



Dr. Joe Gillespie
Academic Editor, *PeerJ*
Department of Microbiology and Immunology
University of Maryland School of Medicine
Baltimore, Maryland, USA

Re: revision of *PeerJ* manuscript # 2021:01:57327:2 (Garrick et al.)

Dear Dr. Gillespie,

Thank you for the opportunity to submit a second revision of our paper. Here we provide a point-by-point description of how reviewers' comments have been addressed. In some instances, we have a different point of view to one of the reviewers regarding how our paper should have been framed, and what is important or worth highlighting. Accordingly, we give some more rationale for why we did what we did. As before, a version of the main text showing all edits as "track changes" has been included. Thank you handling this paper.

Sincerely,

Ryan Garrick, for the authors

Academic editor's comments

>> Two reviewers are very satisfied with your revision (as am I). Great! However, there are some concerns raised by the third reviewer (a few minor by reviewer 2 as well). Please address these and submit a revision ASAP.

Thank you. We have made all suggested minor edits, and made further attempts to clarify our position with respect to lingering concerns of Rev 3, as described below.

Reviewers' comments

Reviewer #1 (Avtzis)

*>> Overall, the manuscript by Garrick and colleagues provides a very thorough and interesting investigation on the intra-population structure of *D. frontalis* across the eastern US, based on the up-to-date approaches in phylogeographic and phylogenetic analyses. For that, I suggest that the manuscript can now be accepted for publication.*

Thank you for the earlier feedback.

Reviewer #2 (Javal)

>> *The authors have responded to all the comments in the previous review. I therefore recommend this article for publication. I have just three minor points to note:*

195-109: this paragraph might have been better placed in the discussion, in the "reconciliation with previous work" section for instance. It would make the introduction shorter, and easier to read. This is only a suggestion, however.

We appreciate this suggestion, but we have opted to leave this information in the Introduction, as we feel that it gives the reader important background information that is directly relevant to our stated goal of this study (see Abstract). Briefly, we wanted to systematically assess several plausible alternative explanations for apparent conflict between two previous studies re: the existence vs. absence of spatial-genetic structure in the southern pine beetle. This section of text provides details on the near equivalency between those studies in terms of some of the most obvious potential reasons for the discrepant findings (i.e., underlying genetic marker set; extent of geographic sampling; sample size adequacy). We think that stating this early, in the Introduction, provides key rationale for focusing on other more nuanced explanations (i.e., information content of loci, null alleles, homoplasy, etc.). This section of text also provides rationale as to why we should care about the discrepant outcomes in terms of the different types of on-the-ground management interventions that would be appropriate under each scenario.

>> *1153-156: it is not clear who sampled the specimens and when.*

Coinciding with submission of the first revised version of this paper, the *PeerJ* journal office asked us for clarification on a similar topic> Based on our response, they have now added the following text into the field permit statement:

"All beetle samples were collected with permission from the U.S. Department of Agriculture Forest Service. Trapping was done as part of the U.S. Forest Service Southern Pine Beetle Prevention and Restoration Program (see: <http://southernforesthealth.net/insects/southern-pine-beetle/southern-pine-beetle-prevention-and-restoration-program>). Each of the collectors were employees of the entities from which the samples were collected, so therefore had permission to collect at those sites: Site; Collector Sicily Island (U.S. National Forest); Jim Meeker, USDA Forest Service Homochitto (U.S. National Forest); JoAnne Barrett, USDA Forest Service Holly Springs (U.S. National Forest); Jim Meeker, USDA Forest Service Tombigbee (U.S. National Forest); Jim Meeker, USDA Forest Service Talladega (U.S. National Forest); Larry Spivey, USDA Forest Service Woolford (Maryland Little Choptank River Sactuary); Heather Disque, Maryland Department of Agriculture) Goat Hill (Pennsylvania Barrens Preserve); Gina Peters and Paul Smith (Pennsylvania Department of Environmental Protection) Warwick (U.S. National Forest); Chip Bates, USDA Forest Service Ponte Vedra (Florida State Land); Jiri Hulcr, University of Florida"

We have also edited the main text in this latest revision to point the reader to this statement, and clarified the collection dates. Taken together, we think this has now addressed the reviewer's concern.

>> *1224: check grammar*

Thank you. We have edited that sentence to improve readability.

Reviewer #3 (James)

>> While some of my previous comments were taken into consideration, my impression is that they were done so only superficially. The focus remains highly technical in nature and still does not adequately address the questions at hand. Greater clarity and an improved framing of the study is needed. Doing the same analysis with just more markers seems like a poor justification to undertake this work, especially when the value of these markers is put into question given the framing regarding previous contradictory results. Similarly, highlighting the contribution of this re-examination as a “roadmap” (line 140) seems strange, as eliminating potential alternate explanations for observed results is just good science.

Our research questions were indeed technical in nature. That was our goal and it has not changed, so we do not agree with the premise here. In general, we are not in favor of retroactively rewriting the research questions, as that seems somewhat disingenuous. In our opinion, the rationale and justification for the framing of the paper were sound, and so we would prefer to remain true to our original goals.

To suggest that we merely did the same analyses as in previously published papers, just with more markers, is an inaccurate characterization of this work. As we have stated before, we 1) **expanded the set of analyses** (e.g., DAPC clustering, individual-based IBD were not previously done), 2) **evaluated novel partitions of the data** (e.g., male-only, female-only sets were not previously analyzed), and 3) **systematically assessed potential causes of failure to detect structure inherent to the markers themselves** (i.e., null alleles, allele size homoplasy). Our intention was always to present a thorough study that focused on understanding the extent to which technical artefacts (signal to noise considerations, etc.) could reconciling discrepant results from previous work. We certainly agreed with the suggestion to add value to the paper via a more detailed discussion of ecological context (timing and outbreak status), and so we did that in the previous revision, at what we felt was appropriate depth (more on this below).

We agree that our reference to a “roadmap” was an overstatement, and we have now removed that text.

>> Line 44 – and outbreak dynamics?

We have now edited the text to include this as well.

>> Line 53 – suggest adding your research objective/question here (e.g., statement on line 72) and starting a new paragraph to present life history information.

We have now reorganized and split this paragraph as suggested.

>> Line 75 – development “and application” of molecular tools?

We have now edited the sentence (which was shifted up) to include application as well.

>> Line 100 – what is a “biologically meaningful distance estimate” and how does one distinguish from an un-meaningful one?

It is one that is based on adequate sample sizes such that reasonable precision and accuracy can be expected (cf. an estimated value that is quite meaningless owing to small sample size effects).

Given our phrasing here (“...in both studies, per-site sample sizes were quite large... Thus, it would seem that biologically meaningful genetic distance estimates should have been attainable”), we feel that there is no strong ambiguity here, and therefore chose to leave it as-is.

>> *Line 107 – context and citation needed for reference to rapid evolution*

Agreed. We now specify rapid evolution in novel environments leading to larger population sizes and/or faster expansion speeds, and provided a representative citation for this “evolution first” scenario (i.e., Szűcs et al. 2017 PNAS 114:13501).

>> *Line 113 – Here you refer to the unexpected absences as the motivating problem, but is there any a priori reason to consider the absence of structure to be more likely than the presence of structure? One could also frame this study to examine the unexpected structure identified in the first study. Is it not plausible that technical errors could have resulted in spurious detection of SGS? Why is the absence of structure so much more suspect than its presence?*

This is a good question. Indeed, in the phylogeographic literature, a broadly distributed panmictic population is often invoked as the “starting condition” for subsequent vicariance events that lead to lineage splitting / emergence of finer-scale structure. One problem we see with this notion (and the reason for our inclination to expect spatial genetic structure in SPB *a priori*) is that there are very few empirical examples of species that truly have a widely distributed panmictic population (i.e., spanning a broad range of latitudes and longitudes, and the inherent environmental heterogeneity that this encompasses). While the American eel and European eel (Côté et al. 2013 Mol Ecol 22:1763) are exceptions to this, these species have unusual traits (e.g., catadromous life cycles) that play a large role in maintaining this rare widespread panmixia. Even for highly mobile taxa such as microbes, isolation-by-distance (e.g., among OTUs within a species) is common. Accordingly, we think the most pertinent question is less about “is there IBD” and more about “over what spatial scale is IBD detectable”? We consider this, as well the border version of it that also applies to clusters (i.e., *how* is genetic variation spatially structured, cf. *is* genetic variation spatially structured?) to be a useful *a priori* expectation for our work on SPB. For this reason, we focused on “...explanations for an unexpected absence of clusters or clines...”.

>> *Line 121 – “albeit ephemerally” – do you know this for sure? Suggest remove*

We have tried to strike a balance here, between wanting to highlight the possibility of this re-partitioning of genetic variation being transient vs. completely absent over time, by replacing with “perhaps ephemerally”.

>> *Line 122 – reference needed for this idea*

We have now added citations that draw a comparison with impacts of postglacial expansion into newly available habitats (Hewitt 1996 Biol J Linn Soc 58:247, and Hewitt 2004 Philos Trans R Soc Lond B Biol Sci 359, 183). For completeness, we also noted a counter-example where range expansion *generates* spatial-genetic structure, rather than overwrites it (Excoffier & Ray 2008 TREE 23: 347).

>> *Line 126 – why “notwithstanding ... panmixia”? You refer to panmixia above as a potential alternate hypothesis. You could have both panmixia and weak markers, no?*

Thank you for pointing this out, it was a poor choice of word that obscured our intended meaning. We have edited this to now read “... and acknowledging the potential for genuine panmixia (i.e., a null hypothesis that may be true)...”. We think this makes it clearer that we recognize the real possibility of that a lack detectable

spatial-genetic structure may simply be because there really is none, and that we use this latter scenario as the null hypothesis against which to test alternative explanations.

>> *Line 131 – “only those loci ...” – Precedent for this approach? Reference needed*

In the previous versions of this paper, we had provided a reference for this approach (i.e., Russello et al. 2012 *Evol Appl* 5:39) in the Materials and Methods (L267), but we see that it might be too late there, and so we have added the same at that citation to an earlier location, here in the Introduction, as suggested.

>> *Line 140 – seems strange to highlight the value of new molecular markers. Suggest re-emphasize conceptual (vs technical) contributions.*

In this instance we disagree with the implication that new, carefully validated and characterized molecular markers are of little value. While this may simply be a difference in personal option, we do not consider our take on this to be unusual. Accordingly, we have chosen to retain our original text (except for removing an earlier reference to a “roadmap”, as mentioned above).

>> *Line 185 – the absence of consideration of outbreak dynamics and demography as an explanation for the differences between the two studies suggests to me that my previous comments were only superficially taken into consideration. Although referred to in the discussion, this concept should be woven into the manuscript more thoroughly. Population connectivity and effective population size vary through time in irruptive species, and both of these factors affect population genetic inference. The focus on uSat number and quality misses the mark and does little to actually address your question.*

We feel that our question was pretty clearly stated, technical in nature, and that we directly address it. This criticism seems to be based on the suggestion that we had a different question at the outset (or that we should now change it retroactively), which is incorrect. Given this, we agree that we have missed the mark on intended research question. We believe it reasonable for us to insist that we stay true to the original goals of this paper. The previous revision made a genuine, substantive, attempt to integrate deeper *discussion* of ecological context. Specifically, in our previous response letter, we wrote the following:

“We agree that outbreak status, and the potential for this to have differed across studies at the time of sampling, warrants an expanded treatment. We have added details to the Discussion that more thoroughly explains the consequences of endemic vs. epidemic conditions on dispersal, and we also included a brief comment on seasonal (i.e., intra-year) changes in population dynamics that might also have contributed to apparent differences among studies.”

“... we certainly see a clear connection between our study species and the concepts that apply to the spruce budworm system. Accordingly, we have added a paragraph to the Discussion that highlights the importance of geographic and ecological context, and also flags the more nuanced view of weak spatial genetic structure in cyclical irruptive insect pests...”

Each of these were addressed via stand-alone newly added paragraphs with relevant citations. Together, these constitute ~17% of the entire Discussion section. Accordingly, rather than being superficial, we think this was done at an appropriate depth.