Reviewer Comments:  
Reviewer 1

The authors study the re-emergence of an important life history trait, migration, by studying the genetics of monarch butterflies. In particular, monarchs are ancestrally from the Americas, where a large population in North America undergoes seasonal long-distance migration. Derived populations of monarchs can be found across the Pacific islands, in almost all cases having lost migratory behavior. In Australia, however, monarchs have redeveloped migratory behavior. The authors surveyed variation in reproductive diapause, specifically delayed oogenesis in females, as it related to genetic variation across families of Australian monarchs. They found a few candidate loci, in particular variants near the gene karst which is known to be expressed in the ovaries of other insects, and thus makes a likely candidate for the observed phenotypic difference.

Overall, I found the manuscript’s motivation and methods clear and think the implications on migration and monarch ecological history are well discussed. As evidence of the quality of this manuscript, I have only a few comments, all of which should be easy to address. I’ve divided them into comments that relate to an area of the text more broadly (minor comments) and those specific to certain lines (line-by-line).

Minor comments:  
Intro: in addition to stated concerns about monarch populations, migratory monarchs apparently hold a high genetic load of deleterious mutations (tolerated because of high heterozygosity and outcrossing in migratory populations). Inbreeding, as becomes more likely in smaller, structured resident populations is likely to result in inbreeding depression, which decreases egg viability and adult longevity, at least in lab settings:  
Mongue, A. J., Tsai, M. V., Wayne, M. L., & de Roode, J. C. (2016). Inbreeding depression in monarch butterflies. Journal of insect conservation, 20, 477-483.

We added this reference and a brief description of the buffering effect of migration on genetic load in monarchs to the introduction (Lines 64-66).

References: Check formatting of citations. I noticed, for instance, that Linnaean names are not italicized, e.g. (not an exhaustive list) refs 16, 18, 23.

We double checked our references section to ensure that species names are italicized. We are happy to fix any other instances or reference format issues found during final copy editing!

Figure 1: You may wish to adjust the color (or at least opacity) of the monarch population categories, specifically the ancestral migratory population. The key shows blue on my monitor but because the highlighted area is over the brown landmass of North America, it comes out as a dark brown in the map.

We darkened the opacity of the ancestral monarch publication in Figure 1—hopefully this will be more consistent across devices.  
  
Line-by-line:  
L93: “less pronounced in males monarchs”, should be male monarchs

Fixed!

L102: “a homolog of which is known to be expressed in silkworm ovaries” since its between species homology, it might be more precisely called an ortholog, no?

While this specific sentence was re-worked and should no longer be an issue, there are a few instances where we left homolog (over ortholog) given that we are unsure if the ancestral divergence point between the *Karst* homolog in monarchs and other species are speciations or duplications (given that ancestral, within species duplications which persisted following speciation would result in paralogs across species if we understand the terminology correctly—I’ve taught this in evolution dozens of times, but it’s still seriously confusing sometimes!). We are very happy to change this to orthologs if the editor prefers, however!

L116: “arisen multiple times after the species’ split from Danaus erippus” just a grammar style choice, but I find the apostrophe to be a bit clunky there. Incidentally the sentence works just fine without it: “after the species [noun] split [past tense verb] from Danaus erippus” is syntactically valid.

Changed as suggested!

L122: remove the comma be for “and were recorded”

Done.

L131: “females may mate multiple times” could use a citation. Smith’s book on sperm competition has some data showing that monarchs are highly polyandrous:  
Smith, R. L. (Ed.) (1984). Sperm competition and the evolution of animal mating systems. Burlington, MA: Academic Press

We added this reference, thanks for the suggestion!

155: “envelops” should be “envelopes”

Fixed

L258: again “ortholog” would be more appropriate here than “homolog” in my opinion

See above.

L321: “life-spans” does not need a hyphen.

Fixed

L356: “than that underlying phenotypically fixed traits” should be “than those…traits”

Re-worded to hopefully be clearer (Lines 405-406).  
  
  
  
Reviewer 2  
The re-emergence of migratory behaviour in the ‘non-native’ Australian population of monarch butterflies provides a fascinating contrast to the loss of such behaviour in parts of the North American native range, and an opportunity to uncover the genetic basis of this phenotype.  
  
This study makes a first step to answering this question by performing a GWAS on a sample of female butterflies, bred under controlled conditions, that vary in their degree of reproductive maturity, which is correlated to migratory phenotype.  
  
The main finding is that one of the three significant outlier loci coincides with Karst, a protein involved in reproductive development. The tentative argument for its involvement is that migratory (and diapausing) insect phenotypes divert energy resources from egg production to fat stores. The fact that Karst has not been implicated with migratory behaviour in an earlier study of North American populations is suggested to be due to different axes of genetic variation among the two populations.  
  
I am not an expert on GWAS but I did not find any obvious errors in the methods. I assume that the use of principal components as covariates deals with the issue of non-independence among closely related individuals.  
  
The authors identify the limitations imposed by the use of RAD rather than whole genome sequence on detailed analysis of the linked haplotypes, and the small number of independent maternal families (n=32).  
  
My feeling is that the finding is interesting but, being based on rather limited data and comparisons, the conclusion that variation in Karst is important for migratory behaviour is a little tenuous.

We have added more context around *Karst*’s possible association with migration, in particular it’s possible involvement in the *Hippo* pathway which influences wing shape, a trait which is strongly correlated with diapause in our system and is a classic part of the migratory syndrome (Lines 319-322). We also calculated and added the percentage of variance in reproductive development explained by variation at *Karst* (lines 252-260, 280-286, and 304)*.*   
  
It is not clear why a study of the genetics of migratory behaviour was conducted entirely on a non-migratory population. Minimally, would it not be possible to genotype the outlier Karst SNPs in a migratory population sample to test your interpretation?

This was an excellent suggestion! We added an extensive analysis of pre-existing data which we had previously genotyped for a study of monarch genetic variation in North America and across the Pacific using the same lab and sequencing protocols. Specifically, we genotyped *Karst* loci in these samples at the most strongly at associated loci in order to determine the presence of *Karst* diapause-associated alleles in the ancestral monarch populations. We significantly re-worked the discussion around our finding that migratory *Karst* alleles are present in North America at low frequencies and on some Pacific islands, thus supporting the maintenance of migratory-associated variation despite repeated bottlenecks during the spread of Monarchs across the Pacific.

For clarification, we focused on this population initially because diapause when exposed to migratory cues is a known segregating trait in this population. Fully migratory populations in Australia likely all undergo diapause, as in other fully migratory populations, and thus association testing is not possible there. Hopefully adding comparisons to other migratory populations strengthens this aspect of the paper, but we also added a note that future work sequencing *Karst* loci in migratory Australian and New Zealand populations would be helpful (lines 374-375).  
  
Very many loci are involved in reproductive development. Therefore, getting a hit in one of them is not so unusual. Moreover, the link between reproductive status and migratory behaviour requires clearer support in this ms.

We added an additional logical link between *Karst* and migration (Lines 319-322, as mentioned above). We also demonstrate that a *Karst* variation explains a substantial and highly significant portion of phenotypic variance for diapause (Lines 283-286). As mentioned in lines 94-103, diapause is a very commonly studied aspect of migration, but we did add the cavate that future work can and should examine genotype-phenotype associations between other traits in the migratory syndrome in North America (lines 370-374).  
  
I appreciate that whole genome data is not available for the Australian populations, but presume that such resequencing data is available for North American samples. If so, it would be appropriate to describe the haplotype variation in Karst in North America and relate this to the SNPs found in this study.

Given that haplotyping our data is not ideal given our data type, we instead compared *Karst* allele frequencies in our studies to those in North America and elsewhere (see above).  
  
In terms of presentation, the ms is generally well written, with only minor suggestions:  
Line 193. Reproductive Score is the y-axis of Fig. 2b, so it would help to expand a little on its interpretation (e.g., with a RS of 1 corresponding to all females reproductively mature, and RS of 0.25 corresponding to…)

We added some text clarifying how the y-axis for figure 2b was calculated (lines 203-204).

Line 143. Table 1 indicates 164.

Typo fixed!

Line 163. 17+11=28 maternal families. Fig. 2b is consistent with 32. It would be useful to add a maternal families column to Table 1.

We added this column and fixed the inconsistencies, which we caused by pulling from the wrong data source instead of the colony results.

Line 226. We obtained a total of…sequence reads.

Re-worded for clarity.

Line 228. Retained…. called genotypes and genotype likelihoods.

Re-worded for clarity.

Line 274. Replace ‘delineate’ with ‘distinguish’.

Re-worded for clarity.

Line 343. Can you be more specific and provide supporting evidence for the Ne statement.

We added a reference and clarification (the relationship between Ne and total population mutation rate is due to mutation/drift equilibrium; lines 343-395).

Fig. 2 caption. A: Three stages of egg development used to phenotype adult females as reproductively mature

Re-worded for clarity.