**Title:** Novel genetic association with migratory diapause in Australian monarch butterflies

**Authors:** William Hemstrom1,2, Micah Freedman3, Myron P. Zalucki4, Michael Miller1

**Author Affiliations:**

1. Department of Animal Science, University of California, Davis; Davis, California, USA
2. Department of Biology, Colorado State University; Fort Collins, Colorado, USA
3. Department of Ecology & Evolutionary Biology, University of Toronto; Toronto, ON, Canada
4. School of the Environment, The University of Queensland; Brisbane, QLD, Australia

**Corresponding Author:**

William Hemstrom: hemstrow@gmail.com, 503-730-5529

# Abstract

*Background*

Monarch butterflies (*Danaus plexippus*) are a charismatic and culturally important North American butterfly species famous for their dramatic and unique migratory life history. While non-migratory populations of the species are widespread and apparently stable, migratory populations in North America have recently seen declines, prompting concern that the migratory phenomenon in North America may be at risk of disappearing. In contrast, a relatively recently-established monarch population in Australia has rapidly re-acquired a migratory life history following hundreds of generations of residency and successive bottlenecks as the species island-hopped across the Pacific during the late 1800s and early 1900s. The process by which migration re-emerged in Australian monarchs is not currently known.

*Results*

We raised and sequenced individuals from Queensland, Australia under environmental conditions associated with migration initiation and found strong variance in reproductive diapause, a key migratory trait, between families which was associated with the spectrin beta chain protein *Karst*. This protein is known to be involved in reproductive development in monarchs but has not previously been identified to be associated with migratory life history variance. The most strongly associated migratory SNPs are also present at a low frequency in North America, suggesting that the Australian population is leveraging standing variation which persisted across repeated bottlenecks as Monarchs spread across the Pacific.

*Conclusions*

Our results provide an intriguing example of how the temporary loss of migration—in this case likely over hundreds of generations—may not entail the loss of genetic variation associated with this complex life history strategy.

*Keywords*: monarch butterflies, migration, genomics, reproductive diapause, GWAS

# Background

Migratory species are acutely at risk from global anthropogenic environmental change (1,2). The loss of breeding, wintering, or any transitional habitats may all cause declines in migrants (3), as may shifts in resource phenology due to seasonal changes which can cause resource availability and species life-histories to fall out of synchrony (4). It is therefore not surprising that widespread declines have been observed in many migratory taxa (3,5,6).

Population-level adaptability in migratory life-history should therefore be expected to increase long-term population viability by conferring resistance to phenological and environmental changes (7). Migratory life-history adaptation can occur along three distinct axes: 1) *temporal*, or adaptation that changes the phenology of migration; 2) *spatial*, or adaptation that alters migratory pathways and orientation; and 3) *residency*, or adaptation that shifts populations between migratory and resident life histories. Inter-population variation and adaptation in each of these pathways is well known from nature: temporal changes or variation in migration are well documented in many taxa including as birds (8–13), fish (14), mammals (15), and butterflies (16,17); as is short-stopping, extension of migratory pathways, or complete re-direction of migration (18–25); and shifts towards residency or partial migration have been observed and/or predicted in birds (7,26–30), butterflies (31), mammals (32–34), and fish (35).

Adaptation via a transition to residency is of both particular interest and concern for the conservation of migratory species. While switching to a resident life history can allow species to avoid the increasing risks of migration and the intrinsically large energetic costs of migration in general (36–38), there is reasonable concern that the “extinction” of migration may not be easily reversible in some species. Permanent loss of migratory behavior is concerning because, in general, variation in migratory tendency is expected to increase long-term population persistence (39). Beyond this, migratory species often provide critical ecosystem services (40,41) and have cultural and/or ecological significance (42–45) that may not survive the transition to permanent residency.

The loss of migratory behavior is a major concern for monarch butterflies (*Danaus plexippus*) (46,47). Specifically, North American migratory monarch populations have experienced pronounced declines in counts of overwintering individuals over the last several decades, sparking a fear that those populations may go extinct in the near future (48), although not all agree (49). However, many individual butterflies ostensibly from the North American migratory population are now year-round residents in some areas in the southern United States (31), where they can survive due to the year-round presence of introduced tropical milkweeds (e.g. *Asclepias curassavica*, *Gomphocarpus sp., Calotropis sp.*) (50), and there are widespread, non-migratory monarch populations that are not of any immediate conservation concern (51). While the species as a whole may therefore not be at risk, their migratory behavior seemingly is. That said, in addition to the damage to the long-term population persistence that could result from the loss of migration in North American populations due to the loss of life-history variation (as discussed above), North American monarchs carry large deleterious genetic loads which are likely buffered by increased effective population sizes and reduced inbreeding facilitated by migration-induced gene flow (52). Moreover, beyond the ecological and demographic risks the loss of the publicly beloved annual monarch migration and concomitant overwintering would be culturally costly (53).

In contrast, there is a population of monarchs in Australia in which migratory behavior has apparently re-emerged after it was previously lost. This population, which was established sometime before 1871 (54), is the product of a series of sequential introductions that crossed the Pacific Ocean from North America beginning sometime in the early 1840s after the introduction of their host plants enabled individuals to survive after being blown from the mainland (Figure 1) (54–56). Monarchs are not known to be migratory on any Pacific islands, and no migratory behavior was observed in Australia until the 1930s, when non-reproductive overwintering aggregations were observed in northern New South Wales (NSW) (57). These and other individuals were later shown to migrate hundreds of kilometers each year to their wintering grounds from more southerly, inland locations in NSW (57,58). Thus, more than 90 years (and hundreds of generations) after expanding out of North America, and after more than 90 years without any evidence for seasonal migration, migration re-emerged in these populations. Understanding how this happened could provide crucial context around the “extinction” of migration in monarchs and other species.

To better understand how migration re-emerged in Australian monarchs, we studied continuously breeding monarch populations in Queensland. While these populations do not experience substantial seasonal fluctuations in resource availability, and thus do not migrate (57), recent work has shown that reproductive diapause can be induced in some, *but not all* individuals from these populations if they are exposed to a reduction in photoperiod during development (59), a trait that these non-migratory populations do not display in the wild. In brief, reproductive diapause, where individuals delay their reproductive maturation, is a critical and fitness-relevant phenotype for migratory monarchs because it allows them to invest more heavily in lipid storage and increase their average longevity (60–62). Like other migratory behaviors, diapause is initiated by changes in day length, temperature, and other seasonal fluctuations (63–65) and is thought to be directly controlled by juvenile hormone (JH) titers during development (60,61,66).

While diapause is one of many migration-associated traits (such as wing morphology, resting and flight metabolic rates, lifespan, and lipid storage), it is in particular linked extremely tightly to migration. Nearly all wild migratory monarch females are in diapause: previous work has found that between 82 and 100% of female migratory monarch butterflies are non-reproductive (47,63,67,68). In contrast, previous rearing trials under typical North American summer breeding conditions have yielded nearly 100% reproductively mature females (63). These trends are less pronounced in male monarchs, which are therefore not the focus of this study (47,67). Given that phenotyping the entire migratory syndrome is challenging, diapause serves as an excellent proxy for individual migratory status and has been used as such or as a focal trait for the study of migration in monarchs in many studies (47,59,63,66,67,69).

The variance in diapause induction in Queensland monarchs allows for a direct search for genetic associations with this trait and thus for an examination of the genetic underpinnings of the re-evolution of migration in Australian monarchs. Here, we used reduced representation sequencing to search for genes associated with migratory diapause in Queensland monarch butterflies. We found three strongly associated loci, one of which co-located with a spectrin beta chain protein *Karst*, which is involved in actin filament binding, is known to be expressed in silkworm ovaries, and is expressed during monarch diapause termination. This gene had not previously been identified as associated with migration in North American monarch butterflies, and we did not find that any previously identified migratory genes were associated with diapause in our study; however, using data we previously published from North American and Pacific monarchs we found evidence that the *Karst* SNPs associated with migration are nonetheless present in North America.

# Methods

## Background

Although the ancestral range of monarch butterflies is believed to be in North and Central America, they have fairly recently expanded into South America and the Caribbean, and very recently to several locations throughout the Atlantic and Pacific (70–72). In North America, the species is well known for its unique, multi-generational migratory life history wherein individuals dispersing northward in the spring do so over three to four generations and then return to their wintering grounds in a single step (73,74).

In Southern Florida, the Caribbean, Central America, northern South America, and throughout most of their introduced range monarchs are non-migratory (75–77). Residency in these populations appears to be a derived trait which has arisen multiple times after the species split from *Danaus erippus*, their closest extant relative (71). The Pacific expansion of monarchs is relatively recent, with historical records indicating that they established first in Hawaii in the 1840s, likely as a result of individuals blown in during storm events that were able to survive on recently introduced milkweeds (55,78). Genetic evidence is consistent with an introduction to Hawaii during this time period (56). As described above, they then spread across the Pacific and reached Australia in approximately 1871(54) and were recorded migrating again in Australia sometime in the 1930s (57) (Figure 1).

## Sample collection, incubation, and phenotyping

We collected 22 female monarch butterflies from Pinjarra Hills, Queensland, Australia (27°32′26.7″S, 152°54′22.7″E) between the 5th and 9th of July 2018. Individuals in all life stages were present at the field site during our sample collections, consistent with continuous, year-round breeding. We enclosed each of these 22 females individually on either *Asclepias curassavica* or *Gomphocarpus sp.* host plants found on-location and subsequently successfully collected eggs produced by 21 of them. While these eggs represent 21 total maternal families, each may consist of multiple groups of half-siblings since females may mate multiple times (79) and lay clutches with mixed parentage. We did not remove *Ophryocystis elektroscirrha* (a protozoan parasite) spores, which were uniformly present on our eggs (consistent with previous reports from this location (59)).

We then incubated all eggs using the same incubators and following the “decreasing photoperiod” experimental scheme of Freedman et al. (59) in order to induce reproductive diapause and ensure that our data could be later pooled with that produced in the earlier study. Briefly, larvae were reared at a constant temperature of 28°C under a photoperiod regime that declined from 14:10 L:D to 12:12 L:D over the course of 30 days (Δ4 mins/day). We released all male butterflies following emergence, since phenotyping diapause in male monarchs can be challenging due to the difficulty in extracting and weighing the ejaculatory duct, seminal vesicle, and accessory glands, which are likely the best indicator of male reproductive development (63). This left us with a total of 164 females from 20 maternal families.

We assessed reproductive diapause by determining the degree of oocyte development in each of our adult female butterflies after 70 degree days of development, which is sufficient time for females to develop mature ovaries under normal summer conditions (80). We observed that individuals fell into several defined bins: 97 out of 164 had fully developed, chorionated eggs (visible ridges along the exterior of the egg), while 67 females did not. Of these latter individuals, a few (11) had no yolk in their eggs whatsoever, but most had some degree of yolking. Since vitellogenesis generally only begins to a substantial degree after eclosion in monarchs (81), we classified individuals as either fully reproductively mature (with chorionated eggs), partially reproductively mature (with more than a tiny amount of yolking), or reproductively immature (with no or only very small amounts of yolking). Images of egg development in each individual are available online (see Data Availability, Figure 2a). Wings and bodies from all individuals were preserved in dry coin envelopes or 95% ethanol, respectively. We note that while we refer to the postponement of reproductive development under otherwise suitable conditions as diapause, other authors refer to this process in monarchs as oligopause (82) or quiescence, since diapause implies an extended refractory period that we did not test for.

We supplemented these samples with preserved, dry butterfly wings from each of the “decreasing photoperiod” female individuals previously scored for reproductive development by Freedman et al. (59), constituting an additional 40 total female butterflies from 12 maternal families, which were scored by the authors as reproductively chorionated, yolked, and unyolked, respectively. Together, these two datasets contained 204 monarch butterflies from 32 maternal families (see Table 1).

## Sequencing and genotyping

We removed and extracted DNA from one leg from each of our samples collected in 2018 or from a wing base from each of the samples collected by Freedman et al. (59) using the magnetic bead protocol of Ali et al. (83). We quantified the resulting DNA on a BioTek Instruments FLx800 Fluorescence Reader using Thermo Fisher Scientific Quant-iT PicoGreen dsDNA Reagent, then prepared Restriction Associated Digest (RAD) libraries using the Pst1 restriction enzyme according to Ali et al. (83). We sequenced these 150 bp paired-end sequencing libraries using an Illumina Hi-Seq 4000.

We aligned the resulting raw sequence data to the “MEX\_DaPlex” monarch butterfly genome assembly (84) using the mem algorithm of the Burrows-Wheeler Aligner (85). We then filtered out PCR duplicates, improperly paired, or poorly mapped reads using SAMtools (86). For some downstream pedigree reconstruction, we then called genotypes using the ANGSD software (87) package with the following parameters: -doMajorMinor 1 (determine major and minor alleles using a genotype likelihood approach), -doMaf 2 (determine minor allele frequencies), -SNP\_pval 1e-8 (keep only loci with a SNP p-value ), -doGeno 4 (call genotypes), -doPost 2 (calculate genotype posterior probabilities using a uniform prior), -postCutoff 0.95 (keep only loci where the highest genotype posterior probability ), -minQ 20 (keep only loci with a sequencing quality ), -minMapQ 20 (keep only loci with a mapping quality ), -minInd 130 (keep only loci sequenced in at least 130 individuals), and -minMaf 0.05 (keep only loci with a minor allele frequency ).

## Statistical analysis

We first confirmed the parentage of our samples using the Colony2 program, assuming polygamous, random mating (88). In order to determine the degree to which diapause status was biased between families, we fit a basic linear model in R version 4.2.2 (89) using phenotypic status as the response variable and maternal family as a fixed effect. Phenotypic status was coded as 0, 1, or 2 for reproductively immature, partially mature, or mature (as described above), respectively. We used an ANOVA test in R to determine if maternal ID significantly improved model fit. To determine the expected phenotypic distribution for each family sorted by mean reproductive rank (calculated from the individual 0, 1, 2 scores rescaled to 0, 0.5, and 1 for immature, partially mature, or mature, respectively such that a family with 100% reproductively mature individuals would have a score of 1) under a null distribution, we randomly permuted individuals between families 10,000 times and then ranked each family for reproductive score.

We conducted a Genome-Wide Association Study (GWAS) to determine if any loci were significantly associated with reproductive status. Since our data was generally of low coverage, we used a frequency test in ANGSD (87) based on genotype likelihoods rather than called genotypes and thus increased the effective number of loci we were able to analyze. We used the following parameters: -doMajorMinor 1, -SNP\_pval 1e-12, -GL 1 (calculate genotype likelihoods using the SAMtools approach), -minQ 20, -minMapQ 20, -minInd 102 (50% of the individuals with called phenotypes), -minMaf 0.05, -doMaf 2 (assume a known minor allele), -doAsso 1 (do an association test using a frequency test), and -yquant coding phenotypic status as above. Since our samples were composed of many groups of full and half-siblings and came from two genetically different years, we also used the -cov argument and supplied the first 20 principal components derived from PCAngsd (90), which we ran using the default parameters using genotype likelihoods derived from ANGSD using the same parameters plus -doGlf 2 to save genotype likelihoods. We calculated p-values for the resulting likelihood ratio test score and corrected those values using the false discovery rate approach of Benjamini and Hochberg (91) in R.

We constructed quantile-quantile and Manhattan plots of the resulting p-values using snpR (92) and identified candidate adaptive genes as those co-locating within 50kb of any SNPs with a corrected . Since the “MEX DaPlex” reference genome uses RefSeq gene IDs but most monarch studies use the monarch official gene set IDs, we used BLAST (93) to identify official gene set genes with protein sequences matching those of the RefSeq genes we identified via the built-in tool on MonarchBase (94).

To determine if diapause is part of a generalized migratory syndrome in Australia, we re-analyzed the connection between reproductive status and wing morphology in our 2016 samples. To determine if wing shape or size was correlated with reproductive status, we fit a pair of linear mixed effect models with either wing shape or size as response variables, the number of yolked oocytes as a fixed effect, and maternal family as a random effect using the R package nlme (95).

## Genotyping Karst in North American and other Pacific monarchs

To determine if the migration-associated variation we observed in the Australian monarchs was present in the North American and Pacific island populations from which the Australian population is derived, we used data which we previously published on monarchs from these locations (56), which was originally produced using the same protocol and restriction enzyme as used here. After downloading this data from the NCBI, we aligned and filtered it as described above. We then removed reads not located within 100kb of *Karst* from both this dataset and the data described above using SAMtools (86) and called genotype likelihoods with ANGSD (87) for the region directly within *Karst* using the same options as above save for: -GL 2 (use the GATK method for calling likelihoods), -minInd 102 (keep only loci called in at least 102 individuals), and -doGlf 2 (produce a beagle formatted genotype likelihood file). We then imputed the *Karst* genotypes using beagle version 3.3.2 using the default options (96) in order to generate genotype calls. Given that poorly sequenced individuals and loci can bias downstream inference (97), we then removed calls for which the highest genotype calling confidence was less than 95%, removed individuals and loci which were called in less than 60% of loci or individuals, respectively, and identified and removed loci with minor alleles sequenced in less than five individuals with snpR (92). We identified six *Karst* loci with false discovery rates of less than 0.05 for association with diapause in our Australian samples and subsequently calculated allele frequencies for these six loci across all sampling locations.

## Linear mixed modeling of diapause association with Karst

To determine the degree to which *Karst* explained variation in diapause status in the Australian monarchs, we constructed linear mixed models using the imputed genotypic data for the six diapause-associated *Karst* loci described above. Specifically, we used the lmerTest R package (98) to construct two linear mixed effect models with quantitative diapause status as the response variable and maternal and paternal IDs from Colony2 (as described above) as random effects. In one model, we also included additive genotypes for the six *Karst* loci as fixed effects. We then used a two-way ANOVA test in R compare the two models and determine if adding genotypic effects improved model fit.

# Results

## Phenotypic results and sequencing

Of the 204 total adult female monarch butterflies reared in both these studies and by Freedman et al. (59), 103, 78, and 23 were classified as reproductively fully mature, partially mature, and immature, respectively. We obtained 396,512,006 total sequencing reads across all individuals, 99.4% of which mapped to the monarch reference genome and 60.7% of which were retained after filtering. From this, we called 179,735 and 437,259 SNP genotypes and likelihoods, respectively. Using the previously published North American and Pacific monarch samples (56), we called a total of 685 imputed SNPs in *Karst* in 189 samples after filtering, of which 82 were from North America. 124 of the monarch samples collected for this study also passed filtering in the imputed *Karst* dataset.

## Genetic basis of diapause

Maternal family strongly improved model fit (ANOVA, , see Figure 2b). From a total of 437,259 SNPs used in the association test, we identified three genomic regions which were significantly associated with reproductive development after false discovery rate correction on chromosomes 17, 19, and 29 (Figure 2d). A total of 26 genes were within 50kb of these regions (Table S1), one of which (LOC116776761/DPOGS204613/*Karst*) was located directly under the significantly associated SNPs on chromosome 29 (minimum p-value = 0.013 (Figure 3). No annotated genes were directly under the peaks of association on chromosomes 17 and 19 (Figure 3). Genotypes for the six *Karst* loci with a false discovery rate less than 0.05 explained roughly 13.6% (marginal r2 = 0.135) of the variation in reproductive development after accounting for the effects of maternal and parental family. Given the tight physical linkage between these SNPs, all but one (SNP 1) were dropped from the model due to rank-deficiency. The relationship between diapause and additive genotype at that locus was highly significant (p < 0.0001), and the model with genotypic data was significantly preferred over the model without (two-way ANOVA, p < 0.0001).

We found that wing length was significantly negatively correlated with the number of yolked oocytes in Freedman et al.’s (58) earlier study ( = 0.034) after accounting for maternal family. Wing shape was also negatively correlated with the number of yolked oocytes, but not significantly so after accounting for maternal family ( = 0.26).

## Karst genotypes in North America and across the Pacific

The *Karst* SNPs associated with diapause in Australia were present in North America and across the Pacific at varying frequencies (Figure 4). In general, the migration associated alleles varied strongly in frequency across the Pacific, were absent or very rare in Hawaii, and moderate in frequency in both Australia and in North America, with four out of the six more common in Australia and two more common in North America (Figure 4b). The most strongly associated locus (SNP1) was more common in Australia than in North America, absent in Hawaii and several Pacific islands, and most abundant on Rota and Fiji (Figure 1a).

# Discussion

## Evolution of diapause control in Australian monarchs

We found that a region on Chromosome 29 containing the gene DPOGS204613 (homologous to the gene *Karst* in *Drosophila* and hereafter referred to as such) was strongly associated with diapause initiation in monarch butterflies in Queensland, Australia (Figure 3c), explaining roughly 14% of the variation in reproductive development in our study. According to the most recent annotation (84), *Karst* codes for a spectrin beta chain protein involved in actin filament binding. *Karst* has not been previously identified as associated with migratory life history variation in North American monarchs. Specifically, *Karst* was not among the 536 genes located in regions of the genome found to be significantly associated with contrasts between migration and residency in North and South American, Pacific, and European monarch populations by Zhan et al. (71), although we found that the *Karst* genotypes associated with diapause in Australia are present in North America at low frequencies (Figure 4).

The involvement of *Karst* in migratory diapause initiation is perhaps not surprising. Juvenile hormone mediated actin contraction is known to be involved in oocyte formation (99), and mutations in *Karst* are known to cause serious issues during oogenesis in *Drosophila* (100). The involvement of *Karst* in oocyte development is likely conserved in Lepidoptera: a predicted homolog of *Karst*, BGIBMGA012171, has been identified as an ovary-specific expressed protein in silkworms (*Bombyx mori*) (101). More directly, while it does not overlap with any of genes identified as involved in photoperiodic responses by Iiams et al. (64), *Karst* has been observed to be differentially expressed during diapause termination in western North American monarchs (66). Finally, *Karst* has been shown to be a key regulator of the *Hippo* pathway in *Drosophila*, in which RNAi knockouts of *Karst* have been shown to produce noticeable wing overgrowth (102). Thus, variation at the *Karst* locus in monarchs may contribute to differences in reproductive development as well as variation in wing size/elongation.

Given that we found that the *Karst* loci associated with diapause in Australia were present in North America, the *Karst* migratory variance was either maintained as monarchs passed through repeated population bottlenecks during their Pacific expansion or was secondarily re-introduced by immigrant individuals from North America (which would have likely been anthropogenic in origin). The latter seems unlikely given that (1) a secondary re-introduction of migratory monarch butterflies from North America would have resulted in the transfer of other migration associated and neutral genetic variation; (2) we did not find any of the previously known migratory associations in Australia; and (3) monarchs in Australia are generally very different from those in North America at neutral loci (56). Regardless, it is surprising that *Karst* has not been previously identified to differ between migratory and non-migratory monarchs outside of Australia, suggesting that the mechanisms governing migration in monarchs worldwide are not yet fully understood.

In their 2014 study, Zhan et al generally focused their discussion of monarch migratory genetics on Collagen IV -1, one of three genes they found that were strong outliers for association with migration. This gene is involved in muscle functioning in insects (103), and Zhan et al. hypothesized that divergence at this gene was driven by selection for increased muscle efficiency to facilitate long-distance migration. While they mentioned that the remaining significantly associated genes were enriched for the “morphogenesis, neurogenesis, and extracellular matrix/basement membrane” functional terms, they were not otherwise discussed (*Karst* does not have any of these functional terms). It is possible that *Karst* also contributes to diapause and migration in butterflies outside of Australia, but that the quantitative effect of other genes (a product of the effect sizes of the migratory alleles and their frequencies) dwarfs that of *Karst*, thus concealing the relative impact of the latter gene.

However, it does make some sense that variance in Collagen IV -1 would not be key to migratory life history variation in populations that are *newly* migratory. Variance in genes that code for traits such as muscle strength or endurance, such as Collagen IV -1, could make already migratory individuals more fit, but would not itself *cause* or *enable* migratory behavior. *Karst* and other genes that are associated with reproductive diapause may therefore be involved earlier in the evolution of migration, since individuals that do not delay their reproductive investment have much shorter lifespans and often cannot complete their full migration (60). Genes that trigger migratory behavior or control orientation, navigation, and directed flight would also fall into this category to varying extents. That we found *Karst* specifically to be associated with diapause in Australia is therefore not surprising for a newly migratory population (104,105).

Regardless, it is important to note that it is unlikely that *Karst* alone controls diapause induction under declining daylength in Australian monarchs. We detected two other strong associations with diapause, but it is not clear to which genes these outliers correspond (Figure 3a-b). There are a few potential reasons for this: first, it is possible that the causal variants are in transcription factor binding sites or other promoter regions which are not well annotated. Alternatively, the sequencing data in this study is relatively low in resolution, and we therefore may not have data for the actual causal variants with which our loci are in partial linkage disequilibrium. Higher resolution sequencing could again help clarify the causal genes for these regions. Additionally, while our study features many *individuals*, it features only 32 maternal families and thus far fewer independent samples. Our power is therefore limited, and it is entirely likely that we failed to detect many causal loci for diapause onset. Additional studies with larger sample sizes are therefore still needed to better understand the mechanisms underlying diapause induction in Australian monarch butterflies. Lastly, while the correlation between diapause and wing morphology which we observed does support the use of diapause as a proxy for the general migratory syndrome in Australian monarchs, future work examining correlations between *Karst* or genotypes at other loci and other migratory traits, such as directed flight, could shed further light on the evolution of migration in Australian monarchs, as could sequencing and phenotyping monarchs from other migratory populations in Australia and New Zealand.

## Persistence of a migratory life history in Australian monarchs

Our work suggests that even after ceasing migration for hundreds of generations, monarchs may maintain genetic variation that underlies migratory plasticity, as they did in Australia despite the loss of allelic variance during multiple successive bottlenecks. This gives us some hope that the contemporary loss of migration that we have observed in many migratory species may be reversible over relatively short evolutionary timescales. At the very least, we can be assured that for monarchs in particular, migratory populations in Australia constitute a reservoir of migratory alleles that could potentially be tapped for North America if needed.

While monarchs may be able to quickly recover from the loss of migratory behavior, it is unlikely that this is the case for most other migratory species of conservation concern, particularly for vertebrates. For example, if we assume the latter hypothesis (migratory variance in Australian monarchs is the result of a new mutation), it is clear that monarchs have had ample evolutionary opportunity to re-acquire migratory mutations: monarchs have a generation time of approximately seven generations per year if continuously breeding (74), which means that monarchs in Australia went through roughly ~ 350 generations in the 50 or so years since they were first reported on the continent, and although they likely experienced a strong bottleneck initially, they were probably at a relatively large effective size for most of that time. Since the rate at which new mutations appear (and the rate at which standing variation is lost) in a population is proportional to both effective populations size and generation time (i.e. “mutation-drift equilibrium) (106), monarchs have had a large opportunity space for the generation of new migratory alleles and the maintenance of old ones.

Additionally, the maintenance of ancestral North American migratory variance in Australian monarchs was likely only possible because selection was probably not actively acting against migration in the Pacific, where monarchs are not exposed to substantial seasonal changes in day length or average temperature and thus do not receive the primary cues thought to be associated with migration initiation. Environmentally triggered, phenotypically plastic migratory variation, therefore, is probably nearly neutral in the Pacific. In contrast, individually fixed migratory variance, such as wing morphology, was likely selected against across the Pacific, thus driving the observed, repeated decreases in wing size and length observed in newly non-migratory monarch populations (107). Environmentally triggered migratory-associated genetic variation is therefore more likely to be maintained in non-migratory populations than that which underlies phenotypically fixed traits.

# Conclusions

This study suggests that reproductive diapause in Australian monarch butterflies is partially influenced by a novel genetic mechanism via the *Karst* gene, and that the previously identified variable migratory genomic regions were not associated with diapause in this populations. Diapause associated variation at *Karst* is also present in North America and is therefore likely ancestral, maintained over successive bottlenecks during the species’ expansion across the Pacific. Our work is consistent with the hypothesis that genetic variance which enables migration, rather than that which bolsters the fitness of already migratory individuals, is more likely to be observed in recently evolved migratory populations.

# Declarations

## Ethics and Consent to Participate

Permits are not required to collect or raise monarch butterflies in Queensland, Australia. All samples were collected on land belonging to the University of Queensland, and thus did not need any additional permits.

## Consent to Publish: not applicable

## Availability of Data and Materials

The raw sequencing data produced for this study are available from the NCBI under accession number PRJNA984345. The filtered genotypes and the scripts used to produce both them and this paper are available at <https://github.com/hemstrow/aus_monarchs>. The exact script used to produce this paper from the genotypic data, including all non-map figures, is available at <https://github.com/hemstrow/aus_monarchs/blob/master/paper/paper_draft.Rmd>.

## Competing Interestes

The authors declare no competing interests.

## Funding

Travel and sequencing funding was provided by the Jastro-Shields Graduate Research Award and the UC Davis Department of Animal Science. MF was funded by NSF Postdoctoral Fellowship Award 2010658.

## Author Contributions

WH and MM designed the study with input from MF and MZ. WH collected, raised, and phenotyped individuals. MZ provided, lab space, equipment, and expertise for each of those steps. WH analyzed the data and wrote the manuscript. MF, MZ, and MM provided feedback on the analysis and manuscript.

*Acknowledgements*

We would like to thank Dr. Shawan Chowdhury for his assistance with monarch feeding and maintenance during the rearing process and Dr. Sean O’Rourke for assistance during library preparation and sequencing.

# References

1. Both C, Van Turnhout CAM, Bijlsma RG, Siepel H, Van Strien AJ, Foppen RPB. Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. Proceedings of the Royal Society B: Biological Sciences. 2010 Apr 22;277(1685):1259–66.

2. Chowdhury S, Zalucki MP, Amano T, Woodworth BK, Venegas-Li R, Fuller RA. Seasonal spatial dynamics of butterfly migration. Ecology Letters. 2021 Sep 1;24(9):1814–23.

3. Robbins CS, Sauer JR, Greenberg RS, Droege S. Population declines in North American birds that migrate to the neotropics. Proceedings of the National Academy of Sciences. 1989 Oct 1;86(19):7658–62.

4. Jones T, Cresswell W. The phenology mismatch hypothesis: are declines of migrant birds linked to uneven global climate change? Journal of Animal Ecology. 2010 Jan 1;79(1):98–108.

5. Limburg KE, Waldman JR. Dramatic Declines in North Atlantic Diadromous Fishes. BioScience. 2009 Dec 1;59(11):955–65.

6. Runge CA, Martin TG, Possingham HP, Willis SG, Fuller RA. Conserving mobile species. Frontiers in Ecology and the Environment. 2014 Sep 1;12(7):395–402.

7. Pulido F, Berthold P. Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. Proceedings of the National Academy of Sciences. 2010;107(16):7341–6.

8. Rubolini D, AP M, Rainio K, Lehikoinen E. Intraspecific consistency and geographic variability in temporal trends of spring migration phenology among European bird species. Climate Research. 2007;35(1–2):135–46.

9. Gordo O. Why are bird migration dates shifting? A review of weather and climate effects on avian migratory phenology. Climate Research. 2007;35(1–2):37–58.

10. Chu JJ, Gillis DP, Riskin SH. Community science reveals links between migration arrival timing advance, migration distance and wing shape. Journal of Animal Ecology. 2022 Aug 1;91(8):1651–65.

11. Prytula ED, McKellar AE, Schwitters L, Reudink MW. Rapid advancement of spring migration and en route adjustment of migration timing in response to weather during fall migration in Vaux’s Swifts (*Chaetura vauxi*). Can J Zool. 2022 Jan 1;100(1):56–63.

12. Pinszke A, Remisiewicz M. Long-term changes in autumn migration timing of Garden Warblers *Sylvia borin* at the southern Baltic coast in response to spring, summer and autumn temperatures. The European Zoological Journal. 2023 Jul 11;90(1):283–95.

13. Mondain-Monval TO, Amos M, Chapman JL, MacColl A, Sharp SP. Flyway-scale analysis reveals that the timing of migration in wading birds is becoming later. Ecology and Evolution. 2021 Oct 1;11(20):14135–45.

14. Kovach RP, Gharrett AJ, Tallmon DA. Genetic change for earlier migration timing in a pink salmon population. Proceedings of the Royal Society B: Biological Sciences. 2012 Jul 11;279(1743):3870–8.

15. Shuert CR, Marcoux M, Hussey NE, Heide-Jørgensen MP, Dietz R, Auger-Méthé M. Decadal migration phenology of a long-lived Arctic icon keeps pace with climate change. Proceedings of the National Academy of Sciences. 2022 Nov 8;119(45):e2121092119.

16. Culbertson KA, Garland MS, Walton RK, Zemaitis L, Pocius VM. Long-term monitoring indicates shifting fall migration timing in monarch butterflies (*Danaus plexippus*). Global Change Biology. 2022 Feb 1;28(3):727–38.

17. Fitchett JM, Pandazis A, Pillay S. Advance in the timing of the annual migration of the brown-veined white butterfly through Johannesburg, South Africa, over the period 1914–2020. International Journal of Biometeorology. 2022 Nov 1;66(11):2251–8.

18. Ambrosini R, Rubolini D, AP M, Bani L, Clark J, Karcza Z, et al. Climate change and the long-term northward shift in the African wintering range of the barn swallow *Hirundo rustica*. Climate Research. 2011;49(2):131–41.

19. La Sorte FA, Thompson III FR. Poleward shifts in winter ranges of north american birds. Ecology. 2007 Jul 1;88(7):1803–12.

20. Visser ME, Perdeck AC, Van Balen J, Both C. Climate change leads to decreasing bird migration distances. Global Change Biology. 2009 Aug 1;15(8):1859–65.

21. Sonnleitner J, LaZerte SE, McKellar AE, Flood NJ, Reudink MW. Rapid shifts in migration routes and breeding latitude in North American bluebirds. Ecosphere. 2022 Dec 1;13(12):e4316.

22. Rushing CS, Royle JA, Ziolkowski DJ, Pardieck KL. Migratory behavior and winter geography drive differential range shifts of eastern birds in response to recent climate change. Proceedings of the National Academy of Sciences. 2020 Jun 9;117(23):12897–903.

23. Dufour P, Åkesson S, Hellström M, Hewson C, Lagerveld S, Mitchell L, et al. The Yellow-browed Warbler (*Phylloscopus inornatus*) as a model to understand vagrancy and its potential for the evolution of new migration routes. Movement Ecology. 2022 Dec 14;10(1):59.

24. Dufour P, de Franceschi C, Doniol-Valcroze P, Jiguet F, Guéguen M, Renaud J, et al. A new westward migration route in an Asian passerine bird. Current Biology. 2021 Dec 20;31(24):5590-5596.e4.

25. Jia Y, Liu Y, Jiao S, Guo J, Lu C, Zhou Y, et al. Shifting of the Migration Route of White-Naped Crane (*Antigone vipio*) Due to Wetland Loss in China. Remote Sensing. 2021;13(15).

26. Plummer KE, Siriwardena GM, Conway GJ, Risely K, Toms MP. Is supplementary feeding in gardens a driver of evolutionary change in a migratory bird species? Global Change Biology. 2015 Dec 1;21(12):4353–63.

27. Møller AP, Jokimäki J, Skorka P, Tryjanowski P. Loss of migration and urbanization in birds: a case study of the blackbird (*Turdus merula*). Oecologia. 2014;175(3):1019–27.

28. Haran R, Kiat Y, Izhaki I. Patterns of partial migration of the Dead Sea Sparrow (*Passer moabiticus*) along the Great Rift Valley in Israel. Journal of Arid Environments. 2021 Sep 1;192:104534.

29. Falchi V, Cerritelli G, Barbon A, Catoni C, Cutroneo A, dell’Omo G, et al. Inter and intra-population variability of the migratory behaviour of a short-distance partial migrant, the Eurasian Stone-curlew *Burhinus oedicnemus* (Charadriiformes, Burhinidae). Journal of Ornithology. 2023 Jan 1;164(1):85–100.

30. Bonnet-Lebrun AS, Manica A, Rodrigues ASL. Effects of urbanization on bird migration. Biological Conservation. 2020 Apr 1;244:108423.

31. Howard E, Aschen H, Davis AK. Citizen Science Observations of Monarch Butterfly Overwintering in the Southern United States. Rust M, editor. Psyche. 2010;2010:689301.

32. Van Der Ree R, McDonnell MJ, Temby I, Nelson J, Whittingham E. The establishment and dynamics of a recently established urban camp of flying foxes (*Pteropus poliocephalus*) outside their geographic range. Journal of Zoology. 2006 Feb 1;268(2):177–85.

33. Fleming TH, Eby P, Kunz TH, Fenton MB. Ecology of bat migration. Bat ecology. 2003;156:164–5.

34. Jones JD, Kauffman MJ, Monteith KL, Scurlock BM, Albeke SE, Cross PC. Supplemental feeding alters migration of a temperate ungulate. Ecological Applications. 2014 Oct 1;24(7):1769–79.

35. Adams CE, Chavarie L, Rodger JR, Honkanen HM, Thambithurai D, Newton MP. An opinion piece: the evolutionary and ecological consequences of changing selection pressures on marine migration in Atlantic salmon. Journal of Fish Biology. 2022 Apr 1;100(4):860–7.

36. McWilliams SR, Guglielmo C, Pierce B, Klaassen M. Flying, fasting, and feeding in birds during migration: a nutritional and physiological ecology perspective. Journal of Avian Biology. 2004 Sep 1;35(5):377–93.

37. Rankin MA, Burchsted JCA. The Cost of Migration in Insects. Annual Review of Entomology. 1992 Jan 1;37(1):533–59.

38. Wikelski M, Tarlow EM, Raim A, Diehl RH, Larkin RP, Visser GH. Costs of migration in free-flying songbirds. Nature. 2003;423(6941):704.

39. Pulido F. The Genetics and Evolution of Avian Migration. BioScience. 2007 Feb 1;57(2):165–74.

40. López-Hoffman L, Chester CC, Semmens DJ, Thogmartin WE, Rodríguez-McGoffin MS, Merideth R, et al. Ecosystem Services from Transborder Migratory Species: Implications for Conservation Governance. Annu Rev Environ Resour. 2017 Oct 17;42(1):509–39.

41. Semmens DJ, Diffendorfer JE, López-Hoffman L, Shapiro CD. Accounting for the ecosystem services of migratory species: Quantifying migration support and spatial subsidies. Ecological Economics. 2011 Oct 15;70(12):2236–42.

42. Burger J, Gochfeld M, Niles L, Powers C, Brown K, Clarke J, et al. Complexity of bioindicator selection for ecological, human, and cultural health: Chinook salmon and red knot as case studies. Environmental Monitoring and Assessment. 2015 Feb 10;187(3):102.

43. Schuetz JG, Johnston A. Characterizing the cultural niches of North American birds. Proceedings of the National Academy of Sciences. 2019 May 28;116(22):10868–73.

44. Semmens DJ, Diffendorfer JE, Bagstad KJ, Wiederholt R, Oberhauser K, Ries L, et al. Quantifying ecosystem service flows at multiple scales across the range of a long-distance migratory species. Ecosystem Services. 2018 Jun 1;31:255–64.

45. Reynolds HL, Clay K. Migratory species and ecological processes. Envtl L. 2011;41:371.

46. Brower LP, Malcolm SB. Animal Migrations: Endangered Phenomena. American Zoologist. 1991 Feb 1;31(1):265–76.

47. Satterfield DA, Maerz JC, Hunter MD, Flockhart DTT, Hobson KA, Norris DR, et al. Migratory monarchs that encounter resident monarchs show life-history differences and higher rates of parasite infection. Ecology Letters. 2018 Nov 1;21(11):1670–80.

48. Semmens BX, Semmens DJ, Thogmartin WE, Wiederholt R, López-Hoffman L, Diffendorfer JE, et al. Quasi-extinction risk and population targets for the Eastern, migratory population of monarch butterflies (*Danaus plexippus*). Scientific Reports. 2016;6(1):23265.

49. Crossley MS, Meehan TD, Moran MD, Glassberg J, Snyder WE, Davis AK. Opposing global change drivers counterbalance trends in breeding North American monarch butterflies. Global Change Biology. 2022 Aug 1;28(15):4726–35.

50. Batalden RV, Oberhauser KS. Potential changes in eastern North American monarch migration in response to an introduced milkweed, Asclepias curassavica. Monarchs in a changing world: biology and conservation of an iconic butterfly. 2015;215–24.

51. Walker A, Oberhauser KS, Pelton EM, Pleasants JM. Danaus plexippus (errata version published in 2022). The IUCN Red List of Threatened Species. 2022;2022: e.T159971A219149911.

52. Mongue AJ, Tsai MV, Wayne ML, de Roode JC. Inbreeding depression in monarch butterflies. Journal of Insect Conservation. 2016 Jun 1;20(3):477–83.

53. Gustafsson KM, Agrawal AA, Lewenstein BV, Wolf SA. The Monarch Butterfly through Time and Space: The Social Construction of an Icon. BioScience. 2015 Jun 1;65(6):612–22.

54. Zalucki MP, Clarke AR. Monarchs across the Pacific: The Columbus hypothesis revisited. Biological Journal of the Linnean Society. 2004 May 1;82(1):111–21.

55. Pierce AA, Zalucki MP, Bangura M, Udawatta M, Kronforst MR, Altizer S, et al. Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies. Proceedings of the Royal Society B: Biological Sciences. 2014 Dec 22;281(1797):1–9.

56. Hemstrom WB, Freedman MG, Zalucki MP, Ramírez SR, Miller MR. Population genetics of a recent range expansion and subsequent loss of migration in monarch butterflies. Molecular Ecology. 2022 Sep 1;31(17):4544–57.

57. James DG, James TA. Migration and Overwintering in Australian Monarch Butterflies (Danaus plexippus (L.) (Lepidoptera: Nymphalidae): a Review with New Observations and Research Needs. The Journal of the Lepidopterists’ Society. 2019 Dec 1;73(3):177–90.

58. James DG. Induction of Reproductive Dormancy in Australian Monarch Butterflies *Danaus plexippus* (L.). Aust J Zool. 1983;31(4):491–8.

59. Freedman MG, Dingle H, Tabuloc CA, Chiu JC, Yang LH, Zalucki MP. Non-migratory monarch butterflies, *Danaus plexippus* (L.), retain developmental plasticity and a navigational mechanism associated with migration. Biological Journal of the Linnean Society. 2018 Jan 15;123(2):265–78.

60. Herman WS, Tatar M. Juvenile hormone regulation of longevity in the migratory monarch butterfly. Proceedings of the Royal Society of London Series B: Biological Sciences. 2001;268(1485):2509–14.

61. Denlinger DL. Insect diapause: from a rich history to an exciting future. Journal of Experimental Biology. 2023 Feb 28;226(4):jeb245329.

62. Numata H, Shintani Y. Diapause in Univoltine and Semivoltine Life Cycles. Annu Rev Entomol. 2023 Jan 23;68(1):257–76.

63. Goehring L, Oberhauser KS. Effects of photoperiod, temperature, and host plant age on induction of reproductive diapause and development time in *Danaus plexippus*. Ecological Entomology. 2002 Dec 1;27(6):674–85.

64. Iiams SE, Lugena AB, Zhang Y, Hayden AN, Merlin C. Photoperiodic and clock regulation of the vitamin A pathway in the brain mediates seasonal responsiveness in the monarch butterfly. Proceedings of the National Academy of Sciences. 2019 Dec 10;116(50):25214–21.

65. Saunders DS. Dormancy, Diapause, and the Role of the Circadian System in Insect Photoperiodism. Annu Rev Entomol. 2020 Jan 7;65(1):373–89.

66. Green II DA, Kronforst MR. Monarch butterflies use an environmentally sensitive, internal timer to control overwintering dynamics. Molecular Ecology [Internet]. 2019 Jul 24;0(ja). Available from: https://doi.org/10.1111/mec.15178

67. Majewska AA, Altizer S. Exposure to Non-Native Tropical Milkweed Promotes Reproductive Development in Migratory Monarch Butterflies. Insects. 2019;10(8).

68. Herman WS. Studies on the adult reproductive diapause of the monarch butterfly, *danaus plexippus*. The Biological Bulletin. 1981 Feb 1;160(1):89–106.

69. Tenger-Trolander A, Lu W, Noyes M, Kronforst MR. Contemporary loss of migration in monarch butterflies. Proceedings of the National Academy of Sciences. 2019 Jul 16;116(29):14671 LP – 14676.

70. Opler PA. A field guide to western butterflies. Houghton Mifflin Harcourt; 1999.

71. Zhan S, Zhang W, Niitepõld K, Hsu J, Haeger JF, Zalucki MP, et al. The genetics of monarch butterfly migration and warning colouration. Nature. 2014 Oct 1;514:317.

72. Haeger J, Jordano D, Zalucki M. Monarchs across the Atlantic Ocean: what’s happening on the other shore. Monarchs in a changing world: biology and conservation of an iconic butterfly Cornell University Press, Ithaca, NY. 2015;247–56.

73. Malcolm S, Cockrell B, Brower L. Spring recolonization of eastern North America by the monarch butterfly: successive brood or single sweep migration. Biology and conservation of the monarch butterfly. 1993;38:253–67.

74. Cockrell BJ, Malcolm SB, Brower LP. Time, temperature, and latitudinal constraints on the annual recolonization of eastern North America by the monarch butterfly. Biology and conservation of the monarch butterfly. 1993;38:233–51.

75. Dockx C. Directional and stabilizing selection on wing size and shape in migrant and resident monarch butterflies, *Danaus plexippus* (L.), in Cuba. Biological Journal of the Linnean Society. 2007;92(4):605–16.

76. Altizer S, Davis AK. Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. Evolution. 2010 Apr 1;64(4):1018–28.

77. North J. The monarch butterfly: biology & conservation. Cornell university press; 2004.

78. Clarke AR, Zalucki MP. Monarchs in Australia: On the Winds of a Storm? Biological Invasions. 2004 Mar 1;6(1):123–7.

79. Smith RL. Sperm competition and the evolution of animal mating systems. Elsevier; 2012.

80. Zalucki MP. The effects of age and weather on egg laying in *Danaus plexippus* L. (Lepidoptera: Danaidae). Researches on Population Ecology. 1981 Dec 1;23(2):318–27.

81. Pan ML, Wyatt GR. Control of vitellogenin synthesis in the Monarch butterfly by juvenile hormone. Developmental Biology. 1976;54(1):127–34.

82. James DG. Ovarian dormancy in *Danaus plexippus* (l.) (lepidoptera: nymphalidae)—oligopause not diapause. Australian Journal of Entomology. 1982 Feb 1;21(1):31–5.

83. Ali OA, O’Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, et al. Rad capture (Rapture): Flexible and efficient sequence-based genotyping. Genetics. 2016;202(2):389–400.

84. Ranz JM, González PM, Clifton BD, Nazario-Yepiz NO, Hernández-Cervantes PL, Palma-Martínez MJ, et al. A de novo transcriptional atlas in *Danaus plexippus* reveals variability in dosage compensation across tissues. Communications Biology. 2021 Jun 25;4(1):791.

85. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England). 2009 Jul 15;25(14):1754–60.

86. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Jun 8;25(16):2078–9.

87. Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: Analysis of Next Generation Sequencing Data. BMC Bioinformatics. 2014 Nov;15(1):356.

88. Jones OR, Wang J. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources. 2010 May 1;10(3):551–5.

89. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria; 2022. Available from: https://www.R-project.org/

90. Meisner J, Albrechtsen A. Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. Genetics. 2018 Oct 1;210(2):719–31.

91. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995 Jan 1;57(1):289–300.

92. Hemstrom W, Jones M. snpR: User friendly population genomics for SNP data sets with categorical metadata. Molecular Ecology Resources. 2023 May 1;23(4):962–73.

93. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990 Oct 5;215(3):403–10.

94. Zhan S, Reppert SM. MonarchBase: the monarch butterfly genome database. Nucleic Acids Research. 2013 Jan 1;41(D1):D758–63.

95. Pinheiro J, Bates D, DebRoy S, Sarkar D, Team RC. Linear and nonlinear mixed effects models. R package version. 2007;3(57):1–89.

96. Browning SR, Browning BL. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. The American Journal of Human Genetics. 2007 Nov 1;81(5):1084–97.

97. Hemstrom W, Grummer JA, Luikart G, Christie MR. Next-generation data filtering in the genomics era. Nature Reviews Genetics. 2024 Nov 1;25(11):750–67.

98. Kuznetsova A, Brockhoff PB, Christensen RHB. {lmerTest} Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software. 2017;82(13):1–26.

99. Luo W, Liu S, Zhang W, Yang L, Huang J, Zhou S, et al. Juvenile hormone signaling promotes ovulation and maintains egg shape by inducing expression of extracellular matrix genes. Proceedings of the National Academy of Sciences. 2021 Sep 28;118(39):e2104461118.

100. Zarnescu DC, Thomas CM. Apical Spectrin Is Essential for Epithelial Morphogenesis but Not Apicobasal Polarity in Drosophila. Journal of Cell Biology. 1999 Sep 6;146(5):1075–86.

101. Chen J e, Li J ying, You Z ying, Liu L li, Liang J she, Ma Y ying, et al. Proteome Analysis of Silkworm, *Bombyx mori*, Larval Gonads: Characterization of Proteins Involved in Sexual Dimorphism and Gametogenesis. J Proteome Res. 2013 Jun 7;12(6):2422–38.

102. Fletcher GC, Elbediwy A, Khanal I, Ribeiro PS, Tapon N, Thompson BJ. The Spectrin cytoskeleton regulates the Hippo signalling pathway. The EMBO Journal. 2015 Apr 1;34(7):940–54.

103. Schnorrer F, Schonbauer C, Langer CCH, Dietzl G, Novatchkova M, Schernhuber K, et al. Systematic genetic analysis of muscle morphogenesis and function in *Drosophila*. Nature. 2010;464(7286):287–91.

104. Merlin C, Liedvogel M. The genetics and epigenetics of animal migration and orientation: birds, butterflies and beyond. el Jundi B, Kelber A, Webb B, editors. Journal of Experimental Biology. 2019 Feb 6;222(Suppl\_1):jeb191890.

105. Green DA. Monarch Butterfly Migration as an Integrative Model of Complex Trait Evolution. The American Naturalist. 2021 Jul 1;198(1):142–57.

106. Crow JF. An introduction to population genetics theory. Scientific Publishers; 2017.

107. Freedman MG, Dingle H, Strauss SY, Ramírez SR. Two centuries of monarch butterfly collections reveal contrasting effects of range expansion and migration loss on wing traits. Proceedings of the National Academy of Sciences of the United States of America. 2020 Nov 17;117(46):28887–93.

# Tables and Figures

## Table 1

Total count of immature, partially mature, and mature female monarch butterflies reared and sequenced in each sampling year.

| Year | Immature | Partially Mature | Mature | Total Individuals | Maternal Families |
| --- | --- | --- | --- | --- | --- |
| 2016 | 12 | 22 | 6 | 40 | 12 |
| 2018 | 11 | 56 | 97 | 164 | 20 |
| Total | 23 | 78 | 103 | 204 | 32 |

## Figure 1:

Approximate ancestral North American and introduced migratory and non-migratory ranges with introduction dates for monarch butterflies in the Pacific, with sampling location noted in blue. Photo courtesy of William Hemstrom.

## Figure 2:

Genome-wide association results for diapause in Australian monarch butterflies. A) Example egg development for the three stages used to phenotype adult females as reproductively mature, partially mature, and immature. B) Average reproductive development scores for maternal families. Average scores were significantly more skewed than expected by chance (orange line). C) Quantile-quantile plot for association results showing a few strong outliers but otherwise conformance to the expected distribution of p-values. D) Manhattan plot showing the strength of association with reproductive status genome-wide. Values above the blue line indicate significant associations after false discovery rate p-value adjustment.

## Figure 3:

Distribution of genome-wide association test p-values adjacent to each of the three outlier regions (top). Genes that are at least partially within 50kb of outlier SNPs are noted (bottom), with CDS regions highlighted in black. One specific gene (LOC133319386) had no match in the monarch official gene set, and so is instead listed with its RefSeq ID.

## Figure 4:

Distribution of diapause associated *Karst* alleles in North America and across the Pacific. A) Frequencies for the diapause associated allele for the most *Karst* SNP locus most strongly associated with reproductive status (SNP1) in North America and across the Pacific to Australia. Circles with a red “X” indicate no detected diapause alleles. B) Frequencies of diapause allele at all six of the significant *Karst* SNP loci in North America, Australia, Hawaii, and on other islands across the Pacific.