

Restricted dispersal in a sea of gene flow

L. Benestan, K. Fietz, N. Loiseau, P.-E. Guerin, E. Trofimenko, P. S. Rühs, C Schmidt, W. Rath, A. Biastoch, A. Pérez-Ruzafa, P. Baixauli, A. Forcada, E. Arcas, P. Lenfant, S. Mallol, R. Goñi, L. Velez, M. Höppner, S. Kininmonth, D. Mouillot, O. Puebla and S. Manel

Article citation details

Proc. R. Soc. B **288**: 20210458.
<http://dx.doi.org/10.1098/rspb.2021.0458>

Review timeline

Original submission: 28 September 2020
1st revised submission: 24 February 2021
2nd revised submission: 15 April 2021
Final acceptance: 23 April 2021

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2020-2408.R0 (Original submission)

Review form: Reviewer 1 (Eric Crandall)

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Excellent

Quality of the paper: Is the overall quality of the paper suitable?

Good

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

Summary

The authors conduct continuous sampling of four marine species along 1,000 km of the Northwestern Mediterranean coast, and characterize each individual's genome using RAD methodology. They then measure isolation-by-distance in each species across increasing spatial scale, and compare this to data simulated under a variety of dispersal kernels and parameters, finding that these simulations generally yield results similar to the significant IBD detected in two of the four species. Greatest similarity is found with simulations that include the possibility of long-distance dispersal, two events of which they also empirically observed in their data. Lagrangian simulations of larval dispersal underscore the importance of local dispersal with occasional longer-distance events.

Major Comments

Overall, this is a well-conceived and well-executed study that significantly extends themes from Benestan's, Puebla's and Manel's prior research and will make a nice contribution to this journal and to seascape genetics in general. In general the analyses are sound and well-developed and described. My only major issue is that there is some important information missing regarding biological and geographical aspects of the system, as well as regarding the population genetic simulations, which I mention in my specific comments.

Specific Comments

L58 - Not sure Waples and Gaggiotti is the best citation for this, as they didn't really explore marine parameter space. What about, in addition, Gagnaire et al. 2015?

L77 - New paragraph, and even then this is a bit of an abrupt transition. I invite the authors to provide more relevant oceanographic and geographic information about the region and species sampled.

L78 - Please provide more a priori justification for your continuous sampling scheme for the reader either here or in the methods.

L83 - Please provide more information about these diverse life-history traits beyond PLD, such as some estimate of reproductive output per female, reproductive mode (brooding, spawning, etc.), reproductive seasonality, larval behavior, recruitment conditions etc. Probably in a table in the manuscript or supplement. Also, what if anything is known about the spatial distribution of these species, or their census population sizes in the region? How close do they come to the psuedo-continuous distribution in the IBD model?

L109 & L121 - I know DArTSeq is proprietary, but a little more information about library prep and bioinformatic processing would be helpful. I believe it is basically ddRAD?

L127 - Please cite the 1995 paper

L154 - Readers will need more detail about the model here:

- 1) In IBDSim individuals give birth to thousands of gametes, which are subject to the stated emigration rates, but this is not obvious in the manuscript.
- 2) What happens when a node is occupied by multiple individuals following dispersal?
- 3) Were samples taken before or after dispersal?
- 4) How many samples were taken and with what spacing? This question is especially important to understanding how the results compare to the empirical data

L166 - How many samples were taken and with what spacing?

L175 - Stokes is capitalized as it is referring to George Stokes

L201 - Give yourselves some credit! Use the active voice.

L208 - Would be helpful to see a global F_{st} for each species.

L212 - State which two species here.

L236 - I guess you mean Table S4? Can you put some indicator of per simulation significance in this table?

L241 - I don't see any data at all in panel 3A? And very little in 3B?

L273 - The presence of a coastal boundary layer would also be expected to reduce overall dispersal distances (Nickols et al. 2012)

L293 - Pinsky et al 2017 bridged this divide first, and while they are cited in the previous sentence, I think they deserve more explicit mention here.

L300 - Crandall et al. 2012, sampling at larger spatial scales and using coalescent analysis, found evidence of IBD in a snail species with PLD equivalent to that in *P. elaphas*.

L310 - This information, preferably with citations, would be useful at the end of the introduction.

L315 - Crandall et al. 2019 used a coalescent sampler (Migrate) to detect evidence for stepping-stone dispersal when IBD was not evident. They found clear support for a stepping-stone model in 18 of 26 species along the Hawaiian islands, where IBD had only been found in 4 species previously.

L320 - Well put. Overall, a well-written and comprehensive discussion.

L247 - While the archiving described here might meet the minimum standards of the journal, may I suggest that the authors submit their sampling metadata to GEOME (geome-db.org; Deck et al. 2017)? GEOME provides two easy to use ways to then upload the raw genomic reads to INSDC

SRA, and can also link in the SNP calls on Dryad. Doing so will facilitate FAIR access to these data for future researchers (Wilkinson et al. 2019)

Figure 4 - Why aren't results for the lobster shown?

Fig S1B - What are the units of the color scale, and could a clearer scale be used (white->blue or whatever)?

Fig S2 - *S. cabrilla* is difficult to read

Fig S6 - If I'm reading this correctly, the scale of significant IBD slopes is much greater in this species than those in the current manuscript. Might be worthy of discussion?

Eric Crandall

October 20, 2020

Literature Cited

Crandall ED, Toonen RJ, ToBo Laboratory, Selkoe KA. 2019. A coalescent sampler successfully detects biologically meaningful population structure overlooked by F-statistics. *Evolutionary Applications* 12:255–265. DOI: 10.1111/eva.12712.

Crandall ED, Trembl EA, Barber PH. 2012. Coalescent and biophysical models of stepping-stone gene flow in neritid snails. *Molecular Ecology* 21:5579–5598. DOI: 10/f4dvqs.

Deck J, Gaither MR, Ewing R, Bird CE, Davies N, Meyer C, Riginos C, Toonen RJ, Crandall ED. 2017. The Genomic Observatories Metadatabase (GeOMe): A new repository for field and sampling event metadata associated with genetic samples. *PLoS Biology* 15:e2002925. DOI: 10.1371/journal.pbio.2002925.

Gagnaire P-A, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F, Arnaud-Haond S, Bierne N. 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* 8:769–786. DOI: 10.1111/eva.12288.

Nickols KJ, Gaylord B, Largier JL. 2012. The coastal boundary layer: predictable current structure decreases alongshore transport and alters scales of dispersal. *Marine Ecology Progress Series* 464:17–35. DOI: 10/f396dw.

Pinsky ML, Saenz-Agudelo P, Salles OC, Almany GR, Bode M, Berumen ML, Andréfouët S, Thorrold SR, Jones GP, Planes S. 2017. Marine Dispersal Scales Are Congruent over Evolutionary and Ecological Time. *Current biology* : CB:1–16. DOI: 10.1016/j.cub.2016.10.053.

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, Hoen PAC t, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. 2019. Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Nature Ecology and Evolution* 3:1–9. DOI: 10/bdd4.

Review form: Reviewer 2

Recommendation

Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Good

General interest: Is the paper of sufficient general interest?

Good

Quality of the paper: Is the overall quality of the paper suitable?

Marginal

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

Comments are provided in the attachment. (See Appendix A)

Review form: Reviewer 3

Recommendation

Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Good

General interest: Is the paper of sufficient general interest?

Good

Quality of the paper: Is the overall quality of the paper suitable?

Good

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

No

Is it clear?

Yes

Is it adequate?

No

Do you have any ethical concerns with this paper?

No

Comments to the Author

In this study, Benestan et al. explore how IBD patterns change across spatial scales in four co-distributed marine species using impressive genomic data sets. They supplement these empirical analyses with IBD simulations and biophysical models of dispersal. One of the main findings of the study is that dispersal can be limited even when there is little to no genetic structure at regional spatial scales. I enjoyed reading this paper. I found the analyses to be sound, the writing to be generally clear, and the figures to effectively visualize the key results. I also agree with the authors that researchers are often too quick to state that extensive dispersal prevails in the marine environment, based on low levels of genetic structure in most marine species with a pelagic larval state. This paper helps to clarify why such a simple statement is not always true. However, I am unconvinced by some of the framing/motivation of the paper--- namely, that there is an existing disconnect between population genetic data and the results of direct parentage studies. I provide more details below along with a few other comments.

- In the Introduction, the authors ask whether evidence of restricted dispersal revealed by parentage studies can be reconciled with low genetic structure observed in marine species (L 64-66). This question is used as the primary motivation for the study. This question implies that previous empirical work has revealed a disconnect between these two approaches, but I am not convinced this is true. There are a few important points to note here:

First, within parentage studies, there has been a progression over the course of several decades towards empirical estimates of dispersal kernels (see Almany et al. 2013, 2016; D'Aloia et al. 2013, 2015; Williamson et al. 2016 for examples). All of these studies have documented leptokurtic dispersal kernels in marine fishes—that is to say, they all provide evidence of short distance dispersal, but they also estimate that there is a lower level of long distance dispersal (with the average scale of dispersal varying across study species). The notion that most dispersal is limited,

but that some connections occur over larger spatial scales, is therefore well-established in the species that have been the subjects of large parentage studies. I found this manuscript to overlook this major finding from genetic parentage studies (eg it is missing from the Discussion on lines 320-327).

Second, in some of the few fish species that have been intensively studied via parentage analysis, there are complementary population genetic analyses, and these findings are largely congruent. For example, in the clownfish *Amphiprion percula*, dispersal estimates based on parentage work are remarkably similar to those based on IBD demonstrating not only that both datasets can provide consistent dispersal estimates, but also that short distance dispersal can occur despite “extensive gene flow among populations” (Pinsky et al. 2017). In another example, D’Aloia et al. (2020) explore how parentage-based dispersal estimates can help explain population genetic patterns at different spatial scales in the goby *Elacatinus lori*.

In light of these comments, I suggest that the authors should (1) incorporate these earlier findings into their Introduction and Discussion and (2) articulate the major contributions of this paper, given this earlier work. I do not think the novel contribution here is reconciling population genetic approaches with direct parentage approaches as the Abstract implies.

- The authors should clarify what they mean by “continuous” sampling designs and provide more details on their own sampling regimes. Can the authors provide some summary of the average sampling intervals for each species? Looking at Figure 1, there is clearly variability in continuity of sampling across the 4 species. For example, *D. sargus* could be viewed as 3 main clusters of sampling, whereas *M. sumuletus* is more continuous throughout the study area. It is also unclear to me what each circle represents in Figure 1. Is it one individual per site, or can it be multiple? I did not find this information in the supplementary materials so at a minimum the mean and s.d. (and perhaps range) of # samples per site should be reported if it is not one individual.

- Another suggestion regarding terminology is to be clearer about the definition of a “regional” spatial scale. I wondered whether this is a clear spatial definition, or whether it may depend upon the species. For example, in the extreme, what is regional versus local scale for a cosmopolitan species versus an endemic?

- Given that this is a short-format paper the Discussion could be improved by staying focused on the major implications of the study. The final two paragraphs (out of only 5 total) seemed to stray from the focus of the paper (especially the text on uncovering the genomic architecture of adaptation). I think a brief mention of the implications for MPA design is warranted, but otherwise I suggest removing most of this text. This would give you space to discuss your results and how it relates to the large body of existing work on marine dispersal.

-One area that I felt deserved more attention is further explanation for why IBD disappears at large spatial scales. There is a reference to another theory paper in the Introduction (L73), but more discussion in the present manuscript is warranted since this was a key result in two of your study species. I understand why IBD signals are obscured when there is strong genetic structure, but could you provide more explanation for why this happens in the absence of structure? Your IBD simulations show that the decrease in IBD at large spatial scales is not necessarily because of long distance dispersal, as this pattern can emerge in the stepping stone simulations. So, what else could explain it?

- Finally, I was unable to access the data via the dryad link in the manuscript.

Minor comments:

L56: "It can persist..." is ambiguous language. Please reword to clarify whether "it" refers to "weak genetic structure"

L166: "Samples were taken linearly" ... I found this confusing. Could you please rephrase to make this clearer?

L204-206: This wording should be revised, as there is genetic structure in one of your study species.

L257-258: These are clearly not parent offspring pairs (kinship coefficient would be ~ 0.25) so reference to such pairs should be removed from this sentence.

L302: Remove "nevertheless" from sentence as it adds confusion.

L309 and remainder of this paragraph: Suggest replacing "adult dispersal" with "adult migration" or "adult movement." Dispersal has a specific definition, and I don't think the type of movement covered here is referring to dispersal.

References:

Almany, G. R., Hamilton, R. J., Bode, M., Matawai, M., Potuku, T., Saenz-Agudelo, P., ... & Russ, G. R. (2013). Dispersal of grouper larvae drives local resource sharing in a coral reef fishery. *Current biology*, 23(7), 626-630.

Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., ... & Nanninga, G. B. (2017). Larval fish dispersal in a coral-reef seascape. *Nature Ecology & Evolution*, 1(6), 1-7.

D'Aloia, C. C., Bogdanowicz, S. M., Francis, R. K., Majoris, J. E., Harrison, R. G., & Buston, P. M. (2015). Patterns, causes, and consequences of marine larval dispersal. *Proceedings of the National Academy of Sciences*, 112(45), 13940-13945.

D'Aloia, C. C., Andrés, J. A., Bogdanowicz, S. M., McCune, A. R., Harrison, R. G., & Buston, P. M. (2020). Unraveling hierarchical genetic structure in a marine metapopulation: A comparison of three high-throughput genotyping approaches. *Molecular Ecology*.

Pinsky, M. L., Saenz-Agudelo, P., Salles, O. C., Almany, G. R., Bode, M., Berumen, M. L., ... & Planes, S. (2017). Marine dispersal scales are congruent over evolutionary and ecological time. *Current Biology*, 27(1), 149-154.

Williamson, D. H., Harrison, H. B., Almany, G. R., Berumen, M. L., Bode, M., Bonin, M. C., ... & Jones, G. P. (2016). Large-scale, multidirectional larval connectivity among coral reef fish populations in the Great Barrier Reef Marine Park. *Molecular Ecology*, 25(24), 6039-6054.

Decision letter (RSPB-2020-2408.R0)

20-Nov-2020

Dear Dr Benestan:

I am writing to inform you that your manuscript RSPB-2020-2408 entitled "Restricted dispersal in a sea of gene flow" has, in its current form, been rejected for publication in *Proceedings B*.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

- 1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
- 2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
- 3) Line numbers in your main document.
- 4) Data - please see our policies on data sharing to ensure that you are complying (<https://royalsociety.org/journals/authors/author-guidelines/#data>).

To upload a resubmitted manuscript, log into <http://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Dr Sasha Dall
mailto:proceedingsb@royalsociety.org

Associate Editor
Board Member: 1
Comments to Author:

We have received a third review of your paper, which needs to be taken into account when revising your manuscript. As you will see, the referee feels you need to better frame the context of your study in order to provide a more accurate description of what the main contribution of your study is.

Reviewer(s)' Comments to Author:
Referee: 1
Comments to the Author(s)
Summary

The authors conduct continuous sampling of four marine species along 1,000 km of the Northwestern Mediterranean coast, and characterize each individual's genome using RAD methodology. They then measure isolation-by-distance in each species across increasing spatial scale, and compare this to data simulated under a variety of dispersal kernels and parameters, finding that these simulations generally yield results similar to the significant IBD detected in two of the four species. Greatest similarity is found with simulations that include the possibility of long-distance dispersal, two events of which they also empirically observed in their data. Lagrangian simulations of larval dispersal underscore the importance of local dispersal with occasional longer-distance events.

Major Comments

Overall, this is a well-conceived and well-executed study that significantly extends themes from Benestan's, Puebla's and Manel's prior research and will make a nice contribution to this journal and to seascape genetics in general. In general the analyses are sound and well-developed and described. My only major issue is that there is some important information missing regarding biological and geographical aspects of the system, as well as regarding the population genetic simulations, which I mention in my specific comments.

Specific Comments

L58 - Not sure Waples and Gaggiotti is the best citation for this, as they didn't really explore marine parameter space. What about, in addition, Gagnaire et al. 2015?

L77 - New paragraph, and even then this is a bit of an abrupt transition. I invite the authors to provide more relevant oceanographic and geographic information about the region and species sampled.

L78 - Please provide more a priori justification for your continuous sampling scheme for the reader either here or in the methods.

L83 - Please provide more information about these diverse life-history traits beyond PLD, such as some estimate of reproductive output per female, reproductive mode (brooding, spawning, etc.), reproductive seasonality, larval behavior, recruitment conditions etc. Probably in a table in the manuscript or supplement. Also, what if anything is known about the spatial distribution of these species, or their census population sizes in the region? How close do they come to the psuedo-continuous distribution in the IBD model?

L109 & L121 - I know DArTSeq is proprietary, but a little more information about library prep and bioinformatic processing would be helpful. I believe it is basically ddRAD?

L127 - Please cite the 1995 paper

L154 - Readers will need more detail about the model here:

- 1) In IBDSim individuals give birth to thousands of gametes, which are subject to the stated emigration rates, but this is not obvious in the manuscript.
- 2) What happens when a node is occupied by multiple individuals following dispersal?
- 3) Were samples taken before or after dispersal?
- 4) How many samples were taken and with what spacing? This question is especially important to understanding how the results compare to the empirical data

L166 - How many samples were taken and with what spacing?

L175 - Stokes is capitalized as it is referring to George Stokes

L201 - Give yourselves some credit! Use the active voice.

L208 - Would be helpful to see a global F_{st} for each species.

L212 - State which two species here.

L236 - I guess you mean Table S4? Can you put some indicator of per simulation significance in this table?

L241 - I don't see any data at all in panel 3A? And very little in 3B?

L273 - The presence of a coastal boundary layer would also be expected to reduce overall dispersal distances (Nickols et al. 2012)

L293 - Pinsky et al 2017 bridged this divide first, and while they are cited in the previous sentence, I think they deserve more explicit mention here.

L300 - Crandall et al. 2012, sampling at larger spatial scales and using coalescent analysis, found evidence of IBD in a snail species with PLD equivalent to that in *P. elaphas*.

L310 - This information, preferably with citations, would be useful at the end of the introduction.

L315 - Crandall et al. 2019 used a coalescent sampler (Migrate) to detect evidence for stepping-stone dispersal when IBD was not evident. They found clear support for a stepping-stone model in 18 of 26 species along the Hawaiian islands, where IBD had only been found in 4 species previously.

L320 - Well put. Overall, a well-written and comprehensive discussion.

L247 - While the archiving described here might meet the minimum standards of the journal, may I suggest that the authors submit their sampling metadata to GEOME (geome-db.org; Deck et al. 2017)? GEOME provides two easy to use ways to then upload the raw genomic reads to INSDC SRA, and can also link in the SNP calls on Dryad. Doing so will facilitate FAIR access to these data for future researchers (Wilkinson et al. 2019)

Figure 4 - Why aren't results for the lobster shown?

Fig S1B - What are the units of the color scale, and could a clearer scale be used (white->blue or whatever)?

Fig S2 - *S. cabrilla* is difficult to read

Fig S6 - If I'm reading this correctly, the scale of significant IBD slopes is much greater in this species than those in the current manuscript. Might be worthy of discussion?

Eric Crandall
October 20, 2020

Literature Cited

Crandall ED, Toonen RJ, ToBo Laboratory, Selkoe KA. 2019. A coalescent sampler successfully detects biologically meaningful population structure overlooked by F-statistics. *Evolutionary Applications* 12:255–265. DOI: [10.1111/eva.12712](https://doi.org/10.1111/eva.12712).

Crandall ED, Trembl EA, Barber PH. 2012. Coalescent and biophysical models of stepping-stone gene flow in neritid snails. *Molecular Ecology* 21:5579–5598. DOI: [10/1365-3113.2012.02712.x](https://doi.org/10.1111/j.1365-3113.2012.02712.x).

Deck J, Gaither MR, Ewing R, Bird CE, Davies N, Meyer C, Riginos C, Toonen RJ, Crandall ED. 2017. The Genomic Observatories Metadatabase (GeOME): A new repository for field and sampling event metadata associated with genetic samples. *PLoS Biology* 15:e2002925. DOI: [10.1371/journal.pbio.2002925](https://doi.org/10.1371/journal.pbio.2002925).

Gagnaire P-A, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F, Arnaud-Haond S, Bierne N. 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* 8:769–786. DOI: [10.1111/eva.12288](https://doi.org/10.1111/eva.12288).

Nickols KJ, Gaylord B, Largier JL. 2012. The coastal boundary layer: predictable current structure decreases alongshore transport and alters scales of dispersal. *Marine Ecology Progress Series* 464:17–35. DOI: 10/f396dw.

Pinsky ML, Saenz-Agudelo P, Salles OC, Almany GR, Bode M, Berumen ML, Andréfouët S, Thorrold SR, Jones GP, Planes S. 2017. Marine Dispersal Scales Are Congruent over Evolutionary and Ecological Time. *Current biology* : CB:1–16. DOI: 10.1016/j.cub.2016.10.053.

Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, Hoen PAC t, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. 2019. Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Nature Ecology and Evolution* 3:1–9. DOI: 10/bdd4.

Referee: 2

Comments to the Author(s)

Comments are provided in the attachment

Referee: 3

Comments to the Author(s)

In this study, Benestan et al. explore how IBD patterns change across spatial scales in four co-distributed marine species using impressive genomic data sets. They supplement these empirical analyses with IBD simulations and biophysical models of dispersal. One of the main findings of the study is that dispersal can be limited even when there is little to no genetic structure at regional spatial scales. I enjoyed reading this paper. I found the analyses to be sound, the writing to be generally clear, and the figures to effectively visualize the key results. I also agree with the authors that researchers are often too quick to state that extensive dispersal prevails in the marine environment, based on low levels of genetic structure in most marine species with a pelagic larval state. This paper helps to clarify why such a simple statement is not always true. However, I am unconvinced by some of the framing/motivation of the paper--- namely, that there is an existing disconnect between population genetic data and the results of direct parentage studies. I provide more details below along with a few other comments.

- In the Introduction, the authors ask whether evidence of restricted dispersal revealed by parentage studies can be reconciled with low genetic structure observed in marine species (L 64-66). This question is used as the primary motivation for the study. This question implies that previous empirical work has revealed a disconnect between these two approaches, but I am not convinced this is true. There are a few important points to note here:

First, within parentage studies, there has been a progression over the course of several decades towards empirical estimates of dispersal kernels (see Almany et al. 2013, 2016; D’Aloia et al. 2013, 2015; Williamson et al. 2016 for examples). All of these studies have documented leptokurtic dispersal kernels in marine fishes – that is to say, they all provide evidence of short distance dispersal, but they also estimate that there is a lower level of long distance dispersal (with the average scale of dispersal varying across study species). The notion that most dispersal is limited, but that some connections occur over larger spatial scales, is therefore well-established in the species that have been the subjects of large parentage studies. I found this manuscript to overlook this major finding from genetic parentage studies (eg it is missing from the Discussion on lines 320-327).

Second, in some of the few fish species that have been intensively studied via parentage analysis, there are complementary population genetic analyses, and these findings are largely congruent. For example, in the clownfish *Amphiprion percula*, dispersal estimates based on parentage work are remarkably similar to those based on IBD demonstrating not only that both datasets can provide consistent dispersal estimates, but also that short distance dispersal can occur despite “extensive gene flow among populations” (Pinsky et al. 2017). In another example, D’Aloia et al. (2020) explore how parentage-based dispersal estimates can help explain population genetic patterns at different spatial scales in the goby *Elacatinus lori*.

In light of these comments, I suggest that the authors should (1) incorporate these earlier findings into their Introduction and Discussion and (2) articulate the major contributions of this paper, given this earlier work. I do not think the novel contribution here is reconciling population genetic approaches with direct parentage approaches as the Abstract implies.

- The authors should clarify what they mean by “continuous” sampling designs and provide more details on their own sampling regimes. Can the authors provide some summary of the average sampling intervals for each species? Looking at Figure 1, there is clearly variability in continuity of sampling across the 4 species. For example, *D. sargus* could be viewed as 3 main clusters of sampling, whereas *M. sumuletus* is more continuous throughout the study area. It is also unclear to me what each circle represents in Figure 1. Is it one individual per site, or can it be multiple? I did not find this information in the supplementary materials so at a minimum the mean and s.d. (and perhaps range) of # samples per site should be reported if it is not one individual.

- Another suggestion regarding terminology is to be clearer about the definition of a “regional” spatial scale. I wondered whether this is a clear spatial definition, or whether it may depend upon the species. For example, in the extreme, what is regional versus local scale for a cosmopolitan species versus an endemic?

- Given that this is a short-format paper the Discussion could be improved by staying focused on the major implications of the study. The final two paragraphs (out of only 5 total) seemed to stray from the focus of the paper (especially the text on uncovering the genomic architecture of adaptation). I think a brief mention of the implications for MPA design is warranted, but otherwise I suggest removing most of this text. This would give you space to discuss your results and how it relates to the large body of existing work on marine dispersal.

-One area that I felt deserved more attention is further explanation for why IBD disappears at large spatial scales. There is a reference to another theory paper in the Introduction (L73), but more discussion in the present manuscript is warranted since this was a key result in two of your study species. I understand why IBD signals are obscured when there is strong genetic structure, but could you provide more explanation for why this happens in the absence of structure? Your IBD simulations show that the decrease in IBD at large spatial scales is not necessarily because of long distance dispersal, as this pattern can emerge in the stepping stone simulations. So, what else could explain it?

- Finally, I was unable to access the data via the dryad link in the manuscript.

Minor comments:

L56: “It can persist...” is ambiguous language. Please reword to clarify whether “it” refers to “weak genetic structure”

L166: “Samples were taken linearly”... I found this confusing. Could you please rephrase to make this clearer?

L204-206: This wording should be revised, as there is genetic structure in one of your study species.

L257-258: These are clearly not parent offspring pairs (kinship coefficient would be ~0.25) so reference to such pairs should be removed from this sentence.

L302: Remove “nevertheless” from sentence as it adds confusion.

L309 and remainder of this paragraph: Suggest replacing “adult dispersal” with “adult migration” or “adult movement.” Dispersal has a specific definition, and I don’t think the type of movement covered here is referring to dispersal.

References:

Almany, G. R., Hamilton, R. J., Bode, M., Matawai, M., Potuku, T., Saenz-Agudelo, P., ... & Russ, G. R. (2013). Dispersal of grouper larvae drives local resource sharing in a coral reef fishery. *Current biology*, 23(7), 626-630.

Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., ... & Nanninga, G. B. (2017). Larval fish dispersal in a coral-reef seascape. *Nature Ecology & Evolution*, 1(6), 1-7.

D’Aloia, C. C., Bogdanowicz, S. M., Francis, R. K., Majoris, J. E., Harrison, R. G., & Buston, P. M. (2015). Patterns, causes, and consequences of marine larval dispersal. *Proceedings of the National Academy of Sciences*, 112(45), 13940-13945.

D’Aloia, C. C., Andrés, J. A., Bogdanowicz, S. M., McCune, A. R., Harrison, R. G., & Buston, P. M. (2020). Unraveling hierarchical genetic structure in a marine metapopulation: A comparison of three high-throughput genotyping approaches. *Molecular Ecology*.

Pinsky, M. L., Saenz-Agudelo, P., Salles, O. C., Almany, G. R., Bode, M., Berumen, M. L., ... & Planes, S. (2017). Marine dispersal scales are congruent over evolutionary and ecological time. *Current Biology*, 27(1), 149-154.

Williamson, D. H., Harrison, H. B., Almany, G. R., Berumen, M. L., Bode, M., Bonin, M. C., ... & Jones, G. P. (2016). Large-scale, multidirectional larval connectivity among coral reef fish populations in the Great Barrier Reef Marine Park. *Molecular Ecology*, 25(24), 6039-6054.

Author's Response to Decision Letter for (RSPB-2020-2408.R0)

See Appendix B.

RSPB-2021-0458.R0

Review form: Reviewer 2

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Excellent

Quality of the paper: Is the overall quality of the paper suitable?

Excellent

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

Following revisions, this manuscript by Benestan et al. is much improved. The authors have made many changes that clarify their study design and findings, as well as their contribution to the literature. We have a few additional suggestions and questions to further strengthen the manuscript, but otherwise, we think the authors have done a good job of addressing reviewer comments.

A legend is needed to relate point size to the number of individuals in Figure 1.

Is there any chance that differences in library preparation contributed to differences in IBD estimates among species? For example, if DArT-seq targets hypo-methylated sites that are associated with a restriction enzyme cut sites, presumably the DArT loci would have greater functional relevance than the RAD loci and might be associated with environmental differences in the study region.

Last paragraph seems like it was added as an afterthought. We suggest incorporating it into the previous paragraph to summarize the findings and to suggest next steps.

The authors suggest that IBD at small spatial scales in *M. surmuletus* and *S. cabrilla* might be related to the fact that sampling occurred on the spawning grounds in months that align with the known spawning seasons of these fish (lines 344-346). However, Table S1 says that spawning migration is unknown for these species. It seems that the findings of IBD at small scales for these particular species could potentially be due to adult spawning migration (as seems to be the case

for the summer flounder study), larval dispersal or a combination of the two. The authors do make clear that sampling of adults incorporates larval and adult movement (lines 350-351), but we think that the authors need to be very clear that adult and/or larval movement could explain their IBD findings in *M. surmuletus* and *S. cabrilla*, and not just adult movement.

The Dryad link provided in the manuscript and the reviewer responses didn't work. We also couldn't access the `finescale_ibd` repository on GitHub. Please double check the links are correct and the repositories are public.

Line 56. Replace 'high' with 'large'

Line 68. Replace 'have therefore to' to 'have to therefore' or 'therefore have to'

Line 72. Allows one? Allows us?

Line 80. Allow us?

Line 252. Put 'i.e. from marine species sampled continuously at large spatial scale' in parentheses and delete prior ','. Or replace prior ', ' with ', '.

Line 330 and 331. Change 'scale' to 'scales'

Review form: Reviewer 3

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Good

General interest: Is the paper of sufficient general interest?

Good

Quality of the paper: Is the overall quality of the paper suitable?

Excellent

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

Thanks for the thoughtful response to reviewers. I enjoyed rereading the manuscript and think you did a very nice job addressing the many comments! In this version, there is greater clarity about the four study species, the spatial scales of analysis, and how your work fits in within the very large body of marine dispersal research. I think the paper is a nice contribution. A few minor comments:

-I'm sure there are word limits, but in the Abstract it seems like it would be prudent to clarify that the pattern of IBD at small spatial scales was only found in some of your study species. Documenting variability across species is one of the strengths of a multi-species study and an important conclusion.

-I appreciate the greater clarity about spatial scale, especially since it's an important aspect of your study. This is nit-picky, but when you talk about LDD, why is 40 km used as a cut-off? It seems arbitrary to me (including in the referenced paper) and I wonder if you need it here as it doesn't seem to really affect your analyses.

-I suggest you create a new paragraph when you start talking about NGS in the introduction (L 69). I felt this part could use a little more refining. What exactly are you trying to emphasize here? I don't think it's just about relatedness. You don't necessarily need SNPs to identify relationships beyond parent-offspring. Are you trying to emphasize that it can improve genetic structure and IBD estimates? Or are you trying to get at the difference between direct and indirect genetic studies of dispersal. It's unclear. You also don't need SNP data to conduct continuously-sampled study designs (and this seems like a separate issue). I think a revision of this paragraph would help you further articulate how your study fits in.

-Fig. 1 caption: It would help to clarify that size of bubble reflects # individuals per site (and also consider including the range of samples per site). I still have a hard time figuring out how many were collected per site.

-L 109: It would be helpful to include some of the information from the caption of Fig. S1 here. Sampling was 'continuous' but the distribution of distances was not even.

-L288: You may want to clarify that 292 and 294 km are just max distances between individuals. Their "origin" may be somewhere between them, or somewhere even further away (if they're even sibs).

Decision letter (RSPB-2021-0458.R0)

26-Mar-2021

Dear Dr Benestan:

Your manuscript has now been peer reviewed and the reviews have been assessed by an Associate Editor. The reviewers' comments (not including confidential comments to the Editor) and the comments from the Associate Editor are included at the end of this email for your reference. As you will see, the reviewers and the Editors have raised some concerns with your manuscript and we would like to invite you to revise your manuscript to address them.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address all of the comments at this stage. If deemed necessary by the Associate Editor, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please log into <http://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions", click on "Create a Revision". Your manuscript number has been appended to denote a revision.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (<https://royalsociety.org/journals/ethics-policies/>). You should pay particular attention to the following:

Research ethics:

If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

Use of animals and field studies:

If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

Data accessibility and data citation:

It is a condition of publication that you make available the data and research materials supporting the results in the article (<https://royalsociety.org/journals/authors/author-guidelines/#data>). Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article (<https://royalsociety.org/journals/ethics-policies/data-sharing-mining/>). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

If you wish to submit your data to Dryad (<http://datadryad.org/>) and have not already done so you can submit your data via this link

[http://datadryad.org/submit?journalID=RSPB&manu=\(Document not available\)](http://datadryad.org/submit?journalID=RSPB&manu=(Document not available)), which will take you to your unique entry in the Dryad repository.

If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link.

For more information please see our open data policy <http://royalsocietypublishing.org/data-sharing>.

Electronic supplementary material:

All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI. Please try to submit all supplementary material as a single file.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,
Dr Sasha Dall
mailto:proceedingsb@royalsociety.org

Associate Editor Board Member

Comments to Author:

Two referees that reviewed your original submission have evaluated the revised version and agree that you have satisfactorily incorporated their suggestions. Referee 1's comments are aimed at improving the clarity of the manuscript. Referee 2, however, make more substantial comments that need attention. In particular, he/she highlights the different library preparation used for different species, which could lead to differences in the evolutionary forces driving spatial patterns of genetic variation (DArT loci are more likely to be associated to functional variation than RAD loci). A paragraph with this caveat and its potential effects on the interpretation of results should be added to the discussion.

A very important remaining issue is that the data has not been made available. The authors provide a Dryad link but the DOI provided is incorrect. It is unclear to me if this is a genuine mistake or not but, in either case, the authors should provide the required information, which should be subsequently verified to make sure that it is correct.

Reviewer(s)' Comments to Author:

Referee: 3

Comments to the Author(s).

Thanks for the thoughtful response to reviewers. I enjoyed rereading the manuscript and think you did a very nice job addressing the many comments! In this version, there is greater clarity about the four study species, the spatial scales of analysis, and how your work fits in within the

very large body of marine dispersal research. I think the paper is a nice contribution. A few minor comments:

-I'm sure there are word limits, but in the Abstract it seems like it would be prudent to clarify that the pattern of IBD at small spatial scales was only found in some of your study species.

Documenting variability across species is one of the strengths of a multi-species study and an important conclusion.

-I appreciate the greater clarity about spatial scale, especially since it's an important aspect of your study. This is nit-picky, but when you talk about LDD, why is 40 km used as a cut-off? It seems arbitrary to me (including in the referenced paper) and I wonder if you need it here as it doesn't seem to really affect your analyses.

-I suggest you create a new paragraph when you start talking about NGS in the introduction (L 69). I felt this part could use a little more refining. What exactly are you trying to emphasize here? I don't think it's just about relatedness. You don't necessarily need SNPs to identify relationships beyond parent-offspring. Are you trying to emphasize that it can improve genetic structure and IBD estimates? Or are you trying to get at the difference between direct and indirect genetic studies of dispersal. It's unclear. You also don't need SNP data to conduct continuously-sampled study designs (and this seems like a separate issue). I think a revision of this paragraph would help you further articulate how your study fits in.

-Fig. 1 caption: It would help to clarify that size of bubble reflects # individuals per site (and also consider including the range of samples per site). I still have a hard time figuring out how many were collected per site.

-L 109: It would be helpful to include some of the information from the caption of Fig. S1 here. Sampling was 'continuous' but the distribution of distances was not even.

-L288: You may want to clarify that 292 and 294 km are just max distances between individuals. Their "origin" may be somewhere between them, or somewhere even further away (if they're even sibs).

Referee: 2

Comments to the Author(s).

Following revisions, this manuscript by Benestan et al. is much improved. The authors have made many changes that clarify their study design and findings, as well as their contribution to the literature. We have a few additional suggestions and questions to further strengthen the manuscript, but otherwise, we think the authors have done a good job of addressing reviewer comments.

A legend is needed to relate point size to the number of individuals in Figure 1.

Is there any chance that differences in library preparation contributed to differences in IBD estimates among species? For example, if DArT-seq targets hypo-methylated sites that are associated with a restriction enzyme cut sites, presumably the DArT loci would have greater functional relevance than the RAD loci and might be associated with environmental differences in the study region.

Last paragraph seems like it was added as an afterthought. We suggest incorporating it into the previous paragraph to summarize the findings and to suggest next steps.

The authors suggest that IBD at small spatial scales in *M. surmuletus* and *S. cabrilla* might be related to the fact that sampling occurred on the spawning grounds in months that align with the known spawning seasons of these fish (lines 344-346). However, Table S1 says that spawning migration is unknown for these species. It seems that the findings of IBD at small scales for these

particular species could potentially be due to adult spawning migration (as seems to be the case for the summer flounder study), larval dispersal or a combination of the two. The authors do make clear that sampling of adults incorporates larval and adult movement (lines 350-351), but we think that the authors need to be very clear that adult and/or larval movement could explain their IBD findings in *M. surmuletus* and *S. cabrilla*, and not just adult movement.

The Dryad link provided in the manuscript and the reviewer responses didn't work. We also couldn't access the `finescale_ibd` repository on GitHub. Please double check the links are correct and the repositories are public.

Line 56. Replace 'high' with 'large'

Line 68. Replace 'have therefore to' to 'have to therefore' or 'therefore have to'

Line 72. Allows one? Allows us?

Line 80. Allow us?

Line 252. Put 'i.e. from marine species sampled continuously at large spatial scale' in parentheses and delete prior ','. Or replace prior ', ' with ', '.

Line 330 and 331. Change 'scale' to 'scales'

Author's Response to Decision Letter for (RSPB-2021-0458.R0)

See Appendix C.

Decision letter (RSPB-2021-0458.R1)

23-Apr-2021

Dear Dr Benestan

I am pleased to inform you that your manuscript entitled "Restricted dispersal in a sea of gene flow" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Data Accessibility section

Please remember to make any data sets live prior to publication, and update any links as needed when you receive a proof to check. It is good practice to also add data sets to your reference list.

Open Access

You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CC BY licence. Our article processing charge for Open Access is £1700.

Corresponding authors from member institutions

(<http://royalsocietypublishing.org/site/librarians/allmembers.xhtml>) receive a 25% discount to these charges. For more information please visit <http://royalsocietypublishing.org/open-access>.

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Paper charges

An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). The preferred payment method is by credit card; however, other payment options are available

Electronic supplementary material:

All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,

Dr Sasha Dall

Editor, Proceedings B

mailto:proceedingsb@royalsociety.org

Associate Editor:

Board Member

Comments to Author:

Your revised manuscript incorporates all suggestions made by the referees. I have also verified that the dataset is indeed available from the Dryad database. As a result, I think it is a valuable contribution to the field of marine ecology.

Appendix A

In this manuscript, the authors investigate isolation by distance at different geographic scales using genetics, computer simulations and a biophysical model. The topic of the manuscript is interesting, and the results are relevant, with implications for ecology, evolution and conservation. Strengths of the manuscript include the use of a null model to understand the IBD results and the diversity of methods used to investigate the question. Below, we provide suggestions for improvement.

1. We recognize that Proceedings B manuscripts are relatively short, but we felt that additional information about the four study species was needed earlier on in the paper. This information not only provides the relevant context for the hypotheses, but also provides key details for interpreting the biological importance of the authors' findings. For example, information on what is known about dispersal distances, adult spawning timing and location for each species was generally lacking, or provided too late, which made it challenging to understand the rationale for the hypotheses and to put the results within the appropriate context. This would also help to flesh out the 'diversity of life history traits' that the authors mention, rather than only focusing on PLD.
2. Are any of the four species restricted to certain depth ranges within the study region, and was this accounted for when geographic distance was calculated using the bathymetry data for the IBD analysis? An explanation of why Loiselle's kinship coefficient was used for IBD calculations would also be helpful.
3. The authors are using pelagic larval duration (PLD) as a proxy for dispersal distance, but PLD does not actually have a strong relationship to dispersal distance. We were confused about its inclusion. The authors need to explain why this is relevant to the study, especially if adults of each species were sampled. Is dispersal during the larval stage thought to be the dominant dispersal mode for each of these species (see point #1)?
 - a. Nanninga, Gerrit B., and Andrea Manica. "Larval swimming capacities affect genetic differentiation and range size in demersal marine fishes." *Marine Ecology Progress Series* 589 (2018): 1-12.
 - b. Selkoe, Kimberly A., et al. "Emergent patterns of population genetic structure for a coral reef community." *Molecular ecology* 23.12 (2014): 3064-3079. for one example
4. In general, the methods were well documented, but additional details for certain analyses would help readers to follow along more easily. In particular, the biophysical model methods seemed incomplete (further details in point #7). Also, the population structure methods appear to be missing from section 2d.
5. Since samples came from the artisanal fisherman, are the samples from mature/adult individuals that were captured on spawning grounds? The authors state that *D. sargus* were not captured on spawning grounds in lines 306-308, so it seems like this might be true of the other species too? Sampling individuals when they are not on their spawning grounds could miss true population structure and impact IBD results. Related to this, P.

dentatus were captured on their spawning grounds, so at least some discussion of how sampling during different parts of a species' lifecycle might impact IBD seems warranted. For Figure 1, the sampling is so dense in some areas that it is challenging to see on the maps. Is each unique lat/lon coordinate representative of a single individual? If it is not, the authors might consider using point area to indicate the number of individuals captured at each lat/lon coordinate and/or to make points semi-transparent.

6. The author's definitions of short, regional and long distance dispersal in the manuscript seemed arbitrary, despite the citation of papers in which these were defined. The authors report that their results were not sensitive to the inclusion of closely related individuals, but were siblings directly tested for (i.e. Colony, etc)?
7. The author's use of the biophysical model left many questions. The authors cite the Bode et al 2019 study, but also see North E, Gallego A, Petitgas P. Manual of recommended practices for modelling physical-biological interactions in fish early life. Copenhagen: International Council for the Exploration of the Sea; 2009 2009. 1-112 p. It would be helpful if the authors revisited the biophysical model section in the context of what is known to be important for capturing dispersal. For example, are these species demersal spawners? If so, then were particles released on the surface or from the benthos? Why were only 100 particles released? There are limitations to what can be reasonably implemented, but some justification for choices is needed. In general, the biophysical model itself seems to distract from the novelty of this work- the investigation across spatial scales is exciting, and the biophysical model section seems ancillary.
8. The authors could do better to contextualize their study within the dispersal literature, to further discuss why IBD at different spatial scales matters and to discuss the scale at which IBD should be examined. In general, we found the introduction and discussion to be undercited. For example, a huge body of literature has gone into the statement: 'The marine environment is also more homogenous than terrestrial and freshwater environments, and presents fewer geographic barriers to the movement of organisms,' (lines 49-50) yet no citations were provided.

Figure 3. In the figure caption, we suggest pointing out that the dispersal function utilized for each panel is in the inset. What do #75, #76, #77 and #106 mean in the insets?

Line 34. It is unclear as written why discrete or heterogeneous sampling limits regional scale inference.

Line 48. Limited citations given.

Line 49. No citations given.

Line 64. The examples given are fishes, but the literature is not only fishes. We suggest eliminating this clause and citing the existing literature more

Line 67. The use of “unprecedented” should probably be tempered. In terms of spatial scale, the paper cited later, Hoey & Pinsky 2018 was a larger spatial extent, and in terms of direct and indirect dispersal estimates, please see the following work in addition to D’Aloia et al 2015

D'Aloia, C. C., et al. "Seascape continuity plays an important role in determining patterns of spatial genetic structure in a coral reef fish." *Molecular Ecology* 23.12 (2014): 2902-2913.

Line 96. Replace ‘conserve’ with ‘preserve’

Line 118-119. Are these previous de novo references or full genome references? Regardless, pointing readers to where these references exist in the literature will be helpful.

Line 127. More details on settings and implementation of GENODIVE would be good.

Line 133. It is unclear what ‘ “adjusted” distance required to keep those individuals in the analysis’ is. So there was a criteria cutoff?

Line 157. Please provide more details on the implementation of IBDSim. In general, these types of details will be helpful for readers to understand assumptions and also for future researchers reading this for methods.

Line 180. Parcels allows one to?

Line 201. Where did so many samples get lost? We are told 2,034 samples to start, and 1,299 individuals are in the analysis?

Line 205. “Relative” to what?

Line 215. Readers need much more information about these related individuals. It would be nice to see the authors try to parse out what these relationships are, especially because the authors base the “long distance dispersal” results on these relationships. Are they siblings? In which case, how are the authors inferring a dispersal trajectory?

Line 217-222. This should be moved up to line 213 before the authors start talking about highly related and geographically close individuals.

Line 254. This section is not substantial because there isn't life history information with which to contextualize dispersal for the species, the authors have not inferred the relationship (just close kinship, which in itself does not give information regarding dispersal), and the arbitrary use of "long distance" when the literature has given examples of how to define "long distance dispersal".

Line 273. The inferences made from this "secondary peak" are hard to understand when only 100 particles were released, meaning that only a small fraction of possible dispersal trajectories were sampled with this model.

Line 300. See general comments about the PLD framing.

Lines 304-308. Readers will be looking for this sort of information in the Introduction or Methods, so you might consider mentioning and flushing out relevant details much earlier

Line 311. The point being made is unclear, and some life history information would help the readers understand better.

Line 322. This distance seems very arbitrary. Some justification is needed.

Line 323. Defining local vs regional in terms of the spatial scale of this study would be helpful.

Lines 328-336. This brings up the interesting question of scale. For conservation practices, it might be important to create a network of nearby MPAs. On the other hand, perhaps larger scales are more helpful for understanding population structure throughout a species range.

Line 330. Another example of a statement not supported by a citation.

Fig S2. Sorted by latitude going from left to right?

Fig S4. Are the significant slopes plotted on here?

Fig S6. The use of clipart in the figures was great! However, the flatfish in Fig S6 is likely a European flounder and not a summer flounder.

Appendix B

Dear associate Editor,

Thanks for your interest in considering a revised version of our manuscript entitled "**Restricted dispersal in a sea of gene flow**". We also want to thank **Eric Crandall and the two anonymous reviewers for the constructive and competent comments, which contributed to increase the quality of our manuscript. As indicated below, we have addressed all the comments provided by the three reviewers.**

In particular, we now add more biological information on our species and on oceanographic models. We justify the use of a continuous sampling, and we clarify the link between parentage and relatedness methods.

Point-by-point responses to all comment are provided below in bold. We also enclose a version of the manuscript with all changes highlighted (line numbers in the response letter refer to the revised manuscript with track changes).

We hope that this revised version will fulfil your expectations and meet the *Proceedings* requirements.

Sincerely,

Laura Benestan,
On behalf of all co-authors

Reviewer(s)' Comments to Author:

Referee: 1

The authors conduct continuous sampling of four marine species along 1,000 km of the Northwestern Mediterranean coast, and characterize each individual's genome using RAD methodology. They then measure isolation-by-distance in each species across increasing spatial scales, and compare this to data simulated under a variety of dispersal kernels and parameters, finding that these simulations generally yield results similar to the significant IBD detected in two of the four species. Greatest similarity is found with simulations that include the possibility of long-distance dispersal, two events of which they also empirically observed in their data. Lagrangian simulations of larval dispersal underscore the importance of local dispersal with occasional longer-distance events.

Overall, this is a well-conceived and well-executed study that significantly extends themes from Benestan's, Puebla's and Manel's prior research and will make a nice contribution to this journal and to seascape genetics in general. In general, the analyses are sound and well-developed and described. My only major issue is that there is some important information missing regarding biological and geographical aspects of the system, as well as regarding the population genetic simulations, which I mention in my specific comments.

Response: We thank Dr. Crandall for his positive feedback. We agree that more information could be provided on the biological and geographical background of the study and on the simulations. As detailed below, this is now done.

Specific Comments

L58 - Not sure Waples and Gaggiotti is the best citation for this, as they didn't really explore marine parameter space. What about, in addition, Gagnaire et al. 2015?

Response: Agreed, thanks for the suggestion. We now only cite Gagnaire *et al.* (2015) and Hauser *et al.* (2008) (line 56).

L77 - New paragraph, and even then, this is a bit of an abrupt transition. I invite the authors to provide more relevant oceanographic and geographic information about the region and species sampled.

Response: We agree that more biological, geographic and oceanographic background were needed here. We added biological information on the four-study species (lines 96-104 and Table S1). The geographic and oceanographic contexts are now presented at the beginning of the paragraph to

make the transition less abrupt (lines 92-96). The authors on this study include experts on each of the four species, yet surprisingly little is known about these common and widespread species.

L78 - Please provide more a priori justification for your continuous sampling scheme for the reader either here or in the methods.

Response: We agree that this is an important point that requires to be more explicit in the manuscript, and that was also asked by Referee 3 (main comment). We now mention in the Methods that “*our continuous sampling provides the opportunity to explore IBD across a continuum of spatial scales, from just a few km to close to 950 km*” (lines 109-110). Furthermore, we now mention in the introduction lines 75 to 78 that “*although the norm for the study of marine organisms, a population-level approach can generate biases and artefacts when the sampling design artificially discretizes species distributions that are much more continuous in reality.*” Here we added the following references: Prunier *et al.* (2013), Luximon *et al.* (2014) and Battey *et al.* (2020).

L83 - Please provide more information about these diverse life-history traits beyond PLD, such as some estimate of reproductive output per female, reproductive mode (brooding, spawning, etc.), reproductive seasonality, larval behavior, recruitment conditions etc. Probably in a table in the manuscript or supplement.

Response: As detailed above, more biological information on the four-study species is now provided in Table S1 and in the Introduction (lines 98-106).

Also, what if anything is known about the spatial distribution of these species, or their census population sizes in the region? How close do they come to the pseudo-continuous distribution in the IBD model?

Response: We now mention that: “*the four species are reported over the entire study region (Albouy et al. 2015)*” (line 98). To the best of our knowledge, population distribution, sizes and densities have not been documented in detail for the four species in the study area. Again, the authors on this study include experts on each of the four species, yet surprisingly little is known about these common and widespread species. We now add a figure in the electronic supplementary material that shows the distribution of geographic distances between pairs of samples for each species (Fig. S1). We mention in the caption that the distributions present hills and valleys, which suggests here again that densities are not homogenous, but that we were nonetheless able to cover a continuous range distances from 1 to 950 km.

L109 & L121 - I know DArTSeq is proprietary, but a little more information about library prep and bioinformatic processing would be helpful. I believe it is basically ddRAD?

Response: Correct, we now precise that DArTSeq is a variant of double-digest RAD sequencing (line 129) and refer to Killian *et al.* (2012) for the details of the method (line 134). Bioinformatic processing is detailed in the next paragraph and a table indicating each filtering step is presented in the electronic supplementary material (Table S2).

L127 - Please cite the 1995 paper

Response: Thanks, we now add the Loiselle *et al.* (1995) reference (line 151).

L154 - Readers will need more detail about the model here:

- 1) In IBDSim individuals give birth to thousands of gametes, which are subject to the stated emigration rates, but this is not obvious in the manuscript.
- 2) What happens when a node is occupied by multiple individuals following dispersal?
- 3) Were samples taken before or after dispersal?

Response: We agree that more details are needed on the simulations. We now describe the life cycle in the simulations: “*Each individual produces a large number of gametes and dies. Mutation occurs in the gametes, which then disperse and fuse into diploid individuals. Competition brings the number of individuals in each node back to one, and individuals are sampled after the last dispersal step*”. (lines 178-181)

- 4) How many samples were taken and with what spacing? This question is especially important to understanding how the results compare to the empirical data

Response: We now precise that “three hundred samples were taken linearly, i.e. following a line in the center of the lattice, with single spacing between adjacent samples”. (lines 188-189)

L175 - Stokes is capitalized as it is referring to George Stokes

Response: Thanks, Stokes is now capitalized here and throughout the manuscript.

L201 - Give yourselves some credit! Use the active voice.

Response: This statement was deleted from the Results because the number of samples that we collected is already presented in the Methods and we now clarify the number of samples that were collected *versus* genotyped, *versus* considered after filtering in response to another reviewer’s comment below. Nevertheless, we now use the active voice in the Methods when presenting the number of samples that we obtained: “*We obtained a total of 2,043 samples that consisted of clips from pectoral fins (fishes) or pleopods (lobster)*” (line 115-116).

L208 - Would be helpful to see a global F_{ST} for each species.

Response: We agree, yet in the absence of clustering in three species – *Diplodus sargus*, *Mullus surmuletus* and *Palinurus elephas* – we cannot estimate a global F_{ST} unless we subjectively divide the samples into groups. The exception is *Serranus cabrilla*, for which we did identify genetic clusters and present the (global) F_{ST} between the two groups (0.021, line 2346).

L212 - State which has two species here.

Response: We now precise that this refers to *M. surmuletus* and *S. cabrilla* (line 238).

L236 - I guess you mean Table S4? Can you put some indicator of per simulation significance in this table?

Response: Correct this is Table S4 not S6, thanks for pointing this out this is now fixed. As for significance you mean statistical significance of the isolation by distance slope? If so this cannot be represented by a single value because this test is repeated at all spatial scales, i.e. 60 times per simulation. Presenting this would imply presenting 192 panels (one per simulation) like the ones we presented in Fig. 3, which would be very heavy. We decided to present instead the detailed results of four representative simulations in Fig. 3. We now precise in the caption of Fig. 3 lines 606-608 that this selection covers the entire “*range of the results that were observed in the simulations, from a slight increase in isolation by distance at small spatial scales and a pattern that is never significant (a) to a marked increase in isolation by distance at small spatial scales and a pattern that is significant at all spatial scales (d)*”.

L241 - I don't see any data at all in panel 3A? And very little in 3B?

Response: Sorry the symbols were too small and did not show in the pdf conversion, this is now fixed.

L273 - The presence of a coastal boundary layer would also be expected to reduce overall dispersal distances (Nickols et al. 2012)

Response: Thanks for pointing to this reference. The biophysical simulations implement a no-slip boundary condition that is similar in scale to the findings of Nickols et al. 2012. We now mention that this is expected to reduce overall dispersal distances (line 206-207).

L293 - Pinsky et al 2017 bridged this divide first, and while they are cited in the previous sentence, I think they deserve more explicit mention here.

Response: Agreed, we now mention explicitly that “*A step in this direction was taken by Pinsky et al., who showed that dispersal estimates based on parentage analysis are in agreement with those based on IBD in the clownfish Amphiprion percula.*” (lines 322-324)

L300 - Crandall et al. 2012, sampling at larger spatial scales and using coalescent analysis, found evidence of IBD in a snail species with PLD equivalent to that in *P. elephas*.

Response: Thanks for pointing this out. We now add a statement in the Discussion that refers to this study (lines 355-356).

L310 - This information, preferably with citations, would be useful at the end of the introduction.

Response: Agreed, this statement was moved to the beginning of the Methods section where we now explicitly state the implications of our sampling design: “*The fact that we sampled adults implies that our genetic data integrate the effect of dispersal at both the larval and adult stages. This is particularly relevant in our case since the habitats are more or less continuous, which implies that adults can potentially move throughout these habitats*” (lines 112-115), citing Hernandez-Garcia *et al.* (2015), Planes & Lenfant (2002), Abecasis *et al.* (2013) and Goñi & Latrouite 2005.

L315 - Crandall *et al.* 2019 used a coalescent sampler (Migrate) to detect evidence for stepping-stone dispersal when IBD was not evident. They found clear support for a stepping-stone model in 18 of 26 species along the Hawaiian Islands, where IBD had only been found in 4 species previously.

Response: Thanks for pointing out this relevant study, which nicely complements our results. We now state that “*this outcome complements the results of Crandall *et al.* (2019), who provide evidence of stepping-stone dispersal in a variety of marine species*” (lines 355-356).

L320 - Well put. Overall, a well-written and comprehensive discussion.

Response: Thanks!

L247 - While the archiving described here might meet the minimum standards of the journal, may I suggest that the authors submit their sampling metadata to GEOME (geome-db.org; Deck *et al.* 2017). GEOME provides two easy to use ways to then upload the raw genomic reads to INSDC SRA, and can also link in the SNP calls on Dryad. Doing so will facilitate FAIR access to these data for future researchers (Wilkinson *et al.* 2019)

Response: Thanks for your suggestion, the sampling metadata was submitted to GEOME under the Reservebenefit project name.

Figure 4 - Why aren't results for the lobster shown?

Response: The biophysical model “*was not run for the lobster because the drifting parameters (i.e., active vertical migration, resulting drifting depth) were not known to the degree of certainty needed to be reasonably confident about the simulation results*”. Extensive sensitivity experiments that are beyond the scope of this study would have been needed. This is now stated in the text (lines 217-219).

Fig S1B - What are the units of the color scale, and could a clearer scale be used (white->blue or whatever)?

Response: The units of the color scale are now indicated and the colors were changed. See new figure (Fig. S3).

Fig S2 - *S. cabrilla* is difficult to read

Response: The colors and layout were changed in the ADMIXTURE plot. See new figure (Fig. S3).

Fig S6 - If I'm reading this correctly, the scale of significant IBD slopes is much greater in this species than those in the current manuscript. Might be worthy of discussion?

Response: Correct, we now state that “*the strongest IBD pattern in terms of both slope and statistical significance was observed in the *P. dentatus* data from the Northwest Atlantic (Hoey & Pinsky 2018) that we reanalyzed. This species is characterized by a strong homing behavior (Sackett *et al.* 2007) and high residency on spawning grounds (Desfosse 1995), where the samples were collected. This illustrates the fact that our approach, which is based on the sampling of adults, integrates both larval and adult movement. It is therefore preferable to sample during the spawning season when it comes to detect IBD*” (lines 347-352)

Eric Crandall

October 20, 2020

We thank Dr. Crandall for the constructive feedback on our study.

Referee: 2

In this manuscript, the authors investigate isolation by distance at different geographic scales using genetics, computer simulations and a biophysical model. The topic of the manuscript is interesting, and the results are relevant, with implications for ecology, evolution and conservation. Strengths of the manuscript include the use of a null model to understand the IBD results and the diversity of methods used to investigate the question.

Response: We thank Referee 2 for the positive feedback on our manuscript.

Below, we provide suggestions for improvement.

1. We recognize that Proceedings B manuscripts are relatively short, but we felt that additional information about the four-study species was needed earlier on in the paper. This information not only provides the relevant context for the hypotheses, but also provides key details for interpreting the biological importance of the authors' findings. For example, information on what is known about dispersal distances, adult spawning timing and location for each species was generally lacking, or provided too late, which made it challenging to understand the rationale for the hypotheses and to put the results within the appropriate context. This would also help to flesh out the 'diversity of life history traits' that the authors mention, rather than only focusing on PLD.

Response: We agree that more background information on the four-study species could be provided. As detailed in our response to Dr. Crandall above, this is now done in the Introduction (lines 96-104) and in Table S1. As mentioned above the authors of this study include experts on each study species, yet surprisingly little is known about these common and widespread species.

2. Are any of the four species restricted to certain depth ranges within the study region, and was this accounted for when geographic distance was calculated using the bathymetry data for the IBD analysis? An explanation of why Loiselle's kinship coefficient was used for IBD calculations would also be helpful.

Response: We agree that this is important information that should be presented. We now provide the minimum and maximum depths of the four species in Table S1. They are all found in shallow waters, although *Serranus cabrilla* can extend significantly deeper (down to 500 m) than the other three species. We now also precise that all samples were collected at shallow depths (on average 46 m, range 1.8-693.0m; line 111). Finally, we tested the effect of depth on genetic variation with a distance-based Redundancy Analysis and found no effect. This analysis will be presented in a separate study on adaptive genetic variation (Benestan et al. *in prep*). As for the choice of the Loiselle kinship coefficient we now state that it was chosen because it uses a correction for small sample size and tends to show less bias than other coefficients in this case (Wang 2017), and that it correlates well with five other coefficients that we tested on a subset of the data (Table S3, lines 152-153).

3. The authors are using pelagic larval duration (PLD) as a proxy for dispersal distance, but PLD does not actually have a strong relationship to dispersal distance. We were confused about its inclusion. The authors need to explain why this is relevant to the study, especially if adults of each species were sampled. Is dispersal during the larval stage thought to be the dominant dispersal mode for each of these species (see point #1)?

Response: We agree that PLD is not the strongest predictor of dispersal and this is now explicitly stated in the Introduction, citing Selkoe & Toonen (2011) (line 104). Furthermore, we de-emphasized PLD and don't refer to it in our hypotheses anymore. We still mention PLD though because although not the strongest predictor of dispersal it is still relevant in the context of dispersal, particularly when it is extremely long as in *P. elephas*. We also now clarify that the fact that we sampled adults implies that "our genetic data integrate the effect of dispersal at both the larval and adult stages" (lines 112-113). We now add that "This is particularly relevant in our case since the habitats are more or less continuous, which implies that adults can potentially move throughout these habitats" (lines 113-115). We finally expand on adult movement in the Discussion (lines 342-346).

4. In general, the methods were well documented, but additional details for certain analyses would help readers to follow along more easily. In particular, the biophysical model methods seemed incomplete (further details in point #7). Also, the population structure methods appear to be missing from section 2d.

Response: We now provide more information on the biophysical model and as detailed below we now follow your suggestions (lines 209-224). The population structure methods were hidden in the population structure supplementary figure caption, we now present this in the Methods instead (lines 148-149).

5. Since samples came from the artisanal fisherman, are the samples from mature/adult individuals that were captured on spawning grounds? The authors state that *D. sargus* were not captured on spawning grounds in lines 306-308, so it seems like this might be true of the other species too? Sampling individuals when they are not on their spawning grounds could miss true population structure and impact IBD results. Related to this, *P. dentatus* were captured on their spawning grounds, so at least some discussion of how sampling during different parts of a species' lifecycle might impact IBD seems warranted.

Response: Very good point, we now expand considerably on this in the Discussion: “*We suggest that this may be due to the known migratory behavior of D. sargus, which spends the majority of the year in genetically heterogeneous populations before returning to specific and distinct areas for reproduction. Sampling was mostly conducted between June and November, almost exclusively outside of the D. sargus spawning period (March-June), when migration may have contributed to obliterate the IBD signal, and the same is true for P. elephas. The overlap between sampling and spawning season was greater for M. surmuletus and S. cabrilla (June-July), the two species for which we did detect a significant IBD signal. Furthermore, the strongest IBD pattern in terms of both slope and statistical significance was observed in the P. dentatus data from the Northwest Atlantic that we reanalyzed. This species is characterized by a strong homing behavior and high residency on spawning grounds [68], where the samples were collected. This illustrates the fact that our approach, which is based on the sampling of adults, integrates both larval and adult movement. It is therefore preferable to sample during the spawning season when it comes to detect IBD.*” (lines 339-352)

For Figure 1, the sampling is so dense in some areas that it is challenging to see on the maps. Is each unique lat/lon coordinate representative of a single individual? If it is not, the authors might consider using point area to indicate the number of individuals captured at each lat/lon coordinate and/or to make points semi-transparent.

Response: We now clarify that we sampled several individuals per GPS position when possible to increase our power to detect genetic structure (lines 119-121) and Fig. 1 was updated to provide the number of samples per site.

6. The author's definitions of short, regional and long-distance dispersal in the manuscript seemed arbitrary, despite the citation of papers in which these were defined.

Response: We now simply use the term “large” spatial scale and define it as >100 km, as opposed to “small” spatial scales (few tens of km). As for long-distance dispersal we follow Manel *et al.* (2019) and define it as dispersal events beyond 40 km.

The authors report that their results were not sensitive to the inclusion of closely related individuals, but were siblings directly tested for (i.e. Colony, etc)?

Response: This was not done in our original manuscript but we now use Colony to identify closely related individuals, thanks for the suggestion. The same pairs of individuals that were detected as related by the Loiselle coefficient - 2 pairs in *M. surmuletus* and 7 pairs in *S. cabrilla* - were detected as full and half sibs by COLONY (lines 284-286).

7. The author's use of the biophysical model left many questions. The authors cite the Bode et al 2019 study, but also see North E, Gallego A, Petitgas P. Manual of recommended practices for modelling physical-biological interactions in fish early life. Copenhagen: International Council for the Exploration of the Sea; 2009 2009. 1-112 p. It would be helpful if the authors revisited the biophysical model section in the context of what is known to be important for capturing dispersal. For example, are these species demersal spawners? If so, then were particles released on the surface or from the benthos? Why were only 100 particles released? There are limitations to what can be reasonably implemented, but some justification for choices is needed. In general, the biophysical model itself seems to distract from the novelty of this work- the investigation across spatial scales is exciting, and the biophysical model section seems ancillary.

Response: We agree that more details were needed on the biophysical model. We now mention that “*simulated larvae were seeded in eight marine protected areas from the study region that are known spawning grounds for the selected species and represent a total area of approximately 1,072 km² (electronic supplementary material, Fig. S2b). We clarify that the four species are demersal spawners (Table S1). Nevertheless, since these areas are shallow, particles were released at the surface. This was done with a uniform random spatial distribution within each marine protected area for a total of 17,160 particles (~ 100 per model grid box of 16 sq.km) per day. As a result, a total of 2,093,520 larvae were simulated for *D. sargus*, which was seeded over 122 days (March to June), and of 2,625,480 for *M. surmuletus* and *S. cabrilla*, which were seeded over 153 days (March to July, lines 209-216). The biophysical model is an important component of the study in our view because it complements the population genomic analyses with a completely different perspective with a view on larval dispersal specifically. It supports, for our study species and in our study area in particular, our hypothesis that most dispersal occurs over small distances with rare long-dispersal events.*”

8. The authors could do better to contextualize their study within the dispersal literature, to further discuss why IBD at different spatial scales matters and to discuss the scale at which IBD should be examined. In general, we found the introduction and discussion to be undercited. For example, a huge body of literature has gone into the statement: ‘The marine environment is also more homogenous than terrestrial and freshwater environments, and presents fewer geographic barriers to the movement of organisms,’ (lines 49-50) yet no citations were provided.

Response: As indicated above (answer to point 5) and below, we now better contextualize our study within the dispersal literature. We now cite McCauley *et al.* (2015) and Grummer *et al.* (2019) to back up the statement that the marine environment is more homogenous than terrestrial and freshwater environments as well as presenting fewer geographic barriers to the movement of organisms (lines 47-49).

Figure 3. In the figure caption, we suggest pointing out that the dispersal function utilized for each panel is in the inset. What do #75, #76, #77 and #106 mean in the insets?

Response: Thanks for pointing out this. We have now precise in the Fig. 3 caption that the dispersal kernel used for each panel is illustrated in the inset and that the numbers 75, 76, 77 and 106 refer to the simulation ID in Table S4 where the details of the simulation parameters are presented. See updated figure.

Line 34. It is unclear as written why discrete or heterogeneous sampling limits regional scale inference.

Response: As indicated above, we now mention that although the norm for the study of marine organisms, “*a population-level approach can generate biases and artefacts when the sampling design artificially discretizes species distributions that are much more continuous in reality*”, citing Prunier *et al.* (2013), Luximon *et al.* (2014) and Battey *et al.* (2020). (lines 76-78)

Line 48. Limited citations given. Line 49. No citations given.

Response: As indicated above we now cite McCauley *et al.* (2015) for the first statement and Grummer *et al.* (2019) for the second one (lines 47-49).

Line 64. The examples given are fishes, but the literature is not only fishes. We suggest eliminating this clause and citing the existing literature more.

Response: We now eliminate “mostly fishes” and we add citations that include both fish (e.g. Bode *et al.* 2019) and non-fish studies (e.g. Mokhtar-Jamai *et al.* 2013; Coates *et al.* 2014).

Line 67. The use of “unprecedented” should probably be tempered. In terms of spatial scale, the paper cited later, Hoey & Pinsky 2018 was a larger spatial extent, and in terms of direct and indirect dispersal estimates, please see the following work in addition to D’Aloia *et al.* 2015
D’Aloia, C. C., *et al.* “Seascape continuity plays an important role in determining patterns of spatial genetic structure in a coral reef fish.” *Molecular Ecology* 23.12 (2014): 2902- 2913.

Response: Agreed! We now replace “unprecedented” by “continuous” (line 81). We are aware of the 2014 D’aloia paper, we did not cite it because we did not observe the same effect of habitat discontinuity in our data (i.e. samples from the Balearic Islands are not genetically differentiated).

Line 96. Replace ‘conserve’ with ‘preserve’

Response: Thanks, this is now changed.

Line 118-119. Are these previous de novo references or full genome references? Regardless, pointing readers to where these references exist in the literature will be helpful.

Response: We now clarify that these are *de novo* reference genomes that were assembled for this study specifically. The paper describing these reference genomes was not published when we submitted our initial manuscript but it is now out and we now refer to it (Fietz *et al.* 2020; line 130, line 140).

Line 127. More details on settings and implementation of GENODIVE would be good.

Response: As indicated above we now expand on the Loiselle coefficient implemented in GENODIVE (lines 150-151). There are no particular parameters or settings for this analysis.

Line 133. It is unclear what “adjusted” distance required to keep those individuals in the analysis’ is. So, there was a criteria cutoff?

Response: We now clarify this point. (lines 159-161)

Line 157. Please provide more details on the implementation of IBDSim. In general, these types of details will be helpful for readers to understand assumptions and also for future researchers reading this for methods.

Response: As indicated above, we now briefly describe the life cycle underlying the genetic simulations (lines 178-181): “Each individual produces a large number of gametes and dies. Mutation occurs in the gametes, which then disperse and fuse into diploid individuals. Competition brings the number of individuals in each node back to one, and individuals are sampled after the last dispersal step.” In addition, we now provide the full set of parameters for one exemplary simulation in the Supplementary Material (Text S2). The parameters that were changed in the other simulations are listed in Table S4.

Line 180. Parcels allows one to?

Response: Thanks, this is now changed as suggested.

Line 201. Where did so many samples get lost? We are told 2,034 samples to start, and 1,299 individuals are in the analysis?

Response: We now clarify in the Methods that a total of 2,043 samples were sampled. Among these, 1,305 samples were selected for genotyping (280 for *D. sargus*, 312 for *M. surmuletus*, 245 for *P. elephas* and 468 for *S. cabrilla*) (See lines 249-252). “This selection was made to maximize the ratio of continuous spatial coverage over sequencing effort, while also considering several samples per GPS position when possible to increase our power to detect genetic structure” (lines 119-121). We also clarify in the Results that “six samples were discarded due to a high proportion (>30%) of missing data, resulting in a final dataset of 1,299 individuals” (line 229-230).

Line 205. “Relative” to what?

Response: This term was removed.

Line 215. Readers need much more information about these related individuals. It would be nice to see the authors try to parse out what these relationships are, especially because the authors base the “long distance dispersal” results on these relationships. Are they siblings? In which case, how are the authors inferring a dispersal trajectory?

Response: As indicated above, we now present the results of a sibship inference analysis using Colony, which classified these two pairs of individuals as half sibs (probability = 1.00). We are not able to provide a dispersal trajectory though, because this category refers to individuals who share alleles from one parent so they could represent not just siblings but also e.g. uncle-niece or grandmother-grandson relationships. Nevertheless, we know that “the distances of 292 and 294 km would have been covered in one or two generations” (line 288).

Line 217-222. This should be moved up to line 213 before the authors start talking about highly related and geographically close individuals.

Response: We tried but it didn't work well because the highly related and geographically close individuals are relevant when IBD is detected.

Line 254. This section is not substantial because there isn't life history information with which to contextualize dispersal for the species, the authors have not inferred the relationship (just close kinship, which in itself does not give information regarding dispersal), and the arbitrary use of "long distance" when the literature has given examples of how to define "long distance dispersal".

Response: As indicated above we now present life history information in the Introduction (lines 96-104) and in Table S1, define long-distance dispersal citing the literature, and confirm the relationship between the closely related samples using COLONY.

Line 273. The inferences made from this "secondary peak" are hard to understand when only 100 particles were released, meaning that only a small fraction of possible dispersal trajectories was sampled with this model.

Response: As detailed above, we now clarify that up to 2,625,480 particles were simulated: "Simulated larvae were seeded in eight marine protected areas from the study region that are known spawning grounds for the selected species and represent a total area of approximately 1,072 km² (electronic supplementary material, Fig. S2b). Since these areas are shallow particles were released at the surface, with a uniform random spatial distribution within each marine protected area for a total of 17,160 particles (~ 100 per model grid box of 16 sq.km) per day. As a result, a total of 2,093,520 larvae were simulated for *D. sargus*, which was seeded over 122 days (March to June), and of 2,625,480 for *M. surmuletus* and *S. cabrilla*, which were seeded over 153 days (March to July)." (lines 209-216).

Line 300. See general comments about the PLD framing.

Response: As indicated above, we now highlight the limitations of PLD as a predictor of dispersal (line 104) and de-emphasize dispersal in our hypotheses. We still mention PLD though because although not the strongest predictor of dispersal it is still relevant in the context of dispersal, particularly when it is extremely long as in *P. elephas*.

Lines 304-308. Readers will be looking for this sort of information in the Introduction or Methods, so you might consider mentioning and flushing out relevant details much earlier

Response: Agreed, the migratory behavior of *D. sargus* is now mentioned upfront in the Introduction and the spawning season of each species is detailed in Table S1.

Line 311. The point being made is unclear, and some life history information would help the readers understand better.

Response: As indicated above, we now provide more life history information in the Introduction (lines 96-104) and in Table S1. We now bring this point earlier and clarify that our genetic approach integrates the effect of both larval and adult dispersal while the biophysical model is meant to simulate larval dispersal only.

Line 322. This distance seems very arbitrary. Some justification is needed.

Response: We now clarify that this distance is not arbitrary by providing the exact distance between the two pairs of samples (292 and 294 km, respectively).

Line 323. Defining local vs regional in terms of the spatial scale of this study would be helpful.

Response: As indicated above we now use instead "small" (few tens of km) and "large" (>100 km) to refer to the two spatial scales.

Lines 328-336. This brings up the interesting question of scale. For conservation practices, it might be important to create a network of nearby MPAs. On the other hand, perhaps larger scales are more helpful for understanding population structure throughout a species range.

Response: Good point, but we now remove this entire paragraph from the discussion.

Line 330. Another example of a statement not supported by a citation.

Response: This entire paragraph was removed from the discussion.

Fig S2. Sorted by latitude going from left to, right?

Response: Correct, this is now clarified in the figure caption.

Fig S4. Are the significant slopes plotted on here?

Response: Thanks for pointing this out, this was missing in our original submission. The slopes that are significant are now highlighted in bold and this is clarified in the figure caption.

Fig S6. The use of clipart in the figures was great! However, the flatfish in Fig S6 is likely a European flounder and not a summer flounder.

Response: Thanks for noticing this! The figure was changed (Fig. S7).

Referee: 3

In this study, Benestan et al. explore how IBD patterns change across spatial scales in four co-distributed marine species using impressive genomic data sets. They supplement these empirical analyses with IBD simulations and biophysical models of dispersal. One of the main findings of the study is that dispersal can be limited even when there is little to no genetic structure at regional spatial scales. I enjoyed reading this paper. I found the analyses to be sound, the writing to be generally clear, and the figures to effectively visualize the key results. I also agree with the authors that researchers are often too quick to state that extensive dispersal prevails in the marine environment, based on low levels of genetic structure in most marine species with a pelagic larval state. This paper helps to clarify why such a simple statement is not always true.

Response: We thank Referee 3 for the positive feedback on our manuscript

However, I am unconvinced by some of the framing/motivation of the paper--- namely, that there is an existing disconnect between population genetic data and the results of direct parentage studies. I provide more details below along with a few other comments.

- In the Introduction, the authors ask whether evidence of restricted dispersal revealed by parentage studies can be reconciled with low genetic structure observed in marine species (L 64-66). This question is used as the primary motivation for the study. This question implies that previous empirical work has revealed a disconnect between these two approaches, but I am not convinced this is true. There are a few important points to note here:

First, within parentage studies, there has been a progression over the course of several decades towards empirical estimates of dispersal kernels (see Almany et al. 2013, 2016; D'Aloia et al. 2013, 2015; Williamson et al. 2016 for examples). All of these studies have documented leptokurtic dispersal kernels in marine fishes—that is to say, they all provide evidence of short distance dispersal, but they also estimate that there is a lower level of long-distance dispersal (with the average scale of dispersal varying across study species). The notion that most dispersal is limited, but that some connections occur over larger spatial scales, is therefore well-established in the species that have been the subjects of large parentage studies. I found this manuscript to overlook this major finding from genetic parentage studies (e.g it is missing from the Discussion on lines 320-327).

Second, in some of the few fish species that have been intensively studied via parentage analysis, there are complementary population genetic analyses, and these findings are largely congruent. For example, in the clownfish *Amphiprion percula*, dispersal estimates based on parentage work are remarkably similar to those based on IBD demonstrating not only that both datasets can provide consistent dispersal estimates, but also that short distance dispersal can occur despite “extensive gene flow among populations” (Pinsky et al. 2017). In another example, D'Aloia et al. (2020) explore how parentage-based dispersal estimates can help explain population genetic patterns at different spatial scales in the goby *Elacatinus lori*.

In light of these comments, I suggest that the authors should (1) incorporate these earlier findings into their Introduction and Discussion and (2) articulate the major contributions of this paper, given this earlier work. I do not think the novel contribution here is reconciling population genetic approaches with direct parentage approaches as the Abstract implies.

Response: We agree with Referee 3 that the study could be better framed and embedded in the existing literature. We now expand on the insights provided by parentage analyses and slightly reframe our study in this context in the introduction (lines 62-75) to address reviewer points 1 and 2:

“Direct approaches to investigate larval dispersal such as mass mark-recapture or parentage analyses support the predominance of restricted dispersal in a variety of species and indicate that dispersal kernels are often a rapidly decreasing function of distance, with little dispersal beyond a few tens of kilometers [16–23]. This suggests that long-distance dispersal, defined here as dispersal beyond 40 km [24], is rare. Nevertheless, parentage analyses are challenging to scale up because the number of samples that need to be genotyped rises disproportionately with increasing spatial scale (but see e.g. [16,23]). They also provide a snapshot of dispersal over just one generation, and have therefore to be repeated for every new generation analyzed (e.g. [18, 26]). On the other hand, next-generation sequencing provides the opportunity to generate robust estimates of relatedness between any pair of individuals by genotyping a large number of Single Nucleotide Polymorphism (SNP) markers [25]. This approach can be applied at any spatial scale and allows to go deeper in the recent past by identifying relationships that are more distant than parent-offspring. It also offers the possibility to work at the level of individuals in continuous space, as opposed to the population-level analyses that are typically conducted at large spatial scales.”

We also made an effort to better put our study in a broader context in the discussion (lines 321-326):
*“In this sense our approach empirically bridges the results provided by large-scale population genetic [58] and small-scale dispersal [59,60] studies. A step in this direction was taken by Pinsky et al. [57], who showed that dispersal estimates based on parentage analysis are in agreement with those based on IBD in the clownfish *Amphiprion percula*. Nevertheless, this study considered IBD at the population level and at a fixed spatial scale of 200 km [57]. Our continuous approach shows that IBD can develop at substantially smaller spatial scales.”*

Finally, we replaced “reconcile” by “bridge” in the abstract (line 43).

- The authors should clarify what they mean by “continuous” sampling designs and provide more details on their own sampling regimes. Can the authors provide some summary of the average sampling intervals for each species? Looking at Figure 1, there is clearly variability in continuity of sampling across the 4 species. For example, *D. sargus* could be viewed as 3 main clusters of sampling, whereas *M. surmuletus* is more continuous throughout the study area. It is also unclear to me what each circle represents in Figure 1. Is it one individual per site, or can it be multiple? I did not find this information in the supplementary materials so at a minimum the mean and s.d. (and perhaps range) of # samples per site should be reported if it is not one individual.

Response: These are all good points. As indicated above (point 5, Referee 2), Fig. 1 was updated to illustrate the number of samples per site. We also now present the frequency distribution of the geographic distance between pairs of samples for each species (Fig. S1). The distributions present hills and valleys, which suggests that densities are not homogenous across the study area, but we were nonetheless able to cover a continuous range distances from 1 to 950 km.

- Another suggestion regarding terminology is to be clearer about the definition of a “regional” spatial scale. I wondered whether this is a clear spatial definition, or whether it may depend upon the species. For example, in the extreme, what is regional versus local scale for a cosmopolitan species versus an endemic?

Response: We agree that this term can be confusing. We now use “large” spatial scale instead and define it as >100 km, as opposed to “small” spatial scales (few tens of km).

- Given that this is a short-format paper the Discussion could be improved by staying focused on the major implications of the study. The final two paragraphs (out of only 5 total) seemed to stray from the focus of the paper (especially the text on uncovering the genomic architecture of adaptation). I think a brief mention of the implications for MPA design is warranted, but otherwise I suggest removing most of this text. This would give you space to discuss your results and how it relates to the large body of existing work on marine dispersal.

Response: Agreed, the section on genomic architecture of adaptation was removed and we now expand instead on the major implications of the study.

-One area that I felt deserved more attention is further explanation for why IBD disappears at large spatial scales. There is a reference to another theory paper in the Introduction (L73), but more discussion in the present manuscript is warranted since this was a key result in two of your study species. I understand why IBD signals are obscured when there is strong genetic structure, but could you provide more explanation for why this happens in the absence of structure? Your IBD simulations show that the decrease in IBD at

large spatial scales is not necessarily because of long distance dispersal, as this pattern can emerge in the stepping stone simulations. So, what else could explain it?

Response: Good point, we now suggest that “*the decrease of IBD at large spatial scale may result from the fact that it takes more time to reach the mutation-drift equilibrium at large spatial scale (Slatkin 1993; Hardy & Vekemans 1999), and that factors other than dispersal (e.g. demographic history or selection) may shape genetic variation at larger spatial scales (lines 329-331)*”.

- Finally, I was unable to access the data via the dryad link in the manuscript.

Response: Sorry about that, the link is now available at

<https://datadryad.org/stash/share/lkYLCqfssPINOetqZyEPSnW1Fth61u6r9CFKI7Gspw>.

L56: “It can persist...” is an ambiguous language. Please reword to clarify whether “it” refers to “weak genetic structure”

Response: This is now clarified.

L166: “Samples were taken linearly” ... I found this confusing. Could you please rephrase to make this clearer?

Response: We tried to clarify these lines 188-189 as “*Three hundred samples were taken following a line in the center of the lattice, with single spacing between adjacent samples.*” In addition, we now present an exemplary IBDsim input file in which the exact position of the samples is indicated: see Text S2.

L204-206: This wording should be revised, as there is genetic structure in one of your study species.

Response: Agreed, this is now clarified lines 232-234 as “*Large-scale population genetic structure was absent in three species (D. sargus, M. surmuletus and P. elephas) and weak in one species (S. cabrilla)*”.

L257-258: These are clearly not parent offspring pairs (kinship coefficient would be ~0.25) so reference to such pairs should be removed from this sentence.

Response: Agreed, as detailed above this is now clarified with the Colony analyses.

L302: Remove “nevertheless” from the sentence as it adds confusion.

Response: Done.

L309 and remainder of this paragraph: Suggest replacing “adult dispersal” with “adult migration” or “adult movement.” Dispersal has a specific definition, and I don’t think the type of movement covered here is referring to dispersal.

Response: Agreed, we now use “*adult movement*” (see line 351).

We thank Referee 3 for the constructive feedback on our study.

References:

- Abecasis, D., Bentes, L., Lino, P. G., Santos, M. N., & Erzini, K. (2013). Residency, movements and habitat use of adult white seabream (*Diplodus sargus*) between natural and artificial reefs. *Estuarine, Coastal and Shelf Science*, 118, 80-85.
- Albouy C *et al.* 2015 FishMed: traits, phylogeny, current and projected species distribution of Mediterranean fishes, and environmental data. *Ecology* (doi:10.1890/14-2279.1)
- Battey, C. J., Ralph, P. L., & Kern, A. D. (2020). Space is the place: Effects of continuous spatial structure on analysis of population genetic data. *Genetics*, 215(1), 193-214.
- Bode, M., Leis, J. M., Mason, L. B., Williamson, D. H., Harrison, H. B., Choukroun, S., & Jones, G. P. (2019). Successful validation of a larval dispersal model using genetic parentage data. *PLoS biology*, 17(7), e3000380.
- Coates, J. H., Hovel, K. A., Butler, J. L., & Bohonak, A. J. (2014). Recruitment and recovery of pink abalone (*Haliotis corrugata*) in a historically overexploited kelp forest: Are local populations self-sustaining? *Journal of experimental marine biology and ecology*, 460, 184-192.
- Desfosse JC. 1995 Movements and ecology of summer flounder, *Paralichthys dentatus*, tagged in the southern mid-Atlantic bight. 187 p.

- Fietz, K., Trofimenko, E., Guerin, P. E., Arnal, V., Torres-Oliva, M., Lobréaux, S., ... & Puebla, O. (2020). New genomic resources for three exploited Mediterranean fishes. *Genomics*, 112(6), 4297-4303.
- Gagnaire P, Broquet T, Aurelle D, Viard F, Souissi A, Bonhomme F, Arnaud- Haond S, Bierne N. 2015 Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evol. Appl.* 8, 769–786. (doi:10.1111/eva.12288)
- Goñi, R., & Latrouite, D. (2005). Review of the biology, ecology and fisheries of *Palinurus* spp. species of European waters: *Palinurus elephas* (Fabricius, 1787) and *Palinurus mauritanicus* (Gruvel, 1911). *Cahiers de Biologie Marine*, 46(2), 127-142.
- Grummer, J. A., Beheregaray, L. B., Bernatchez, L., Hand, B. K., Luikart, G., Narum, S. R., & Taylor, E. B. (2019). Aquatic landscape genomics and environmental effects on genetic variation. *Trends in ecology & evolution*, 34(7), 641-654.
- Harrison, H. B., Williamson, D. H., Evans, R. D., Almany, G. R., Thorrold, S. R., Russ, G. R., ... & Berumen, M. L. (2012). Larval export from marine reserves and the recruitment benefit for fish and fisheries. *Current biology*, 22(11), 1023-1028.
- Hardy O, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity*, 83, 145–154.
- Hauser L, Carvalho GR. 2008 Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish Fish.* 9, 333–362. (doi:10.1111/j.1467-2979.2008.00299.x)
- Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., ... & Aschenbrenner-Kilian, M. (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. In *Data production and analysis in population genomics* (pp. 67-89). Humana Press, Totowa, NJ.
- Luximon, N., Petit, E. J., & Broquet, T. (2014). Performance of individual vs. group sampling for inferring dispersal under isolation-by-distance. *Molecular ecology resources*, 14(4), 745-752.
- Loiselle BA, Sork VL, Nason J, Graham C. 1995 Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* (doi:10.2307/2445869)
- Manel S *et al.* 2019 Long-Distance Benefits of Marine Reserves: Myth or Reality? *Trends Ecol. Evol.* (doi:10.1016/j.tree.2019.01.002)
- McCauley, D. J., Pinsky, M. L., Palumbi, S. R., Estes, J. A., Joyce, F. H., & Warner, R. R. (2015). Marine defaunation: animal loss in the global ocean. *Science*, 347(6219).
- Mokhtar-Jamaï, K., Coma, R., Wang, J., Zuberer, F., Féral, J. P., & Aurelle, D. (2013). Role of evolutionary and ecological factors in the reproductive success and the spatial genetic structure of the temperate gorgonian *Paramuricea clavata*. *Ecology and Evolution*, 3(6), 1765-1779.
- Planes, S., & Lenfant, P. (2002). Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. *Molecular Ecology*, 11(8), 1515-1524.
- Prunier, J. G., Kaufmann, B., Fenet, S., Picard, D., Pompanon, F., Joly, P., & Lena, J. P. (2013). Optimizing the trade-off between spatial and genetic sampling efforts in patchy populations: Towards a better assessment of functional connectivity using an individual-based sampling scheme. *Molecular Ecology*, 22(22), 5516-5530.
- Sackett DK, Able KW, Grothues TM. 2007 Dynamics of summer flounder, *Paralichthys dentatus*, seasonal migrations based on ultrasonic telemetry. *Estuar. Coast. Shelf Sci.*
- Selkoe, K. A., & Toonen, R. J. (2011). Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, 436, 291-305.
- Slatkin M (1993) Isolation by distance in equilibrium and non- equilibrium populations. *Evolution*, 47, 264–279.
- Walker, D., Porter, B. A., & Avise, J. C. (2002). Genetic parentage assessment in the crayfish *Orconectes placidus*, a high-fecundity invertebrate with extended maternal brood care. *Molecular Ecology*, 11(10), 2115-2122.
- Wang, J. (2017). Estimating pairwise relatedness in a small sample of individuals. *Heredity*, 119(5), 302-313.

Appendix C

Associate Editor Board Member

Comments to Author:

Two referees that reviewed your original submission have evaluated the revised version and agree that you have satisfactorily incorporated their suggestions.

Response: We thank the two reviewers for taking the time to go through the revised version of our manuscript.

Referee 1's comments are aimed at improving the clarity of the manuscript. Referee 2, however, make more substantial comments that need attention. In particular, he/she highlights the different library preparation used for different species, which could lead to differences in the evolutionary forces driving spatial patterns of genetic variation (DART loci are more likely to be associated to functional variation than RAD loci). A paragraph with this caveat and its potential effects on the interpretation of results should be added to the discussion.

Response: This is a good point. Indeed, ddRAD libraries were generated for *Diplodus sargus* and *Mullus surmuletus* while DART-seq libraries were used for *Serranus cabrilla* and *Palinurus elephas*. Nevertheless, since we capture a signal of restricted dispersal in *S. cabrilla* and *M. surmuletus* (i.e. one species per library type), this signal doesn't seem to be associated with library type. This point is now explicitly addressed in the manuscript at the end of the Discussion: “We detected small-scale IBD with both ddRAD-seq (for *M. surmuletus*) and DARTseq (for *S. cabrilla*) despite the fact that DARTseq targets gene-rich regions of the genome compared to ddRAD-seq (Wenzl et al. 2004). This suggests that library type was not a major factor, but we would nonetheless recommend to harmonize library types when possible.” (lines 385-389).

A very important remaining issue is that the data has not been made available. The authors provide a Dryad link but the DOI provided is incorrect. It is unclear to me if this is a genuine mistake or not but, in either case, the authors should provide the required information, which should be subsequently verified to make sure that it is correct.

Response: We sincerely apologize for this, it was an oversight on our part after a change of the dryad file and github page. The data are now accessible in Dryad with the following <https://doi.org/10.5061/dryad.0cfxpnw1s> and the github is currently publicly available at https://github.com/laurabenestan/finescale_ibd.

Reviewer(s)' Comments to Author:

Referee: 3

Comments to the Author(s).

Thanks for the thoughtful response to reviewers. I enjoyed rereading the manuscript and think you did a very nice job addressing the many comments! In this version, there is greater clarity about the four study species, the spatial scales of analysis, and how your work fits in within the very large body of marine dispersal research. I think the paper is a nice contribution. A few minor comments:

Response: Thank you for taking the time to go through the revised version of our manuscript, your comments contributed to significantly improve the quality of the study.

-I'm sure there are word limits, but in the Abstract, it seems like it would be prudent to clarify that the pattern of IBD at small spatial scales was only found in some of your study species. Documenting variability across species is one of the strengths of a multi-species study and an important conclusion.

Response: We agreed and now clarify in the abstract: “Nevertheless, our continuous sampling strategy uncovers a pattern of isolation by distance at small spatial scales (few tens of km) in two species.” (lines 36-37). The length of the abstract is now 200 words, which is exactly the limit.

-I appreciate the greater clarity about spatial scale, especially since it's an important aspect of your study. This is nit-picky, but when you talk about LDD, why is 40 km used as a cut-off? It seems arbitrary to me (including in the referenced paper) and I wonder if you need it here as it doesn't seem to really affect your analyses.

Response: The meta-analysis conducted in the referenced paper (Manel *et al.* 2019) found that **i. the median dispersal distance estimated with either parentage analysis, assignment tests or isolation by distance was 42km (interquartile range = 27–250 km), ii. the interquartile nearest neighbor between marine reserve was 4-40km, and iii. dispersal beyond 40km is poorly documented.** This 40km threshold therefore frames our motivation to investigate spatial dispersal both below and beyond 40km in the context of marine reserves and multispecies genetic structure.

-I suggest you create a new paragraph when you start talking about NGS in the introduction (L 69). I felt this part could use a little more refining. What exactly are you trying to emphasize here? I don't think it's just about relatedness. You don't necessarily need SNPs to identify relationships beyond parent-offspring. Are you trying to emphasize that it can improve genetic structure and IBD estimates? Or are you trying to get at the difference between direct and indirect genetic studies of dispersal. It's unclear. You also don't need SNP data to conduct continuously-sampled study designs (and this seems like a separate issue). I think a revision of this paragraph would help you further articulate how your study fits in.

Response: Thanks for the feedback on this important statement. It is now a new paragraph as suggested. We now also de-emphasize SNPs in particular and better articulate how coupling an individual-based sampling design with accurate relatedness estimates between any pair of individuals provides potential to estimate dispersal at different spatial and temporal scales: ***“Coupling individual-based sampling designs with robust estimates of relatedness between any pair of individuals provides potential to estimate dispersal at different spatial and temporal scales.”*** (lines 74-75).

-Fig. 1 caption: It would help to clarify that size of bubble reflects # individuals per site (and also consider including the range of samples per site). I still have a hard time figuring out how many were collected per site.

Response: We have added the following sentence to Figure 1 legend: ***“Size of circles reflects the number of individuals per site (mean = 2, range 1-31 samples). A total of 615 sites were sampled.”*** (lines 593-594)

-L 109: It would be helpful to include some of the information from the caption of Fig. S1 here. Sampling was 'continuous' but the distribution of distances was not even.

Response: Yes, but since the outcome of sampling is a result we mention this in the Results section instead: ***“The distribution of geographic distance between pairs of samples indicates that our sampling, although not uniform, covers a continuous range of distances from 1 to 950 km (electronic supplementary material, Fig. S1)”*** (lines 242-244).

-L288: You may want to clarify that 292 and 294 km are just max distances between individuals. Their “origin” may be somewhere between them, or somewhere even further away (if they're even sibs).

Response: We agree that the origin of these samples is unknown and this is now explicitly stated in the ms: ***“It should be noted, however that the origin of these samples is unknown (it could be between where they were collected or even further away)”*** (lines 303-304). The distances of 292 and 294 km are not maximum distances though since, precisely, the origin of these samples could possibly be further away than that.

Referee: 2

Comments to the Author(s).

Following revisions, this manuscript by Benestan *et al.* is much improved. The authors have made many changes that clarify their study design and findings, as well as their contribution to the literature. We have a few additional suggestions and questions to further strengthen the manuscript, but otherwise, we think the authors have done a good job of addressing reviewer comments.

Response: Thank you for taking the time to go through the revised version of our manuscript, your comments contributed to significantly improve the quality of the study.

A legend is needed to relate point size to the number of individuals in Figure 1.

Response: Yes. As indicated above, this information is now included in the Figure 1 caption.

Is there any chance that differences in library preparation contributed to differences in IBD estimates among species? For example, if DArT-seq targets hypo-methylated sites that are associated with a restriction enzyme cut sites, presumably the DArT loci would have greater functional relevance than the RAD loci and might be associated with environmental differences in the study region.

Response: Thanks for this good point. As detailed above, this issue is now explicitly addressed in the manuscript.

Last paragraph seems like it was added as an afterthought. We suggest incorporating it into the previous paragraph to summarize the findings and to suggest next steps.

Response: Done.

The authors suggest that IBD at small spatial scales in *M. surmuletus* and *S. cabrilla* might be related to the fact that sampling occurred on the spawning grounds in months that align with the known spawning seasons of these fish (lines 344-346). However, Table S1 says that spawning migration is unknown for these species. It seems that the findings of IBD at small scales for these particular species could potentially be due to adult spawning migration (as seems to be the case for the summer flounder study), larval dispersal or a combination of the two. The authors do make clear that sampling of adults incorporates larval and adult movement (lines 350-351), but we think that the authors need to be very clear that adult and/or larval movement could explain their IBD findings in *M. surmuletus* and *S. cabrilla* and not just adult movement.

Response: Yes, it is unknown whether there are spawning migrations in *M. surmuletus* and *S. cabrilla*, but in any event we sampled within the spawning season for these two species. Our point is not that the signal of restricted dispersal in these two species is due to migration but rather that the lack of such a signal in the other two species may be due to spawning migrations since we sampled outside the spawning season.

“The fact that we sampled adults implies that our genetic data integrate the effect of dispersal at both the larval and adult stages. This is particularly relevant in our case since the habitats are more or less continuous, which implies that adults can potentially move throughout these habitats [32,35–37]” (lines 124-127, already present in our revised manuscript).

*“We suggest that this may be due to the known migratory behavior of *D. sargus*, which spends the majority of the year in genetically heterogeneous populations [37] before returning to specific and distinct areas for reproduction [32]. Sampling was mostly conducted between June and November, almost exclusively outside of the *D. sargus* spawning period (March-June, [35,66]), when migration may have contributed to obliterate the IBD signal, and the same is true for *P. elephas* [35,36]. The overlap between sampling and spawning season was greater for *M. surmuletus* and *S. cabrilla* (June-July), the two species for which we did detect a significant IBD signal.”* (lines 354-361, already present in our revised manuscript)

“The combination of restricted larval dispersal, captured with the biophysical model, and limited adult movement during spawning season is likely to explain our findings.” (lines 361-363, new)

“This illustrates the fact that our approach, which is based on the sampling of adults, integrates both larval and adult movement. It is therefore preferable to sample during the spawning season when it comes to detect IBD” (lines 366-369, already present in our revised manuscript)

The Dryad link provided in the manuscript and the reviewer responses didn't work. We also couldn't access the finescale_ibd repository on GitHub. Please double check the links are correct and the

repositories are public.

Response: As mentioned above we sincerely apologize for this oversight. The Dryad and GitHub repositories are now publicly available at <https://doi.org/10.5061/dryad.0cfxpnw1s> and https://github.com/laurabenestan/finescale_ibd, respectively.

Line 56. Replace 'high' with 'large'

Response: Done.

Line 68. Replace 'have therefore to' to 'have to therefore' or 'therefore have to'

Response: Done.

Line 72. Allows one? Allows us?

Response: Done.

Line 80. Allow us?

Response: Done.

Line 252. Put 'i.e. from marine species sampled continuously at large spatial scale' in parentheses and delete prior ';'. Or replace prior ';' with ';

Response: Done.

Line 330 and 331. Change 'scale' to 'scales'

Response: Done.