Round #1

Decision

by Guillaume Achaz, 2018-07-24 19:28
Manuscript: https://doi.org/10.1101/306456

Recommendation, pending minor revisions

Dear authors,

Two external reviewers and myself have carrefully read your ms entitled "Separate the wheat from the chaff: genomic analysis of local adaptation in the red coral Corallium rubrum" and we all agree that this ms is sound, interesting and certainly deserves to be recommded by PCI Evol Biol. As pointed out by the reviewers, the main strength of this ms is the indepth discussion of the pros and cons of the methods and results you provide. I agree with their opinion.

We nonetheless have few suggestions that could potentially improve the overall quality of this ms. I believe the suggested corrections will be easily addressed and that I will be able to recommend the revised version without any further opinion from the two external reviewers. Please revise and respond to all comments (see below) before I can recommend this article.

Best regards,

Guillaume Achaz

:: Suggestions from G Achaz ::

Major: (A) can the authors provide basic analysis of within samples diversity (sample size, #sites, paiwise difference, Tajima's D, etc).

Done : we completed table 4 with samples sizes, nucleotide diversities and mean number of private alleles. As we focused at the SNP level we did not compute Tajima’s D.

(B) did the authors try to infer model parameter (using a mix of demography and structure) using appropriate methods like the ones implemented in dadi (or any other equivalent approach).

We did not try to infer model parameters, apart from effective size alone, but the results were inconclusive. We agree that it would be very useful to infer gene flow along with effective size. Nevertheless we did not have time to do it there ;t we will work on this question in another article.

(C) please shorten and/or lower the importance of the detection of selection as there isn't obvious positive conclusions from it.

We shortened the part dedicated to this topic in the discussion, from 42 to 21 lines (now lines 451-472).

Minor: l43 "between" is repeated in tandem l70 replace "high" by "strong"?

Done.

l189, l270, table4 what is gene diversity? pairwise differences (i.e. heterozygosity)?

Here gene diversity corresponds to expected heterozygosity (as in Arlequin’s manual), we now explain this.

l194 I don't understand what is the meaning here of 'centered'

The allele frequencies have been centered, meaning that they have been modified to a mean of zero. This is part of the adegenet R package and recommended for multivariate analyses by Jombart et al., 2009. We added an explanation on this point.

l282-283, l294, l345 this seems surprising. Why would the sex contribute to genetic differentiation? Am I missing something obvious?

Regarding sex, the "sex loci" did not contribute to genetic differentiation between populations. Nevertheless as the PCAs are based on differences between individuals, these loci can separate individuals according to sex. More precisely we demonstrated that there is an XX - XY genetic sex determinism in this species ; this is detailed in Pratlong et al. (2017). In this article, we identified some loci which were fixed in almost all females, and were at the heterozygous state in almost all males, and these loci were important contributors to the PCA "sex axis". We added a sentence to explain this.

l309 you mean 20 instead of 10?

No, the depth difference is 10 m for the two sites of BANN, and also those of BANS. We now indicate this in the sentence.

l357 which axis are you refering to? pc1?

No, this is the second axis, which is linked with depth differences ; we precise this.

l384 I would be more cautious as this pattern is not visible in the Fst analysis.

We agree with this. We deleted this part which would require more analyses.

Reviews

Reviewed by anonymous reviewer, 2018-05-16 18:37

This manuscript describes the population genetic analysis of the red coral Corallium rubrum, using RAD-Seq data from three populations in the mediterranean see, and samples at two depth. In order to describe the genetic structure of the species, the authors used a combination of principal components analysis (PCA), bayesian inference (STRUCTURE), analysis of molecular variance (AMOVA), and measures of fixation idenx (Fst). The authors find that geographical location is the main factor explaining the genetic variance, depth showing an effect only in one population. The authors conduct a careful discussion about potential confounding factors, in particular regarding outlier genes with high Fst values.

This is a well-conducted study with an interesting discussion, both with methodological and biological implications, notably with respect to adaptation in the context of climate change.

I only have minor comments:

- l276: "principal components reflected the geographical and depth origin"... As discussed later in this paragraph, the depth only appear to have an effect in one population. It would be more accurate to write something like "principal components reflected the geographical and, to some extant, depth origin"

We modified this sentence : "The positioning of individuals with respect to the first two principal components reflected the geographical, and, partly, the depth origin of the individuals"

- l58: as written, the authors seem to oppose natural selection and gene flow, while the two phenomena are not at the same level... Just like mutations, gene flow can be neutral or selected. Maybe rephrase?

Done : "The evolution of adaptive processes depends on the interaction between different processes, including local selection and gene flow."

- l235-236: a threshold of 1% was used for the chi2 test and then corrected with a FDR or 5%? This is unusual... Either a threshold on the FDR was used, or p-values were adjusted for multiple testing, but one does not "correct" p-values using a threshold.

We only used the FDR approach ; we corrected the text accordingly.

- l271, Wilcoxon test: not clear what was compared to what: Marseille vc. pooled Corsica + Banyuls? A Kruskal-Wallis test + posthoc would be more appropriate here (see "agricolae" package in R for non-parametric posthoc tests)

We initially performed three tests : Marseille vs Corsica, Marseille vs Banyuls, and Corsica vs Banyuls. We now use a Kruskal-Wallis test and we corrected the text accordingly (see lines 287-290).

- l283: Reference is a bit unclear. Do the author mean something like "The 5th axis of the PCA separated all individuals according to their sex, independently of their geographical origin, as observed in a previous study (Pratlong et al. 2017)" ?

Yes, we modified the sentence accordingly.

- l298-299: two times "the results"

Corrected

- l309, "even for populations separated by 10m": unclear which populations are refered to

We precise this point : "even for the Banyuls and Galeria populations separated by 10 m depth"

- l324: it is unclear which test was performed there, as there are two Fst values per population and it seems like the authors made one comparison per population?

We compared here the F_{ST} between shallow samples with the F_{ST} between deep samples in a given site : for example in Marseille : mean FIG8 and ELV12 F_{ST} compared to F_{ST} between MOR40 and MEJ40. There is then one test for each region : Banyuls, Marseille, Corsica. We modified the sentence : " This difference was significant in each of the three regions (p < 1.10-16 with a t-test in all cases). "

- l376, at a genomic SCALE ?

Done

- Figure 1: while the depth of the sampling site is coded in the sample name, it would be ore clear to have it explicitly indicated (for instance using two distinct plotting symbols)

Done

Reviewed by Lucas Gonçalves da Silva, 2018-05-23 14:49

Dear editors and authors,

I would like to congratulate the authors, because the study is extremely interesting, relevant, with well designed, with consistent methods and great results.

I think a very positive point of the work is the discussion around the limitations of the study. The authors assume and describe these limitations, and this is always very important in genomic scale studies.

The only point left to be desired from my point of view was the discussion and characterization of the study in the context of implications for species-focus conservation. This species is considered endangered by IUCN and included within the Mediterranean Red List of Endangered Species. I believe that the isertion of informations about conservation status (introduction and discussion) will help the applicability of work in decision making, and can enrich the work in general.

Done in the introduction : "This species is included in two international conventions (annex III of the Bern Convention and annex V of the European Union Habitats directive), and its harvesting is regulated by national legislations (Ledoux et al., 2013)." We also added a sentence at the end of thce conclusion : "These results would also be useful for the management of this harvested species which is present in two international conservation conventions."

Figures and tables are satisfactory.

L36 - "... we also discuss ..." L42 - "adaptation by species ..." L43 - remove "between" L87 - missing space

Done (but we did not see any missing space on line 87).

I would love to see a niche modeling with this data, applying maxent, maxlike or some other likelihood algorithm for SDM (or more than only one). It is not something to be included in this paper, but I recommend authors to do so, since they have enough data for this. If you are interested, I leave the communication channel open with me via the editor of PCI Ecol Evo. Editors are authorized to send my email contact, if the authors want it.

Thank you for this suggestion. Yes this would be a very interesting subject.

I strongly recommend the acceptance of this paper.

Best Regards.

Dr. Lucas Gonçalves lucas_gonc@yahoo.com.br