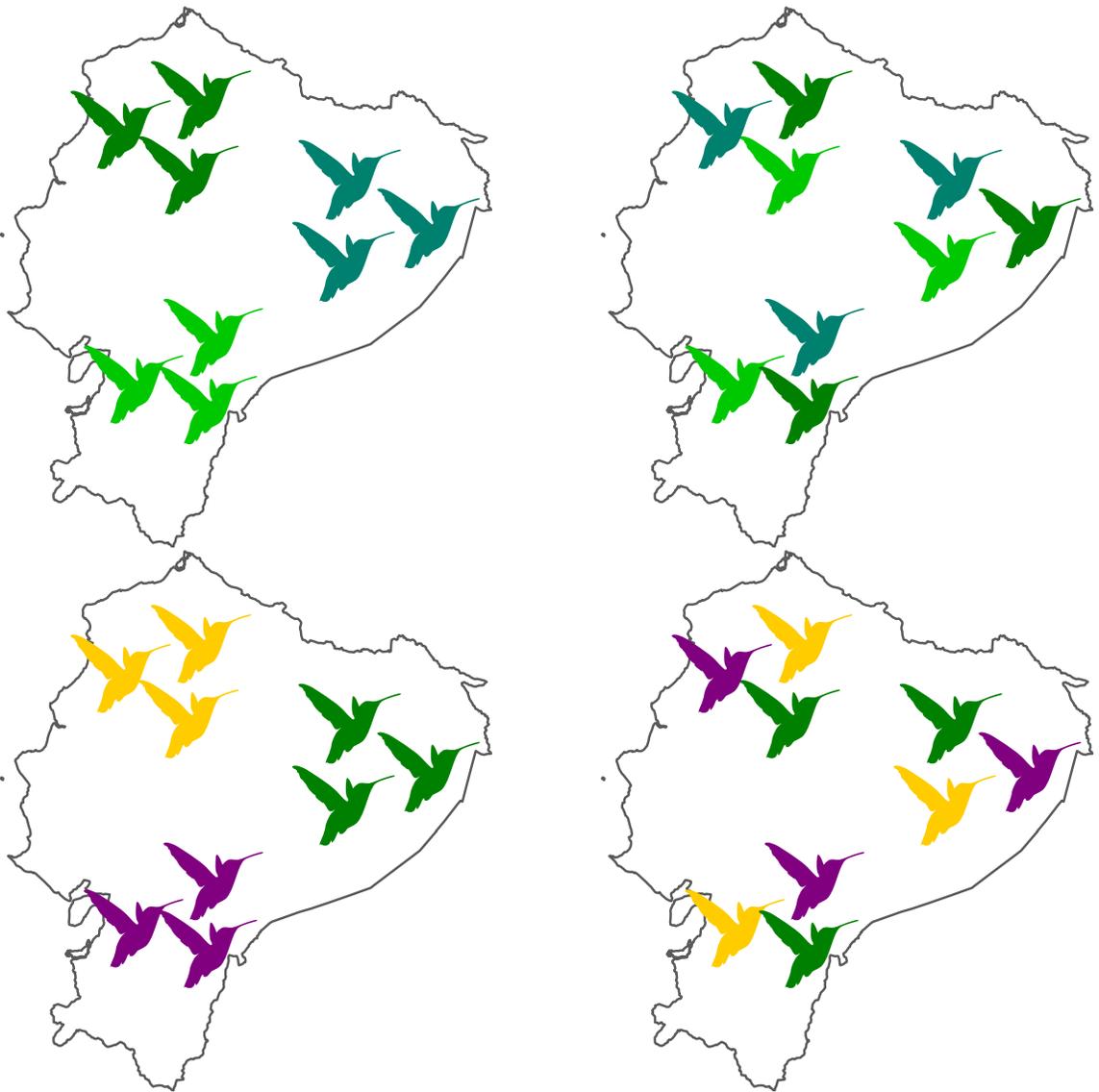
Figure 1: The two top figures have the same (low) total amount of variation but the left one shows a clustering pattern and the right one shows an overdispersion pattern. Similarly, the top bottom figures have the same (high) total amount of variation but the left one shows a clustering pattern and the right one shows an overdispersion pattern.

> the authors do not refer to figure S4 in the manuscript body at all. I think that these specific colors are key support for the crypsis hypothesis and including this information in the main text might strengthen your argument.

Thank you, this is a good idea and we now explicitly mention this, along with a reference to figure S4.

> Otherwise, the authors have fully addressed my previous recommendations as well as those of the other reviewer. I think that the changes to the hummingbird outline figure render the clear, readable, and compelling. Including the hypotheses and predictions as a main body table has certainly improved the readability and flow of the manuscript. I would therefore recommend this paper pending further revisions from other referees.

Thank you!

References

Dalrymple, Rhiannon L., Francis K. C. Hui, Habacuc Flores-Moreno, Darrell J. Kemp, and Angela T. Moles. 2015. "Roses Are Red, Violets Are Blue – so How Much Replication Should You Do? An Assessment of Variation in the Colour of Flowers and Birds." *Biological Journal of the Linnean Society* 114 (1): 69–81. https://doi.org/10.1111/bij.12402.