Reviewer(s)' Comments to Author:  
  
Referee: 1  
  
Comments to the Author(s)  
I read with pleasure this study since it provides clear results about the genetic diversification occurred after a relatively recent colonization of Pacific islands by the Monarch butterfly. In my opinion the results are suitable for Proc R Soc B and I have no technical comments, apart from the minor ones suggested for Figure 1. My main concern is about introduction and discussion which are too focused on the study system without embracing the topic of genetic diversification of island butterflies and of the genetic consequences of occasional dispersal vs migration. This is particularly evident when the authors discuss as unexpected a genetic differentiation among islands located 40 km far to each other, which, in many island systems it appears as a rule (see below). In practice, I mainly suggest the authors to wide their introduction and disussion to fit the aims of a multidisciplinary journal as Proc R Soc B is.

**We have attempted to reframe the introduction and discussion of the manuscript to focus more on the concept of partial migration and its connection to population genetic structure. Throughout the manuscript, we have added references to other study systems in a way that we hope makes our findings more broadly applicable.**  
  
Lines 32-34. Although Monarch butterflies are perhaps an extreme case, in most temperate areas butterflies have largely expanded their distribution after the LGM often showing clear genetic fingerprints for these events (see here <https://doi.org/10.1111/1755-0998.13059> for a recent comparative study in European butterflies).

**Thank you for directing us to this paper. We have added this reference to our introductory paragraph as another example of a clade showing postglacial range expansion.**

Methods. It seems that most of the genetic indexes have been obtained using the snpR package which (according to the text) is not available yet. I tried to access the github link reported in the data accessibility section but I could not find the page. I imagine that the authors will include these functions in the released scripts, so they can also mention in the text where they are now available (e.g. supplementary data, repository, etc).

**The snpR package is now publicly available (https://github.com/hemstrow/snpR) and has an accompanying preprint announcing its publication, which is now cited in the manuscript.**

Line 234-235. Although striking this is not surprising in butterflies. Butterfly populations are known to genetically diverge even between continental islands and their nearby mainland in the Mediterranean (see for example many papers of the Roger Vila’s group) where it is clear that populations from Sicily and neighbouring islands (<https://doi.org/10.1038/srep28828>) and between Sardinia and Corsica and neighbouring mainland (<https://doi.org/10.1111/ddi.12610>) host distinct matrilines. A less recent paper on Hipparchia of the same areas also found the same results using allozymes (<https://doi.org/10.1111/j.1095-8312.1994.tb00982.x>). What the author can say is that although many butterflies show genetic differentiation across relatively short sea channels, this case is surprising given the strong dispersal capabilities of this species. Indeed, comparative studies (cited above) revealed that most (but not all) highly dispersive butterflies tend to show admixed populations over large continental and island areas. This will also help to wide the introduction and the discussion of the study.

**We have updated our discussion so that it now references two of the studies above and highlights that our results have precedent from other systems contrasting population structure in taxa inhabiting mainlands versus islands. We also cite another study demonstrating a similar pattern in a widespread dragonfly. Finally, we make sure to point out that our results are unique because they involve a contemporary (<200 years old) range expansion that has generated a migratory/non-migratory divide which is detectable in population genomic data.**  
  
Lines 294-297. It is a real pity that this assertion, which is fundamental to the economy of the paper in my opinion, cannot be supported by a reference. Are there references supporting similar results

Lines 260-261. Again this is also the case of another migratory butterfly species not mentioned here, Vanessa cardui, also rely on winds for migrations and show almost no genetic differentiation (<https://doi.org/10.1111/j.1600-0587.2012.07738.x>, <https://doi.org/10.1093/biolinnean/blx074>).

**Thank you for bringing these references to our attention. We have updated our discussion so that it includes the paper by Stefanescu et al. (2007) as an example of another butterfly species whose movement patterns are strongly influenced by prevailing winds.**  
  
Figure 1. Letters (a,b,c,d) are missing in the figures. I would include the number of sampled specimens next to the labels indicating the name of each island in the map of Fig. 1(a). Indeed, a high number of sampled specimens could allow to discover rare lineage while (at the other extreme) analyzing a single specimen can not produce any differentiation.

**We have updated Figure 1 so that each panel is labelled separately and so that populations are now labeled with sample numbers in the legend. We chose not to display the full name of the populations on the map due to label overlap issues. We instead noted the full names in the figure legend.**  
  
  
Referee: 2  
  
Comments to the Author(s)  
In this manuscript, the authors used RAD sequencing of monarchs from North America and locations across the Pacific to study the population genetics of monarch butterflies following range expansion from North America. They used a number of population genetic tools, including FST, Tajima’s D and heterozygosity estimates, as well as demographic models implemented in dadi. The authors conclude that monarchs dispersed to Hawaii, and from there dispersed independently to the Mariana Islands and other locations in the Pacific. The FST, NGSadmix, neighbor joining tree, and heterozygosity analyses support the conclusion of serial dispersal from North America. The use of thousands of markers is a significant advancement over previous studies with microsatellites, allozymes and other markers.   
  
While the results are interesting (it is great to see how monarchs dispersed across the Pacific from a North American origin), I have a few major concerns with this manuscript, as well as a number of minor concerns, as specified below.  
  
Major concerns:   
  
1.      First, the authors do not provide a compelling narrative for the study. What is the big question addressed here? Dispersal patterns for monarch butterflies are certainly necessary, but what insights does the study lend to population genetics in general, and how does it provide relevance to other systems. This is not made clear.

**We have updated the introduction and discussion of the manuscript substantially to provide broader context for our results. While most of the paper still involves analysis and interpretation of results that are specific to the monarch butterfly system, we have broadened the overall focus of the paper so that the manuscript now emphasizes (1) patterns of population genetic variation across space in partially migratory species; (2) the importance of seasonal migration and breeding season connectivity as homogenizing forces for spatial genetic variation; (3) general complications associated with demographic inference involve recent divergence events and small founding populations.**

2.      Second, the study is not quite as novel as the authors make it sound. Indeed, this manuscript largely deals with a question that has been addressed in a published paper in Proceedings B already: Pierce et al 2014 showed how monarchs dispersed from a North American origin to found new populations around the globe. Indeed, they found that monarchs underwent a stepwise dispersal from North America to Hawaii to Pacific islands southwest of Hawaii, and the title of their paper was “Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies”. Thus, the finding of stepwise dispersal of monarchs during range expansion in the Pacific is not novel. It is surprising in this context that Pierce et al. 2014 is not cited at all in the introductory framework of the manuscript, which instead refers to evidence in other systems, and to other papers on monarchs, including references 9, 11, 12 and 13. It is not until the fourth paragraph of the introduction that Pierce et al. 2014 is referred to, and when done so, it is said that the evidence is “quite strong” and that investigating additional populations will improve understanding of the cross-Pacific range expansion. Of course that is true, but the authors need to be much more upfront about what is already known, and what they add with the current manuscript. In short, they confirm a serial dispersal that has been previously found, as well as a second serial dispersal from Hawaii to the Mariana Islands.

**We agree that Pierce et al. (2014) already demonstrated that trans-Pacific dispersal in monarchs involved a serial, stepwise expansion process, and we have re-written the introduction so as not to emphasize the novelty of this result. Instead of focusing on serial dispersal itself, the introduction and the goals of the manuscript are now framed around (1) characterizing overall patterns of relatedness between North American and Pacific monarchs, including previously unsampled populations (n = 4 populations) from the Mariana Islands and Norfolk Island.**   
  
3.      Third, while much of the paper deals with serial dispersal across the Pacific, the demographic analysis is focused entirely on North America and Hawaii. Why is this? This needs to be explained. In addition, there are a number of improvements to be made regarding the demographic analysis. Currently, the authors do not explain the different models they are contrasting. Instead, they refer to existing papers without telling the readers what models those existing papers actually used (lines 148-153). There are no visual representations of the models the authors contrasted (these should be provided for all tested models). And the models they discuss the most, because of greatest support, do not appear to be the best models to reflect the actual occurrences of population expansion: both the “found and grow” and the “three epoch” models consider the North American and Hawaiian populations as two populations that diverged from one population. But the Hawaiian population is a relatively small population split from a large North American population. I therefore think the authors should consider using more 2 population models, such as described in the depository: <https://github.com/dportik/dadi_pipeline/blob/master/Two_Population_Pipeline/Models_2D.pdf>. I also suggest that the authors provide a table of log-likelihoods of all the simulated models, and use Tajima's D results to interpret findings regarding demographic history. NAM showed a -ve Tajima's D showing a recent expansion after a bottleneck. HAW population showed a less -ve Tajima's D showing a slight recent expansion after a bottleneck. All other populations show a recent bottleneck. Further comments to demographic analyses are provided below.

**We have rephrased the demographic analysis portion of the methods to stress that we fit all of the two-population models in the Island set from dadi\_pipeline, and clarified that when the proportion of individuals that colonized Hawaii (the parameter *s*) is very low, only a very small number of North American individuals found Hawaii, which is a realistic scenario. A table of log-likelihood/AIC results and the number of completed runs for each model set across each successive run is now provided in Table S4. Additionally, we added a figure depicting the residuals from the derived allele frequency spectra from the optimized parameters for each model vs. the observed data. We included more mention of Tajima’s D in the discussion.**  
  
4.      It is hard to follow the results because the authors have placed most of their figures in the supplementary information. This is really not necessary. Indeed, most figures should be included in the main body of the manuscript. Also, please provide a legend in Figure 1 so it is easier to see what populations the three-letter codes refer to (rather than having that in the caption), and increase font size of the tree so it is legible.  
Minor comments, and further elaborations regarding major comments:

**We have re-arranged and relabeled Figure 1 for readability. Full names and sample sizes for the different populations are now noted in the figure legend. On the tree, sample names are no longer noted—tips are now colored according to the legend instead. The key findings are all now in the main text: structural results (Figure 1), demographic model results between the four best models (Figure 2), and observed/expected SFS plots. While we have kept other plots in the Supplementary material for now for space reasons, we are happy to move them to the main text at the editors’ recommendation.**  
  
5.      Lines 5, 70. It is said that “approximately 280 monarch butterflies” were sampled. I don’t understand “approximately”. Please provide actual number.

**This was updated to read “281 monarch butterflies”**  
  
6.      Line 12: the sentence that says that estimates are concordant with recent expansion but with high uncertainty contradicts itself: if colonization estimated vary from 100 to 100,000 years, then how is that concordant with a recent expansion? The concluding sentence in the abstract (14-17) is hard to follow without some numbers, so please provide more detail in the abstract for a logical progression and conclusion.

**We updated the abstract so that it now simply says that these estimates “do overlap with historical records that indicate a recent expansion.” The range of estimates for establishment timing are obviously very wide, and it is partly due to this high uncertainty that we feel justified in deferring to historical records for estimating establishment timing. Additionally, our models consistently underestimated the number of rare alleles derived found in NAM but not HAW. Since rare alleles should be lost during bottlenecks, this suggests that our models failed to optimize to a small enough founding population size in HAW, which is unsurprising since very small population sizes cause ballooning run times and integration errors in dadi. This futher suggests that the recent introduction times are probably more likely to be true than suggested by the optimum model results alone. We’ve added explanation about this to the discussion section**

7.      Cluster analysis: please provide error rates for each value of K (If available in NGSadmix). The populations ENA and WNA showed unusual patterns with K=9, please address this in the manuscript.

**We added a plot of Evanno et al. (2005)’s deltaK statistic was highest at k = 2 and k = 5, although we were not able to calculate it for K = 3 due to very low likelihood variance between runs.**  
  
8.      Lines 190-191. It is said that two different islands in Hawaii did not show differentiation, but the map only shows one pie chart, not two (for Maui and Oahu).

**Figure1 has been updated so that Maui and Oahu are shown separately. Because there was little population structure within the Hawaiian islands, they look very similar to each other.**  
  
9.      Regarding the dadi models: the conversion of optimized parameters to effective population size is dependent on the used mutation rate. The authors have used the mutation rate of Drosophila melanogaster to convert the optimized parameters to absolute years and migration rates, but Drosophila melanogaster has a different chromosome structure and base composition than lepidopterans. Therefore, it would be more appropriate to use the mutation rate of Heliconius melpomene ( 2.9\*10-9) which is in the same family as the monarch butterfly, and much less divergent than Drosophila melanogaster. It would also be helpful if the authors provided the optimized parameters (raw values without conversion) from the best model.

**We used the *D. melanogaster* parameters so that our results were more directly comparable with Zhan et al (2004)’s demographic results; however, we used a more realistic generation time and mutation rate (from *H. melpomene*) for comparison. The results varied very little on a qualitative scale, since the two changes roughly cancel out. A paragraph about this has been moved from the supplementary material to the main text for clarity.**  
  
10.     Following the recommendation above to provide raw value estimates from the dadi model, it would also be helpful if the authors provided the model and residual plots of the best optimized model. This can be created using Plot\_3D function in dadi\_pipeline.

**We have added Figure 3, which contains the observed sfs and the optimized outputs from the four best models. Since the estimated site frequency spectra vary pretty drastically across the optimized parameter space, and AIC didn’t tend to vary much across this space, we’ve included estimate sfs using the optimized parameters from the four quadrants of the optimized introduction pop size/establishment time parameter space for ease of comparison. We’ve also included a plot in the supplementary material (S5) of the residuals between the observed data and each of these spectra.**

11.     Lines 202, 207 refer to Figure 4. I am assuming Figure 2 is meant to be referred to instead? Likewise, line 212 refers to Figures 2e, 2f. I am assuming Figure S4 should instead be referred to?

**All figure references should now be correct.**  
  
12.     Lines 221-222. The authors state that historical estimates of introduction timing overlap with demographic reconstructions. While I agree, it is clear that the historical estimates (~200 years) are near the very low end of the demographic estimates (100-100,000 years). This needs further discussing.

**We have updated the discussion of the paper to better explain our rationale for accepting a recent establishment date for Pacific monarchs. The enormous range of possible establishment times provided by dadi suggests that the value of model-based inference may be somewhat limited in instances such as these, where we are potentially dealing with very recent events and small founding population sizes. We added two paragraphs explaining (1) why the establishment timing estimates may be biased towards more distant events and (2) additional, non-model-based reasons for why we feel it is more appropriate to defer to historical records. We also noted that the majority of the residual difference between the modeled spectra and the observed spectra are in rare derived alleles present in North America but not Hawaii. Since rare alleles are the most likely to be lost during severe bottlenecks, this implies that the optimizer may not have been able to optimize for a small enough founding population size (which is not surprising given that dadi optimization slows down dramatically or can have integration errors when population sizes are very small). Since small sizes are correlated with recent introductions, this gives some credence to a more recent introduction. We added a paragraph on this to the discussion.**

13.     Lines 228-229: the authors hypothesize that monarchs in Australia likely came from New Caledonia, based on historical records. However, what do the genetic data in this manuscript say? The tree in Figure 1 should be able to help answer this question (I could not readily read the figure because the font was so small, but it seems that New Caledonia could be ancestral to at least some locations in Australia?).

**Our genetic data are consistent with Australia being founded by one of the southwestern Pacific Island groups, which includes New Caledonia, Samoa, New Zealand, and Norfolk Island. Although we are hesitant to read too much into any single analysis, our neighbor joining tree is most consistent with Norfolk Island as the founding population for Australia. However, this result could also be driven partly by periodic gene flow from Australia into Norfolk Island (residents of Norfolk Island report occasional influxes of monarchs and other butterflies from the Australian mainland when prevailing winds are favorable, similar to the scenario implicated in the monarch reaching the United Kingdom).**  
  
14.     Lines 247-248: it is stated that monarchs provide a unique opportunity to compare patterns of population structure of migratory and non-migratory populations. However, there are other species in which this could be done, so this statement should be toned down.

**We have updated the manuscript and tried to limit references to our results being novel or unique. The lone instance where we use this language is in describing this system as a rare example where it is possible to compare population genetic structure in a partially migratory species whose non-migratory descendent lineages became established over contemporary time scales.**  
  
15.     Lines 251-255: this paragraph compares the results from the current analysis with those from previous papers, but it does so without explaining what those other papers found and what molecular markers they used. Please expand.

**We updated the manuscript so that it more thoroughly describes the findings of Hughes and Zalucki (1984) and Zalucki and Hughes (1987).**  
  
16.     Line 264: the authors state that Australian monarchs retain migration-associated behaviors, referring to both a study on monarch movement and a study on diapause and circadian gene expression. It is important to note that diapause induction and circadian gene expression do not necessarily mean retention of migratory behavior: while these biological processes are necessary requirements, they are not sufficient conditions for migratory behavior.

**We agree that diapause induction, circadian clock gene expression, and the ability to cover large distances are necessary but not sufficient conditions for migratory behavior, and we have updated this sentence of the discussion accordingly.**  
  
17.     Lines 287-297. I agree with the authors that the models that include estimates of colonization within the last 200 years are much more likely to be true. That said, I do not think the third point of argument, based on genetic analysis of commercial monarchs is helpful for the argument. The paper referred to here included monarchs from one breeder, creating a huge sampling effect, much greater than what would likely have happened during natural colonization.

**While we agree that this paper (Tenger-Trolander et al. 2019) perhaps shouldn’t be used in the context that we present it (the paper never makes any assertions about degree of divergence), we do feel that the bottleneck effects associated with the establishment of island monarch populations may be comparable to those associated with the founding of these commercial monarch lineages. If necessary, we can also update this section to reference a personal communication with that study’s author about an upcoming analysis that more strongly corroborates our argument here (commercial monarchs are inferred to have been “founded” >1000 years ago using PSMC and other demographic inference).**  
  
18.     Table 1 and Table S2: please provide standard errors.

**Since these metrics are calculated per population, not per sample, they are not typically presented alongside standard errors. While they can be given with standard errors *across loci*, these will be very small when there are many loci, as here, and so will probably misrepresent the true variance of these parameters. We have, however, added bootstrapped *p-*values to our pairwise Fst calculations!**  
  
19.     Table S2: the FST between the populations VIC and NAM seems to be negative, which seems odd, as FST is theoretically between 0 to 1. Please clarify.

**The FST between VIC and NSW is indeed negative. While Fst as traditionally defined (Fst = (Ht – Hs)/Ht) is between 0 and 1. However, Wier and Cockerham’s estimator of Fst (which is ostensibly what is used by genepop) can occasionally return negative Fst values due to weighting quirks. This usually occurs when sample sizes are small and within population variance is larger than between population variance, and is typically understood to essentially equate to Fst = 0. For one other example where Fst values were negative between populations when calculated with WC’s method, see Bortolloto et al 2011 (doi: 10.1186/1471-2156-12-32).**