Editor comments (Joseph Gillespie)
MINOR REVISIONS
Dear Dr. Cerca and colleagues:

Thanks for revising your manuscript. The reviewers are very satisfied with your revision (as am I). Great! However, there are a couple issues to still address and a few minor edits to make. Please address these ASAP so we may move towards acceptance of your work.

Best,

-joe

Dear Joe,

Thank you for your encouraging words. We have now changed the manuscript to meet with the reviewer's comments. This specifically includes:

1. We have included an AIC assessment of the demographic analysis, as argued by #R3. The re-running of the analysis and likelihood, together with AIC, made us obtain slightly different results. Specifically, the models do not reject ancient admixture. We have accommodated these changes by modifying a paragraph in the discussion.
2. We have included a consensus tree in the supplementary, as suggested by #R3.
3. We have established a github page with all source code for this paper. Please find it here: https://github.com/jcerca/Papers/tree/main/Stygocapitella_PeerJ

We have added a note of gratitude in the acknowledgments to the three reviewers. "We are grateful to Diego Fontaneto, Gerardo Perez-Ponce de Leon and an anonymous reviewer for their comments, suggestions and critiques, which have led to the substantial improvement of this manuscript. "

Below we provide a point-by-point answer to the concerns and suggestions raised by #R3, the only reviewer pointing out issues in this round of review.

Thank you for editing our manuscript.

Reviewer 3 (Anonymous)

Basic reporting
No comment

Experimental design
No comment

Validity of the findings
No comment

Comments for the Author

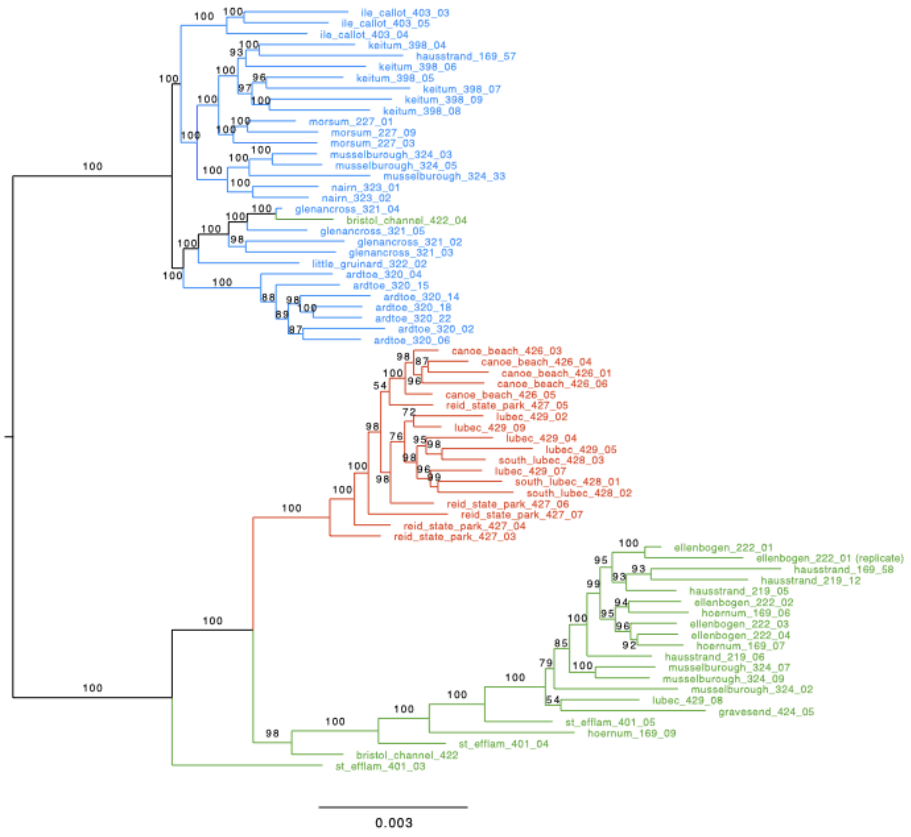
There was considerable improvement to the manuscript in response to reviewers' concerns and I believe it is much better now. That said, the paper could benefit from revision of a few points. Here I start by making some comments that have been apparently overlooked during the first round of reviews and then I make a few suggestions based on the current version of the manuscript.
One of them was a problem already in the first version (model comparison by likelihoods), but the figure in the first version had an unlabelled axis and it was unclear whether this was the case. Now that it is clear, I suggest to use the likelihoods to calculate AIC instead of comparing models by likelihood directly. AIC scores should be comparable across models, but likelihood is not due to the different number of parameters. If anything, I believe this will make the main argument stronger, decreasing support for models including migration in relation to models with migration.

Before pointing out parts to be revised, I would like to commend the authors for the careful revision. While it is clear that great care was taken to address major comments, it seems some of minor issues that I pointed out in the first round were neglected, so I rephrase them here:

We are particularly grateful to #R3 for their attentive read. The reviewer's suggestions have (yet again!) contributed to a substantial improvement of our manuscript. We are also grateful for the kind words of support. Below we answer point-by-point the reviewer's concerns.

1 - I still do not understand why samples are split in A0 and A1 in Figure 2 if alleles are unphased. Whether an allele is marked as "0" or "1" is arbitrary because of the lack of phasing, and therefore grouping them this way makes no sense. I would think that a phylogeny based on the consensus sequence for each individual would be more appropriate.

Thank you for this comment. We have now added this analysis as Suppl. Figure 7:



“**Supplementary Figure 7.** Phylogenomic tree based on 4,737 RADseq loci, where alleles (0 and 1) were turned into a consensus sequence. Bootstrap support is provided for the main branches. Coloration follows species with blue representing Stygocapitella subterranea, green representing S. josemariobrancoi, and orange S. westheidei.”

In the main text we have introduced this analysis by including the following text:

M&M: “Additionally, since RADseq loci (represented by allele 0 and allele 1) are not phased and since the labelling of 0 and 1 are arbitrary, we obtained a consensus sequence for each individual. This was done by running the consambig module included in the EMBOSS pipeline (Rice et al. 2000).”

Results: “Finally, the generated phylogenomic consensus tree shows a similar topology to the that in Figure 2 (Supplementary Figure 7). The three samples causing paraphyly of the lineages in the phylogenomic tree are placed within S. subterranea (Bristol Channel 422 04), as sister to the lineage S. josemariobrancoi and S. westheidei (St. Efflam 401 03), and as the first branch of S. josemariobrancoi (Bristol Channel 422 05).”

2 - "Maps, Google" is still in the references. I would suggest reviewing references more broadly prior to publication. For example, there are two preprints cited and by now they might have been published already.

We have now removed he google maps reference and have added the two published papers, replacing the bioRxiv reference:

“Ferreira MS, Jones MR, Callahan CM, et al (2020) The legacy of recurrent introgression during the radiation of hares. Syst Biol”

“de Medeiros BAS, Farrell BD (2020) Evaluating insect-host interactions as a driver of species divergence in palm flower weevils. Commun Biol. https://doi.org/10.1038/s42003-020-01482-3”

3 - "bad apple" is still mentioned in Supp. Table 2 but not defined in text. In the description of the method to minimize missing data one can guess these are the samples with >45% missing data that have been removed, but since the term "bad apples" is used in the supplement it should also be explicitly defined in the main text (or removed from the supplement).

Thank you for this comment. We have now corrected this inconsistency by removing the bad apple reference. In the summary, it reads:

“**Supplementary Table 2** Specimens used in this study. For each specimen we provide a sampling code, the collection site, a sampling code and the NCBI information for COI, 16S, 18S, ITS1. The column “Present in the final dataset” shows whether the specimen was removed due to >90% missing data, as shown in the final column.”

In addition to those, I have a few comments on this version of the manuscript.

1 - Not sure if I am the one who failed here, but I could not find and download fastsimcoal files that should be in the supplement. Please make sure they are provided in the final version (e. g. together with code) to ensure reproducibility.

Thank you for this comment. The files are available with the review package, and have now been included in the following github page: https://github.com/jcerca/Papers/tree/main/Stygocapitella_PeerJ

2 - While the authors disclosed the mutation rate used in the response to reviewers, it is not mentioned in the text. This is important, since the number of generations estimated and discussed depends on mutation rate. For example, the authors mention that species of Stygocapitella have one generation per year, which is as important as mutation rate to interpret how this relates to actual time scales. I would suggest adding it as a short sentence to methods, not only the supplement.

Thank you for this comment. We have now updated the text on the manuscript:

“When included in the model, gene flow was moduled as asymmetric. Each model was run 10,000 times with an assumed mutation rate of 1.2e-8, and the best fitting scenario was evaluated using likelihood, by running it 100 times.”

3 - While in the response authors mention not using a minimum allele frequency criterion for estimation of Tajima's D, it is not clear in the text when a maf was used or not. The only time maf is mentioned is in line 216, and it seems implicit that this criterion applies to all

downstream analyses. If this is not the case (as seems to be from the response to reviewer's comments), it should be stated explicitly.

We have now made it more explicit that we did not account for minimum allele frequency in this analysis. In specific, on line 258 it reads (addition in bold):
"To gauge population-level patterns and diversity, we selected loci from the all-sites dataset without missing-data at the population-level and estimated summary statistics including nucleotide diversity (π), Waterson's estimator of genetic diversity (S) and Tajima's D using DNAsp v6 (Rozas et al. 2017). T**his dataset was not pruned for minimum allele frequency.** "

4 - Now that it is clear that the y axis in Figure 6 is likelihood, this reveals a problem. Model comparison in a likelihood framework needs to take into account the number of parameters, since a simpler model is preferable to a more complex one if they have the same likelihood. There are different ways of accounting for number of parameters, but the Akaike Information Criteria (AIC) is one of them and widely used in the context of fastsimcoal. AIC = 2*k - 2*L, where k is the number of free parameters (i. e. the number of parameters in the *.est file of fastsimcoal) and L is the log-likelihood. After calculating the AIC for each model, these can be directly compared and the best model should have the lower AIC. There are several tutorials available, I found one of them here, for example:https://speciationgenomics.github.io/fastsimcoal2/
I suspect that after comparing models by the AIC the model with no migration will be favored in relation to others, making the argument in the paper stronger.

Thank you for this comment and for sending us the tutorial which was co-written by one of the co-authors of this manuscript (Ravinet). We have now re-ran the analysis including a likelihood and an AIC. The AIC assessment is included as Supplementary Figure 8, and the likelihood assessment as Figure 6. The new analysis led to slightly different results:
From the four most supported scenarios in the novel analysis, the "no gene flow scenario" displays very recent times of coalescence, which may indicate support for ancient admixture (when running scenarios with no gene flow in an empirical data set with evidence of admixture, times of coalescence times become super recent). The three remaining scenarios indicate ancient admixture. We have re-written a paragraph of the discussion, to allude to the possibility of ancient admixture, but that this is not discernible from incomplete lineage sorting. In essence, we did not alter the conclusions, since we argue that the amount of data (~4,000 SNPs) does not allow us to discern clearly between both patterns.

Paragraph on the discussion "We find clear evidence for shared genetic variation in Stygocapitella. The most conspicuous evidence for this comes from the admixture analysis, which clearly demonstrates admixed populations in the three species (Figure 5). This evidence is further supported by individuals with intermediate positions in the MDS – a test which is robust to missing data (Figure 4). However, several evidences do not support a preponderant role of recent admixture. First, we obtained no evidence for admixture when using F-statistics, since we find only positive F-values (Table 1). Second, contrary to the expectation of ongoing gene flow, we do not observe higher levels of heterozygosity in sympatric populations (Lubec in the USA, Musselburgh in Scotland, Hausstrand in Germany; Table 3) where individuals of different species are found in the same sediment sample in close proximity (volume ranging from 50-500 cm3). Third, admixture often generates

incongruence between mitochondrial and nuclear markers (Melo-Ferreira et al. 2012; Sloan et al. 2017), which is not seen in single-marker trees (Supplementary Figures 2-5). Fourth, models with exclusive recent admixture are generally poorly supported by the demographic analysis (Figure 6). In contrast, three out of the four most supported demographic scenarios suggest ancient admixture, and one supported no gene flow at all (Figure 6). The scenario with no gene flow inferred coalescent times of 451 and 13,834 generations or years (1 generation is expected to be 1 year, Günter Purschke pers. comm)] which are not compatible with estimates of the splitting age of the three Stygocapitella species (~5-30 million years ago; Cerca et al. 2020b). Given that reduced times of coalescence are a typical signature of simulations that do not account for gene flow, when it has occurred in empirical data (Leaché et al. 2019), it is likely that incomplete lineage sorting alone cannot explain the patterns of shared variation among Stygocapitella species. In other words, the demographic analysis supports a scenario that includes ancient admixture. The three scenarios with ancient admixture vary in the presence or absence of admixture after the second coalescence event (S. subterranea and S. josemariobrancoi): in one scenario, admixture is exclusive to the ancestral branch; in the remaining two, gene flow between S. josemariobrancoi and either S. westheidei or S. subterranea occur. Given the lack of support for on-going gene flow between species by the FST, summary statistics, and F-statistics (Tables 1-3, Supplementary Figures 1-5), admixture may have occurred immediately before or after the speciation event of S. westheidei or S. subterranea, but not in recent times (i.e. the last generations). Furthermore, the occurrence of ancient admixture can affect the inference of recent admixture when not take the phylogeny into account (Malinsky et al. 2018; Ferreira et al. 2020), this may explain the incongruence between some of our analysis. Therefore, while the demographic analysis suggests the occurrence of admixture among S. josemariobrancoi and the other species, future studies are necessary to confidently dissect and determine the role of recent gene flow in the system with independent analyses. For example, these studies will benefit from using whole-genome data to determine whether interspecific divergence in regions of the genome show gene-species tree discordance, thereby dissecting ILS and recent hybridization (Joly et al. 2009; Giska et al. 2019). Also, the demographic analysis favouring a preponderant role of ancient admixture does not exclude the occurrence of ILS, and the beforementioned approach would also allow to clarify the relative contribution of ILS and gene flow to shared patterns of variation among species. In sum, to the extent that we can speculate, our data suggests that shared genetic variance is more likely explained by an evolutionary history including incomplete lineage sorting and ancient geneflow.”

Below we copy figure 6 and supplementary figure 8:

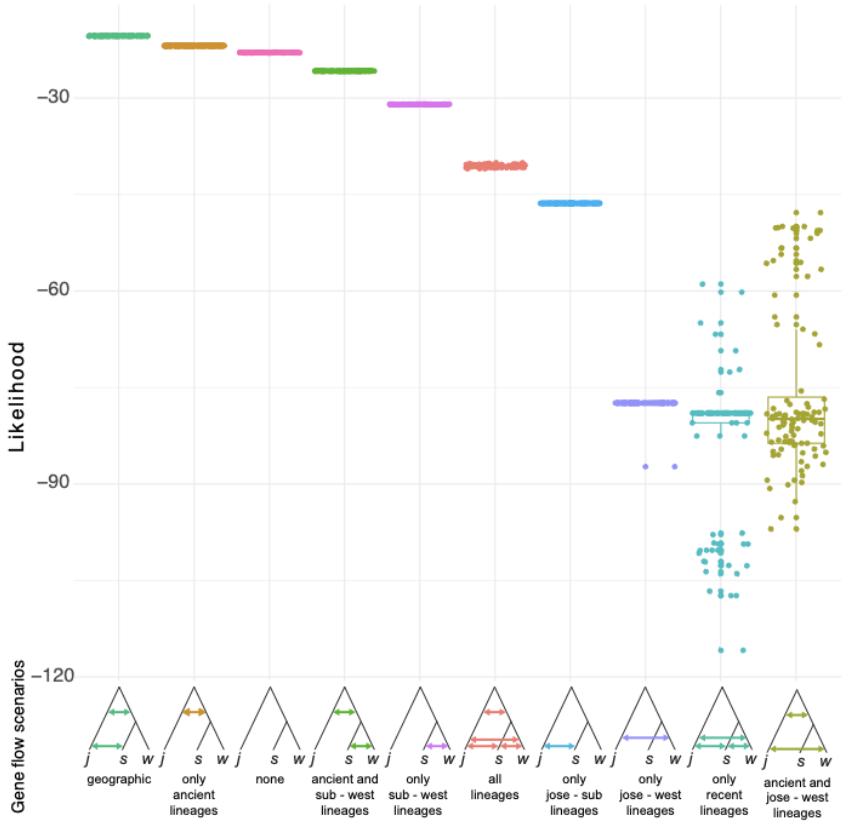
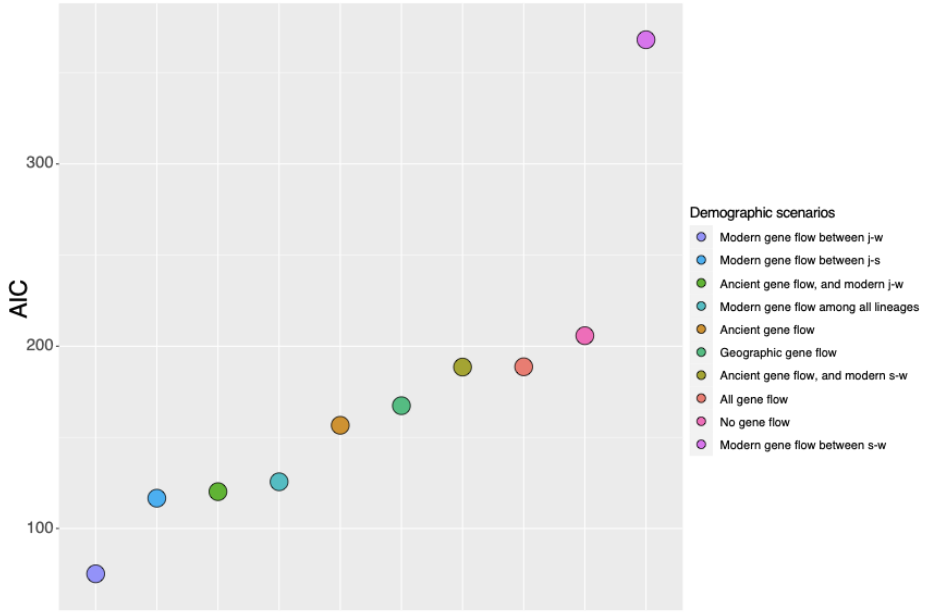


Figure 6: Demographic scenarios considered. The likelihood of different demographic scenarios is displayed on the Y axis. Based on the estimated phylogeny (Figure 2), we modelled scenarios for (from left to right): 1) geographic gene flow (gene flowbetween S. josemariobrancoi and the ancient lineage, and S. josemariobrancoi and S. subterranea); 2) ancient gene flow (gene flow between S. josemariobrancoi and the lineage before the S. subterranea and S. westheidei split); 3) no gene flow at all; 4) ancient gene flow and gene flow between S. subterranea and S. westheidei; 5) gene flow only between S. westheidei and S. subterranea; 6) gene flow in every possible branch; 7) gene flow in sympatric, European linages; 8) gene flow between S. josemariobrancoi and S. westheidei; 9) gene flow between currently existing lineages; 10) ancient gene flow with gene flow between S. josemariobrancoi and S. westheidei.



Supplementary Figure 8. AIC-evaluation of the demographic scenarios. Different models (see main text for details) are depicted in the X axis and have different colours, AIC values are given in the Y axis. Species names are reduced with 's' representing *Stygocapitella subterranea*, 'i' representing *Stygocapitella josemariobrancoi*, and 'w' representing *Stygocapitella westheidei*.

MM 268-275: "Finally, we evaluated various demographic scenarios using fastsimcoal2, using the same dataset for the previous analysis which included running fastsimcoal2 (Excoffier et al. 2013). Fastsimcoal2 uses the site-frequency spectrum (SFS) and a coalescent-simulation framework based on an arbitrary user-defined scenario to infer population sizes, strength of gene flow and times of coalescence. To assess these models we calculated AIC and likelihood. Likelihood is calculated by running the 'best parameters' for each specified scenario multiple times and obtaining the distribution of likelihood estimates. AIC was calculated using a script available in https://speciationgenomics.github.io/fastsimcoal2/. "

Results 420-423: "Finally, while the AIC assessment provide slightly different results (Supplementary Figure 8), the second and third most supported scenarios are the geographic gene flow and ancient gene flow; being thus in agreement with the likelihood results."

5 - line 280, "module as asymmetric" should be "modelled as asymmetric".

6 - There are weird characters following numbers in the paragraph starting in line 294

All changed accordingly.

On Behalf of all co-authors

José Cerca