* -------- Пересылаемое сообщение --------  
  От: PLOS ONE <[em@editorialmanager.com](mailto:em@editorialmanager.com)>  
  Дата: 21.12.2020, 18:06  
  Кому: Vadim Mikhailovitch Khaitov <[polydora@rambler.ru](mailto:polydora@rambler.ru)>  
  Тема: PLOS ONE Decision: Revision required [PONE-D-20-30389] - [EMID:3614987825e8a292]

PONE-D-20-30389  
Species identification based on a semi-diagnostic marker: evaluation of a simple conchological test for distinguishing blue mussels Mytilus edulis L. and M. trossulus Gould  
PLOS ONE  
  
Dear Dr. Khaitov,  
  
Thank you for submitting your manuscript to PLOS ONE. After careful consideration, we feel that it has merit but does not fully meet PLOS ONE’s publication criteria as it currently stands. Therefore, we invite you to submit a revised version of the manuscript that addresses the points raised during the review process.  
  
While both reviewer are excited about the general idea behind this study, both are concerned about methodological approaches (allozyme markers that are not fully diagnostic) and lack of discussion of the hybrid status of many mussel populations / species concept in mytilid mussels. I would suggest to follow the reviewers' advice to incorporate a hybrid category in your analysis and, potentially, to re-analyze parts of the data using diagnostic DNA markers, if possible. All underlying sequence data should be deposited in publicly available databases.  
  
Please submit your revised manuscript by Jan 24 2021 11:59PM. If you will need more time than this to complete your revisions, please reply to this message or contact the journal office at plosone@plos.org. When you're ready to submit your revision, log on to https://www.editorialmanager.com/pone/ and select the 'Submissions Needing Revision' folder to locate your manuscript file.  
  
Please include the following items when submitting your revised manuscript:

* + A rebuttal letter that responds to each point raised by the academic editor and reviewer(s). You should upload this letter as a separate file labeled 'Response to Reviewers'.
  + A marked-up copy of your manuscript that highlights changes made to the original version. You should upload this as a separate file labeled 'Revised Manuscript with Track Changes'.
  + An unmarked version of your revised paper without tracked changes. You should upload this as a separate file labeled 'Manuscript'.

If you would like to make changes to your financial disclosure, please include your updated statement in your cover letter. Guidelines for resubmitting your figure files are available below the reviewer comments at the end of this letter.  
  
If applicable, we recommend that you deposit your laboratory protocols in protocols.io to enhance the reproducibility of your results. Protocols.io assigns your protocol its own identifier (DOI) so that it can be cited independently in the future. For instructions see: <http://journals.plos.org/plosone/s/submission-guidelines#loc-laboratory-protocols>  
  
We look forward to receiving your revised manuscript.  
  
Kind regards,  
  
Frank Melzner  
Academic Editor  
PLOS ONE  
  
Journal Requirements:

When submitting your revision, we need you to address these additional requirements.

1. Please ensure that your manuscript meets PLOS ONE's style requirements, including those for file naming. The PLOS ONE style templates can be found at

https://journals.plos.org/plosone/s/file?id=wjVg/PLOSOne\_formatting\_sample\_main\_body.pdf and

https://journals.plos.org/plosone/s/file?id=ba62/PLOSOne\_formatting\_sample\_title\_authors\_affiliations.pdf

2. We note that you have included the phrase “data not shown” in your manuscript. Unfortunately, this does not meet our data sharing requirements. PLOS does not permit references to inaccessible data. We require that authors provide all relevant data within the paper, Supporting Information files, or in an acceptable, public repository. Please add a citation to support this phrase or upload the data that corresponds with these findings to a stable repository (such as Figshare or Dryad) and provide and URLs, DOIs, or accession numbers that may be used to access these data. Or, if the data are not a core part of the research being presented in your study, we ask that you remove the phrase that refers to these data.

3.We note that [Figure(s) 1] in your submission contain map/satellite images which may be copyrighted. All PLOS content is published under the Creative Commons Attribution License (CC BY 4.0), which means that the manuscript, images, and Supporting Information files will be freely available online, and any third party is permitted to access, download, copy, distribute, and use these materials in any way, even commercially, with proper attribution. For these reasons, we cannot publish previously copyrighted maps or satellite images created using proprietary data, such as Google software (Google Maps, Street View, and Earth). For more information, see our copyright guidelines: http://journals.plos.org/plosone/s/licenses-and-copyright.

We require you to either (1) present written permission from the copyright holder to publish these figures specifically under the CC BY 4.0 license, or (2) remove the figures from your submission:

1.    You may seek permission from the original copyright holder of Figure(s) [1] to publish the content specifically under the CC BY 4.0 license.

We recommend that you contact the original copyright holder with the Content Permission Form (http://journals.plos.org/plosone/s/file?id=7c09/content-permission-form.pdf) and the following text:

“I request permission for the open-access journal PLOS ONE to publish XXX under the Creative Commons Attribution License (CCAL) CC BY 4.0 (http://creativecommons.org/licenses/by/4.0/). Please be aware that this license allows unrestricted use and distribution, even commercially, by third parties. Please reply and provide explicit written permission to publish XXX under a CC BY license and complete the attached form.”

Please upload the completed Content Permission Form or other proof of granted permissions as an "Other" file with your submission.

In the figure caption of the copyrighted figure, please include the following text: “Reprinted from [ref] under a CC BY license, with permission from [name of publisher], original copyright [original copyright year].”

2.    If you are unable to obtain permission from the original copyright holder to publish these figures under the CC BY 4.0 license or if the copyright holder’s requirements are incompatible with the CC BY 4.0 license, please either i) remove the figure or ii) supply a replacement figure that complies with the CC BY 4.0 license. Please check copyright information on all replacement figures and update the figure caption with source information. If applicable, please specify in the figure caption text when a figure is similar but not identical to the original image and is therefore for illustrative purposes only.

The following resources for replacing copyrighted map figures may be helpful:

USGS National Map Viewer (public domain): http://viewer.nationalmap.gov/viewer/

The Gateway to Astronaut Photography of Earth (public domain): http://eol.jsc.nasa.gov/sseop/clickmap/

Maps at the CIA (public domain): https://www.cia.gov/library/publications/the-world-factbook/index.html and https://www.cia.gov/library/publications/cia-maps-publications/index.html

NASA Earth Observatory (public domain): http://earthobservatory.nasa.gov/

Landsat: http://landsat.visibleearth.nasa.gov/

USGS EROS (Earth Resources Observatory and Science (EROS) Center) (public domain): http://eros.usgs.gov/#

Natural Earth (public domain): http://www.naturalearthdata.com/

4. Please include captions for your Supporting Information files at the end of your manuscript, and update any in-text citations to match accordingly. Please see our Supporting Information guidelines for more information: http://journals.plos.org/plosone/s/supporting-information.

[Note: HTML markup is below. Please do not edit.]

Dear Dr. Melzner,

Thank you for evaluation of our manuscript. Our responds to reviewers are below.

We are sorry for not providing initial data. It is provided now. What about technical remarks, “data not shown” phrase is eliminated; information that Inkscape 0.92 was used for producing the map is included in the Fig. 1 legend.

Reviewers' comments:  
  
Reviewer's Responses to Questions

**Comments to the Author**  
  
1. Is the manuscript technically sound, and do the data support the conclusions?  
  
The manuscript must describe a technically sound piece of scientific research with data that supports the conclusions. Experiments must have been conducted rigorously, with appropriate controls, replication, and sample sizes. The conclusions must be drawn appropriately based on the data presented.

Reviewer #1: Partly

Reviewer #2: Partly

2. Has the statistical analysis been performed appropriately and rigorously?

Reviewer #1: Yes

Reviewer #2: Yes

3. Have the authors made all data underlying the findings in their manuscript fully available?  
  
The [PLOS Data policy](http://www.plosone.org/static/policies.action#sharing) requires authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception (please refer to the Data Availability Statement in the manuscript PDF file). The data should be provided as part of the manuscript or its supporting information, or deposited to a public repository. For example, in addition to summary statistics, the data points behind means, medians and variance measures should be available. If there are restrictions on publicly sharing data—e.g. participant privacy or use of data from a third party—those must be specified.

Reviewer #1: No

Reviewer #2: Yes

4. Is the manuscript presented in an intelligible fashion and written in standard English?  
  
PLOS ONE does not copyedit accepted manuscripts, so the language in submitted articles must be clear, correct, and unambiguous. Any typographical or grammatical errors should be corrected at revision, so please note any specific errors here.

Reviewer #1: Yes

Reviewer #2: Yes

5. Review Comments to the Author  
  
Please use the space provided to explain your answers to the questions above. You may also include additional comments for the author, including concerns about dual publication, research ethics, or publication ethics. (Please upload your review as an attachment if it exceeds 20,000 characters)

Reviewer #1: This is an interesting manuscript describing a new semi-diagnostic character that could be used to roughly distinguish between the two mussel species Mytilus edulis and Mytilus trossulus in the field without the need for genotyping. I want to commend the authors for collecting this broad sample set. The analyses are in my opinion well done, although I had a few questions and concerns, especially regarding the use of allozyme markers and the lack of available raw data, that should be addressed.

Thank you for careful evaluation of our manuscript.  
  
- The manuscript is a bit long. Sometimes it is hard to follow what was done.

We shortened Discussion and Materials and Methods as far as we could but made the captions to the figures more detailed and clear.

- L94: unambiguous

Ok, corrected

- L191: Mussels in the Baltic Sea represent a hybrid swarm and are not pure M. trossulus. This is explained later in the discussion, but I think this should be recognized already here.

We introduced the sentence: “Note that the Baltic M. trossulus is more strongly introgressed by M. edulis alleles than other populations.”

What we know is that all Atlantic populations of M. trossulus (as well as M. edulis) are introgressed by alien genes, but the Baltic one is introgressed more than any other. In spite of this, the M. trossulus ancestry of Baltic mussels is clearly evident (see genomic data in <https://doi.org/10.1111/mec.13299> and <https://doi.org/10.1111/jeb.13709>). For us is more important that Baltic mussel is M. trossulus than any swarm. You know that this term has different meanings (unimodal hybrid zone; a randomly mating hybrid population that is independent of further outside recruitment) which confuses the untrained reader.

- L198: I did not find a table with the allozyme and SNP genotypes and also no accession numbers for Illumina sequence data for samples that were new to this study. Could you please provide and deposit the raw data for full reproducibility? I think the input STRUCTURE files should be provided as well.

We are sorry for that. Initial data are provided now.

Data Availability: Initial allozyme data and data on individual shell size and morphotype are available from Saint Petersburg State University database (http://hdl.handle.net/11701/22292). Gulf of Maine SNP genotypes are available from <https://datadryad.org/stash/share/N11WUV2APlT_KAg01Lm1857K_6in2M431xsfh5uOJnQ>

All other relevant data are within the paper and its Supporting Information files.

- L203: Why were the allozyme markers used? These markers are also only partially diagnostic and provide less resolution than multilocus DNA markers. It does not make sense to me to try to confirm a semi-diagnostic characteristic with methods that are not fully diagnostic themselves. Then another part of the dataset is genotyped with a genome-wide SNP panel. I think the authors need to give a better explanation / justification of their analyses and a reason why they did not choose a consistent method across samples.

We inserted the next explanation about allozymes (lines … ): “These four loci were involved in the initial diagnosis between *M. edulis* and *M. trossulus* and description of all contact zones under consideration. They individually show 70–95% allele frequency differences between species, and, being less affected by introgression than most of conventionally used PCR-based “diagnostic” markers are the reliable markers for species identification everywhere.

To use 100 000 DNA markers to perform such a simple classification task as distinguishing between M. edulis and M. trossulus (or genotypes dominated be their genes in case of unimodal hybrid zones) is certainly like to use a sledge-hammer to crack a nut. But if we had that data, what's wrong with using it. Many genetic markers are desirable for detailed analysis of hybrid zones but we do not need more fractional classification of genotypes since the test is based on binary marker and our target contact zone is strongly bimodal.

Our opinion about availability of reliable diagnostic markers for hybridizing species of blue mussels is expressed in the Discussion (lines … ). Our assignment is based not on individual markers but by their complex analysis by Structure method.

- L269: What does a few genotyped samples mean? Please specify the sample size for this analysis.

The sentence is modified: “Finally, we tested how well Ptros, P(edu|E) and P(tros|T) could be predicted using formulas Eq 1-3 and the data on the morphotype proportions among species (P(T|tros), P(T|edu)) in a few (minimum two, see below) genotyped samples. We concede that the assumption that sensitivity and specificity do not depend on the prevalence can be violated in the morphotype test, as it often is in clinical tests [36]”.

Would it not be interesting to see if a general model with pools of all samples is possible, using salinity and location as variables? I don’t completely understand why this was not tested. I would think that such an analysis would be useful for researchers that work on regions not included in this study.

Thank you for noticing that. Exactly such model was tested using extensive Kola data (see “Associations between morphotypes and genotypes around the Kola Peninsula”). The data on other geographical regions do not allow for such an analysis due to lack of samples and (or) salinity variation between sampling localities.

While thinking of a general model with pools of all samples from all sets, using salinity and location as variables, please consider that there is pronounced geographical variation in morphotype frequencies in M. trossulus in addition to salinity related variation in Arctic populations of M. edulis. The predictive power of any model “ignoring” region (technically, treating region as a random factor regulating both intercept and slope of the model) would be unsatisfactory low.

Recommendations for researchers that work on regions not included in this study are provided in the Discussion: they have to evaluate the test de-novo for each new contact zone. To aid such research we developed “calculators”.

- L309-323: This seems redundant. Can you merge this section with the one on the Kola Peninsula?

We would like to keep separate analyses for Kola data and for geographical data because different models are employed.

- L601-603: I don’t think this statement is justified. Mytilus species have emerged as a non-model system for hybrid zone research, so for many scientists knowing the exact genotype composition is very important. I would soften this statement.

We suggest the next wording. “The knowledge about the taxonomic structure of populations and a rough classification of individuals into “species” is often (e.g. in community ecology, biomonitoring, aquaculture studies) more valuable to the blue mussel researchers than the precise information about the genetic affinity (e.g. Structure q-value) of any given mussel.”

- L672: links

OK, corrected. Paragraph is shortened.

- L728: But probably there are some published data on salinity for those locations, even though it was not measured in this study?

The St. Lawrence sample was from Baie-Sainte-Catherine, literature data on salinity in this area are inconclusive (De Vernal, A., St-Onge, G. and Gilbert, D., 2011. Oceanography and Quaternary geology of the St. Lawrence Estuary and the Saguenay Fjord. In IOP Conference Series: Earth and Environmental Science (Vol. 14, No. 1, p. 012004). IOP Publishing). Other subarctic sample was collected beside Hvalsey Church in Davis Strait, not far from the glacier. Literature data on salinity conditions in this area are contradictory: saline according to Ribergaard, M.H., 2007. Oceanographic investigations off west greenland. Danish Metrological Institute Centre for Ocean and Ice (DMI), Copenhagen but fresh according to <https://dce2.au.dk/Pub/arcticenvironment/reports/ArcticReport54.pdf>). This information is provided in the S3 Table.

Reviewer #2: Species identification based on a semi-diagnostic marker: evaluation of a simple conchological test for distinguishing blue mussels Mytilus edulis L. and M. trossulus Gould  
  
PONE-D-20-30389  
  
Khaitov et al.  
  
The manuscript is concerned with the diagnosis of two species (Mytilus edulis, M. trossulus) that are difficult to distinguish based on morphological characters for several reason. For instance, there is a general lack of phenotypic characters in these closely related bivalves, there is large phenotypic plasticity and there is extensive interspecific gene flow. The current taxonomic status of both species is derived from multi-trait investigations of mussel shells and genetic analyses (basically allozyme investigations). Additional support comes from the fact that both species are show ecological differentiation. Despite of ecological and phenotypic differentiation, they are considered part of the Mytilus edulis species complex together with M. galloprovincialis. Given the great importance of mytilid mussels for marine science (they represent up to 90% of the benthic biomass and hence shape marine ecosystems), an easy-to-apply diagnostic marker would be more than desirable for research. This is the justification of the study presented by Khaitov et al.! The authors follow previous investigations in the White Sea and test the “dark stripe” as a semi-diagnostic shell morphological character. There aim is to investigate, whether this trait is diagnostic in all marine environments. The authors apply a mathematical approach that is common on medicine where semi-diagnostic characters are commonly used for diagnosis. Both, the scientific question and the innovative approach makes the study interesting for (marine) taxonomists and ecologists and will finally guaranty the publication. However, I have serious concerns that the manuscript in its current form is already publishable! I have various questions and recommendations to the authors that need to be addressed before I can recommend the publication of this manuscript.

Thank you very much for careful evaluation of our manuscript.

Species concept  
My first question concerns the species concept behind the study. The authors use either allozymes or DNA markers for a priori species identification. To my understanding, this is used to calculate the Ptros-parameter. Although many of the populations are situated in hybrid zones, the individuals are either classified as M. edulis or as M. trossulus. For instance, STRUCTURE-q values (which represent hybrid indices) are used for a priori species diagnosis and the threshold is 0.5 (lines 214 to 215). This means, the authors classify F1 hybrid like individuals and early-backcross-generations as one or the other species. Furthermore, species diagnostic allozymes may mask the extend of genetic admixture at other genomic loci, e.g. in the Baltic. Consequently, classification into two categories that incorporate ANY hybrid status seems not justified. I had expected classification base on other admixture thresholds into at least three categories (including at least one hybrid category) and with reference to a valid species concept (e.g., referring to genetic characters). To avoid any misunderstanding: I agree that small proportions of introgression do justify allocation to one or the other species; I disagree that F1 like genotypes (including backcross genotypes) are not treated as hybrids but as one or the other species. In this context, what is said the paragraph about the Baltic and Norway (lines 669-687) could be discussed in another light if the hybrid category is considered. This paragraph already suggests that the pronounced hybrid character of these populations make the approach used in this manuscript doubtful.

Following your advice, we introduced analyses of hybrid zones and of associations between morphotypes and q-values (“Genotypic structure of samples” section of Results). (Although much of the data has already been published and analysed, the inclusion of new samples and visualisation methods allows reanalyzes.) We also included additional reminders about method of classification into M. edulis or as M. trossulus in the Discussion (lines … and … of the revisited manuscript).

Behind the brackets: you rice here fundamental questions which could not be answered honestly in few words.

We use a standard population genetic method to distinguish between hybridizing taxa. Structure identifies in the data distinct genetic entities in HW and gametic equilibria (called clusters) and estimates the probabilities of individuals belonging to these clusters (q-values, interpreted as contributions of each cluster genome to individual genotypes). Technically, species are these clusters. Their identity to M. edulis and M. trossulus was proved in previous studies cited in the manuscript. Classification of genotypes into ancestry classes (e.g. purebreds, different types of early generation hybrids) by q-values is tricky and neither approach can be considered universal.

M. edulis and M. trossulus behave as sympatric species rather than any “hybrid swarm” in our target contact zone in Russia as well as in America and Scotland (and Greenland, doi: 10.1007/s00300-015-1785-x.). This is a known fact. To our mind this makes M. edulis and M. trossulus a reliable model to illustrate the statistical method for identification of cryptic species in sympatry by semi-diagnostic binary marker. With singular binary marker it is not practical to distinguish between more than two categories.

The case of the Baltic mussel you seems to be familiar with, is somehow different and makes you think about the species concept. If M. edulis and M. trossulus hybridize so easily in the Baltic Sea is it reasonable to distinguish them at all? Or probably it is more honest to say they all are hybrids, members of the unitary and indivisible “hybrid swarm”? Perhaps the truth lies in the middle. It is reasonable to classify mussels from the unimodal hybrid zone by taxonomic ancestry e.g. by scores of q-values, but it is improbable that any class would exclusively include purebreds as well as particular categories of early-generation hybrids. We classified them into M. edulis- and M. trossulus-like (what seems reasonable) and founded that differences in morphotype frequencies are quite small. That’s why we did not recommend our method for individual taxonomic assignment of mussels in the Baltic contact zone, as well as in Western Norway. No harm, no foul. But at least we gave it a try. What's wrong with the fact that we can identify the proportion of species (their genes) in populations and select exclusively M. trossulus-like mussels by morphotypes in the Baltic?

The “dark stripe character” in obvious hybrid individuals  
[I am aware that “dark stripe” does not fully describe the complexity of the prismatic layer closed to the ligament but I would like to use the term for simplicity.] Including a hybrid category would allow testing how the dark-stripe-character is developed in genetic hybrids! I am not sure whether the length of the “stripe” gives an information?

Please see the new “Genotypic structure of samples” section of Results. Hybrids are intermediate in morphotype frequencies.

The length of the strip was considered in other publication (DOI: 10.1086/697944). In quantitative terms, morphotypes could be expressed as the Z-index (Z = A/L, where A is the distance from the umbo to the anterior boundary of the prismatic layer under the ligament nympha and L is the distance from the umbo to the posterior boundary of the ligament). The continuous expression of a morphological trait allows a slightly better discrimination of M. edulis- and M. trossulus-like mussels than the discrete (E-morphotype vs. T-morphotype) expression in the White Sea. Measuring the length of the strip is borrowing. it cannot be recommended as an express method for species identification.

What is interesting, analysis of museum collections showed that there is a downward trend in the Z – index (among E-morphotypes) in the White Sea mussels with time (since 1970s). We do not know is it due to genetic (introgression) or environmental (e.g. acidification) factors. If there are other historical collections of mussel shells representing time series available it will be interesting to analyze them for morphotypes and the Z-index.

Morphological characters  
Just a question: Is the presence/absence of the “dark stripe” the basis for PT (T-morphotype)

Yes.

Taxonomic structure of populations  
I like the idea that the taxonomic structure of a population might be more important for some scientific questions than the genetic affinity of a given mussel (lines 601-603). However, I do not see the reason why mussels in a mixed population should be identified as either M. trossulus or M. edulis (lines 612-615). Again, why is a status of being a hybrid not considered? Hybrids may have properties that allow surviving in extreme marine environments and Hybrids are the rule (not the exception) in mytilid mussel contact zones.

Please see the new “Genotypic structure of samples” section of Results and above

Length of the discussion  
The discussion takes 12.5 pages and is overly long to my opinion. I feel that it need to be shortened substantially. This particularly refers to the part from line 747 to 876 (= 6 pages). I recommend summarizing the applications and the limitations of the morphotype test in a concise way.

We shortened discussion about 10%.

Terminology  
The parameters Ptros, PT, P(E/edu), P(T/tros), P(tros/T), P(edu/E) are parts of the calculations and the abbreviations are justified and logic. However, it is hard to keep the overview about their meaning. I ask the authors to give the short explanations (as in lines 229-236) as often as possible to help the readers.

We made the captions to the figures more detailed and clear. Predictive values are now explained there too.

6. PLOS authors have the option to publish the peer review history of their article ([what does this mean?](https://journals.plos.org/plosone/s/editorial-and-peer-review-process#loc-peer-review-history)). If published, this will include your full peer review and any attached files.  
  
  
If you choose “no”, your identity will remain anonymous but your review may still be made public.  
  
  
**Do you want your identity to be public for this peer review?** For information about this choice, including consent withdrawal, please see our [Privacy Policy](https://www.plos.org/privacy-policy).

Reviewer #1: No

Reviewer #2: No

[NOTE: If reviewer comments were submitted as an attachment file, they will be attached to this email and accessible via the submission site. Please log into your account, locate the manuscript record, and check for the action link "View Attachments". If this link does not appear, there are no attachment files.]  
  
While revising your submission, please upload your figure files to the Preflight Analysis and Conversion Engine (PACE) digital diagnostic tool, <https://pacev2.apexcovantage.com/>. PACE helps ensure that figures meet PLOS requirements. To use PACE, you must first register as a user. Registration is free. Then, login and navigate to the UPLOAD tab, where you will find detailed instructions on how to use the tool. If you encounter any issues or have any questions when using PACE, please email PLOS at [figures@plos.org](mailto:figures@plos.org). Please note that Supporting Information files do not need this step.

*In compliance with data protection regulations, you may request that we remove your personal registration details at any time.*[*(Remove my information/details)*](https://www.editorialmanager.com/pone/login.asp?a=r)*. Please contact the publication office if you have any questions.*

22