ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Population genetics of a recent range expansion and subsequent loss of migration in monarch butterflies

William B. Hemstrom[1] | Micah G. Freedman[2,3] | Myron P. Zalucki[4] |
Santiago R. Ramírez[2,3] | Michael R. Miller[1]

[1]Department of Animal Science, University of California, Davis, California, USA

[2]Department of Evolution and Ecology, University of California, Davis, California, USA

[3]Center for Population Biology, University of California, Davis, California, USA

[4]School of Biological Sciences, The University of Queensland, Brisbane, Australia

Correspondence
Micah G. Freedman, Department of Ecology & Evolution, University of Chicago, Chicago, IL, USA.
Email: micahfreedman@uchicago.edu

Abstract

Range expansions—whether permanent or transient—strongly influence the distribution of genetic variation in space. Monarch butterflies are best known for long-distance seasonal migration within North America but are also established as nonmigratory populations around the world, including on Pacific Islands. Previous research has highlighted stepwise expansion across the Pacific, though questions remain about expansion timing and the population genetic consequences of migration loss. Here, we present reduced-representation sequencing data for 275 monarchs from North America ($n = 85$), 12 Pacific Islands ($n = 136$) and three locations in Australia ($n = 54$), with the goal of understanding (i) how the monarch's Pacific expansion has shaped patterns of population genetic variation and (ii) how loss of migration has influenced spatial patterns of differentiation. We find support for previously described stepwise dispersal across the Pacific and document an additional expansion from Hawaii into the Mariana Islands. Nonmigratory monarchs within the Mariana Islands show strong patterns of differentiation, despite their proximity; by contrast, migratory North American samples form a single genetically panmictic population across the continent. Estimates of Pacific establishment timing are highly uncertain (~100–1,000,000 years ago) but overlap with historical records that indicate a recent expansion. Our data support (i) a recent expansion across the Pacific whose timing overlaps with available historical records of establishment and (ii) a strong role for seasonal migration in determining patterns of spatial genetic variation. Our results are noteworthy because they demonstrate how the evolution of partial migration can drive population differentiation over contemporary timescales.

KEYWORDS
monarch butterfly, population genomics, range expansion, serial dispersal

---

William B. Hemstrom and Micah G. Freedman contributed equally to this work.

1 | INTRODUCTION

Over extended timescales, geographical range expansions generally involve decreasing relatedness and increasing contributions of genetic drift in populations further from the original source population (Excoffier et al., 2009; Hewitt, 1996). This pattern is especially evident in serial stepwise expansion events, in which populations are founded in a stepping-stone fashion (Ibrahim et al., 1996; Slatkin & Excoffier, 2012). Serial dispersal is characteristic of many postglacial range expansions into temperate regions and has been shown for species including eider ducks (*Somateria mollissima*) (Tiedemann et al., 2004), ragwort (*Senecio helleri*) (Bettin et al., 2007), rough-skinned newts (*Taricha granulosa*) (Kuchta & Tan, 2005) and European butterflies (Dapporto et al., 2019).

Studies on the population genetics of geographical range expansions tend to focus on expansion events that occur over extended timescales, though the same expansion processes characterize annual movements associated with seasonal migration. Unlike permanent range expansion, however, seasonal migration may involve individuals capable of making round-trip journeys and traversing the entire species range in their lifetime (Dingle, 2014), thereby limiting opportunities for genetic divergence in allopatry. In species that migrate seasonally, patterns of population genetic variation in space are best captured by considering migratory connectivity of breeding populations (e.g., Cohen et al., 2018; Gao et al., 2020). For example, Wilson's warbler (*Wilsonia pusilla*) has eastern and western North American summer breeding populations that are genetically distinct, despite sharing an overwintering range in Central America (Irwin et al., 2011), and anadromous Coho (*Oncorhynchus tshawytscha*) and steelhead (*O. mykiss*) salmon populations show strong genetic differentiation corresponding to the river drainages where they spawn (Prince et al., 2017; Waples et al., 2004). By contrast, Japanese eels (*Anguilla japonica*) that share common spawning grounds but migrate to disparate areas across temperate Asia show little evidence for genetic differentiation over time or space (Gong et al., 2019).

Migratory species that show evidence for partial migration, whereby species comprise both migratory and nonmigratory populations (Chapman et al., 2011), often feature populations that are highly differentiated, both phenotypically (e.g., Altizer & Davis, 2010; Dingle et al., 1980) and genetically (e.g., Gómez-Bahamón et al., 2020; Zhan et al., 2014). The phenomenon of partial migration is common across the tree of life and has been documented in birds (Adriaensen & Dhondt, 1990), insects (Menz et al., 2019) and ungulates (Berg et al., 2019). Although the evolutionary origins of partial migration are sometimes unclear, one recently invoked scenario involves a migratory, geographically widespread lineage giving rise to one or more nonmigratory descendant lineages that become genetically distinct due to mismatches in the timing and/or location of breeding. This scenario has been hypothesized to be an important contributor to patterns of speciation in tropical birds (e.g., Gómez-Bahamón et al., 2020; Kondo et al., 2008) and may also contribute to diversification of other groups.

Monarch butterflies (*Danaus plexippus* [L.]) provide an intriguing system for studying the effects of both global range expansion and loss of migration on spatial population genetic structure. Migratory North American monarchs comprise a single genetically indistinguishable population (Lyons et al., 2012; Talla et al., 2020) and have a summer breeding range that covers most of the North American continent. Over recent evolutionary history, monarchs have expanded their range globally (Ackery & Vane-Wright, 1984; Fernández-Haeger et al., 2015; Pierce et al., 2015; Pierce, Zalucki, et al., 2014a; Vane-Wright, 1993; Zalucki & Clarke, 2004; Zhan et al., 2014), with a southern expansion into South America and the Caribbean, an eastward expansion across the Atlantic and into the Iberian Peninsula, and a westward expansion across the Pacific. In contrast to their migratory North American ancestors, these expansion populations generally form nonmigratory, year-round breeding populations in areas where they become established (Zhan et al., 2014). The exceptions to this pattern are in southern Australia and New Zealand, where monarchs move seasonally and form overwintering clusters akin to those observed in western North America (James, 1993; Wise, 1980). In this paper, we focus exclusively on the monarch's expansion and subsequent loss of migration in populations across the Pacific.

Little is currently known about how contemporary loss of migration has affected fine-scale patterns of population differentiation in monarchs or other taxa (but see Samarasin et al., 2017). Two studies have indirectly addressed this question in monarchs: Hughes and Zalucki (1984) used four allozyme markers and found relatively high overall F_{ST} (0.032) between monarchs in host plant patches over small spatial scales (tens to hundreds of kilometres) in Queensland, Australia. Pierce, de Roode et al. (2014b) used microsatellite markers and showed that monarchs from the Hawaiian archipelago show little differentiation among islands. However, the conclusions of these studies are limited by the spatial scale of sampling and the number of loci studied. Furthermore, the timing of the monarch's Pacific expansion remains uncertain. Demographic simulations indicate that establishment timing may have happened as long as 2000–3000 years ago (Zhan et al., 2014), although these estimates conflict with historical records, which suggest that expansion across the Pacific happened between the years 1840 and 1900 (Freedman et al., 2020; Zalucki & Clarke, 2004). Further demographic modelling that accommodates a broad range of potential establishment scenarios—including variable establishment timing, founding population sizes and past changes in population size—might help to resolve the discrepancy between historical and model-based estimates of expansion timing.

In this study, we sequenced 275 monarch butterflies at >70,000 highly variable genomic sites from the ancestral North American population and many Pacific Island populations, including a number of previously unsampled populations: the Mariana Islands (Guam, Rota and Saipan), Norfolk Island, Victoria and New South Wales. The goals of this study were to understand (i) overall patterns of genetic relatedness among Pacific and North American populations; (ii) the timing of range expansion and the amount of ongoing gene flow between North American and Pacific populations; and (iii) how

migratory and nonmigratory populations differ in their distribution of population genetic variation in space.

2 | METHODS

2.1 | Sample preparation and sequencing

Monarchs were collected as either larvae or adult butterflies from various locations across their current geographical range between 1990 and 2017 (Table S1). DNA was extracted from samples using a magnetic bead-based protocol (Ali et al., 2016) and quantified using Quant-iT PicoGreen dsDNA Reagent (Thermo Fisher Scientific) on an FLx800 Fluorescence Reader (BioTek Instruments). Restriction-associated digest (RAD) DNA libraries were then created using the *Pst*I restriction enzyme according to Ali et al. (2016) and sequenced using 150-bp paired-end sequencing on an Illumina Hi-Seq 4000.

2.2 | Sequence alignment, filtering and genotype calling

We aligned raw sequence data to version 3 of the monarch butterfly genome assembly (Zhan & Reppert, 2013) using the mem algorithm implemented in the Burrows–Wheeler Aligner (Li & Durbin, 2009). Sequence data were sorted and filtered for PCR (polymerase chain reaction) duplicates and improper pairs using SAMTOOLS (Li et al., 2009). We first removed potential paralogous sites using the NGSPARALOG tool by removing all sites within 1 kb of any single nucleotide polymorphism (SNP) with a log ratio test statistic of >10 in any population (Linderoth, 2018). From this, we created five different data sets using different filtering schemes appropriate for different downstream analysis.

2.3 | Data sets

1. For use in demographic reconstruction and directionality index (ψ) calculation, we called genotypes using the SAMTOOLS genotype likelihood model (Li et al., 2009) as implemented in the ANGSD software package with a minimum mapping and base call quality score of 20, a SNP p-value of 1e-8, a uniform genotype prior and a posterior genotype probability cutoff of 0.95 (Korneliussen et al., 2014). To reduce potential bias due to linkage for demographic analyses, we randomly subsampled SNPs such that no locus was within 10,000 bp of another using a custom R script. The resulting SNPs were used to calculate site frequency spectra (SFS) and then projected down to a sample size of 100 gene copies from North America and 10 from Hawaii for demographic analysis and down to 10 gene copies in each population for calculation of ψ using the methods described by Gutenkunst et al. (2009). These projection numbers were picked to maximize the remaining number of SNPs in the data set. The SFS was polarized via reference to whole genome sequence data of the best-sequenced individual of the monarch's sister species *Danaus erippus* (Zhan et al., 2014) by alignment to the monarch genome as described above. While we did not use a Hardy–Weinberg Equlibrium (HWE) filter here, only a very small proportion of our loci were consistently not in HWE across populations ($p < 1 \times 10^{-6}$ in only 86 out of 11,384 loci, calculated using the method of Wigginton et al., 2005).

2. For use in calculating basic diversity statistics (the average number of pairwise differences, or π, observed heterozygosity, or H_O, and the ratio of within-sample heterozygous to homozygous loci, or Het/Hom), the fixation index (F_{ST}), and isolation-by-distance (IBD), we called genotypes as in data set 1, then removed individuals genotyped at <75% of loci. Since strong bottlenecks are likely to cause large differences in allele frequencies between populations, which, in conjunction with very different sample sizes between populations, can result in loci with very low overall minor allele frequencies having relatively high frequencies in individual populations, we did not use a minor allele frequency filter when calling genotypes. Some populations did not remain in the analysis after this filtering step.

3. To calculate Tajima's *D*, we implemented the same filtering steps as described for data set 2, but without using an SNP p-value filter in ANGSD.

4. For analyses that did not require called genotypes (principal component analysis, or PCA, NGSadmix, and a neighbour-joining tree), we used ANGSD as in data set 1, but did not call genotypes and instead estimated the likelihoods with a minor allele frequency filter of 0.05. For the PCA and neighbour-joining tree, the input distance matrix was created using the Identity-by-State approach in ANGSD (Korneliussen et al., 2014). No individuals were removed for these analyses.

We also generated more thoroughly filtered versions of data sets 3 and 4 (see Supplementary Methods, Filtered Data sets), although these additional filtering steps did not meaningfully influence our inferences.

2.4 | Patterns of relatedness among monarch populations

We calculated π, H_O, Het/Hom and Tajima's *D* (Tajima, 1989) within each population using the SNPR package (Hemstrom & Jones, 2021). We calculated F_{ST} between populations using the R implementation of the GENEPOP software package (Rousset, 2008) with a minor allele frequency cutoff of 0.05 and bootstrapped individuals between populations randomly 1000 times to calculate F_{ST} significance levels using the SNPR package (Hemstrom & Jones, 2021). For each of these statistics, the eastern and western North American samples were pooled together, based on results from previous studies (Lyons et al., 2012; Talla et al., 2020). To determine if our relatively

light filtering approach biased our results, we also reran the diversity statistics π, H_O, Het/Hom and Tajima's D, F_{ST}, and IBD analyses using more heavily filtered data sets, as described in the Supporting Methods.

To describe the basic population structure, we created a neighbour-joining tree (Saitou & Nei, 1987) using the ape R package version 5.0 (Paradis & Schliep, 2019) and conducted a PCA. For comparison, NGSadmix was used to generate individual ancestry coefficients for each individual for between one and nine putative population clusters (k) (Skotte et al., 2013). Each value of k was run 10 times, and the results were collapsed into consensus plots using CLUMPP (Jakobsson & Rosenberg, 2007). The POPHELPER (Francis, 2017) and SNPR (Hemstrom & Jones, 2021) R packages were used to run these analyses. We used the method of Evanno et al. (2005) to detect the number of clusters present from the results; however, since this method has reproducibility issues (Gilbert et al., 2012), tends to underestimate the true k unless in the context of a complex hierarchical examination (Janes et al., 2017) and generally does not improve estimates of k (Waples & Gaggiotti, 2006), we also looked at patterns of clustering in general across a range of k values.

2.5 | Serial expansion

To quantify the direction and strength of population spread across the Pacific, we calculated ψ (Peter & Slatkin, 2013) for each pairwise combination of North America, Hawaii, Queensland, Guam, Rota and Norfolk Island populations using the SNPR package (Hemstrom & Jones, 2021).

2.6 | Genetic variation across space in migratory vs. nonmigratory populations

We looked for evidence of IBD between samples from the Mariana Islands (nonmigratory), Hawaii (nonmigratory), Australia (partially migratory), and North America (migratory). To do so we calculated Edwards' angular genetic distance (Edwards, 1971) between each pair of samples from the given populations, and then compared these distances to the geographical distances between samples using a Mantel Test (Mantel, 1967). Here, we expect stronger patterns of IBD within locations where monarchs have ceased seasonal migration. As with F_{ST}, we used data set 2 with an additional minor allele frequency cutoff of 0.05.

2.7 | Demographic history of the monarch's expansion

To describe the patterns of establishment and migration between North America and the Pacific, the demographic reconstruction program δαδι (hereafter DADI, Gutenkunst et al., 2009) was used

to estimate the demographic history of the North American and Hawaiian samples. Briefly, since demographic processes influence the frequency of common or rare alleles across loci, and the SFS describes how many individual loci fall into each possible allele rarity in each population, the SFS can be used to infer historical population processes. DADI therefore uses simulation to compare the SFS predicted under a specific demographic history to the SFS observed from the data in order to evaluate the likelihood of a demographic model and to optimize the parameters of that model. We chose to focus on Hawaii since previous work has suggested that this island group was probably the first in the Pacific colonized by monarchs (Zalucki & Clarke, 2004; Zhan et al., 2014), and thus the timing of the monarch introduction there represents the earliest possible time for any introductions in the Pacific. The general approach that we describe here is similar to that described in Zhan et al. (2014), but with a few key differences: (i) Zhan et al. (2014) pooled samples from six distinct Pacific populations for their analysis, while we only focused on the Hawaiian population; (ii) the parameter values reported in Zhan et al. (2014) represent a single instance of one focal demographic scenario, whereas our analyses considered a large set of candidate demographic scenarios and report results from multiple iterations of each scenario.

2.8 | DADI model selection

We fit a range of possible models to the observed data: (i) each of the models in the "Island Model" set described in the dadi_pipeline (Portik et al., 2017), which contains some models originally published in Charles et al. (2018); (ii) the model described by Zhan et al. (2014) for the same comparison; and (iii) a similar model that allowed for an additional period of growth prior to the establishment of the Hawaiian population and another following establishment (hereafter referred to as the *Three Epoch* model). The latter two models, as well as the two dadi_pipeline models which performed well, are shown schematically in Figure 1. To allow for a more realistic population growth trajectory, we also ran each of the dadi_pipeline models with a logistic population growth equation rather than the exponential growth variant defined in the original models. Each of the dadi_pipeline models and their logistic growth versions were run three times: once with growth allowed in the founding population post-split, once with growth allowed in the founded population post-split, and once with growth allowed in both populations post-split. Note that in each of these (dadi_pipeline) models, a source population splits to form two descendant populations, with an optimized parameter (s) controlling the portion of the population that forms each descendant population. When s is optimized to be very small (as it typically was), the founded population represents only a very small proportion of the ancestral population, as is probably realistic for the founding of the Hawaiian population from the North American population. Schematic depictions of the dadi_pipeline models are available in Portik et al. (2017) and Charles et al. (2018).
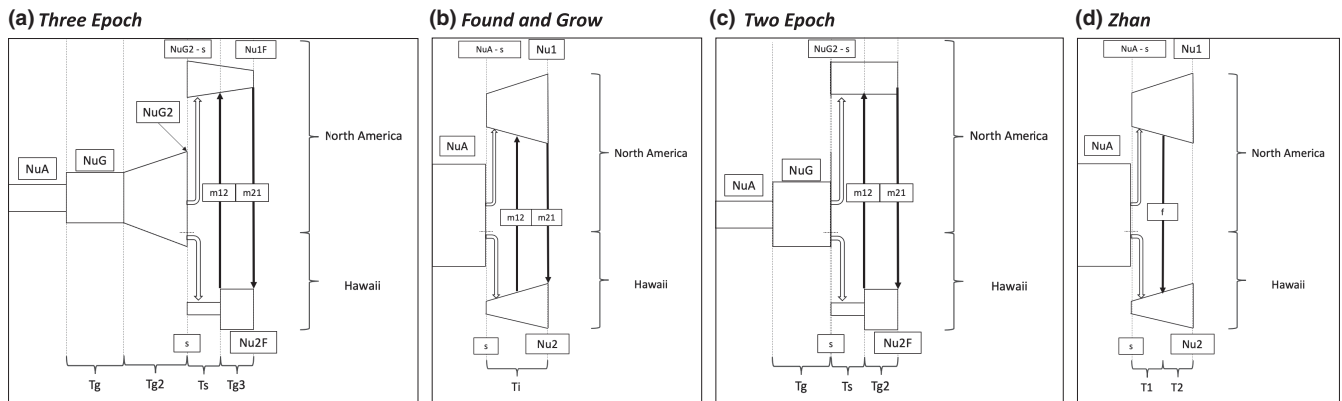
FIGURE 1 Visual depiction of the four best performing models from DADI simulations. (a) *Three epoch*, which allows for multiple changes in the size of the ancestral north American population prior to establishment in Hawaii, the north American and Hawaiian populations to change in size after the Hawaiian establishment, and a constant migration rate between north American and Hawaii after a brief time lag. (b) *Found and grow*, which assumes a constant ancestral north American population size. (c) *Two epoch*, which has a one-time admixture rather than constant migration. (d) the model from Zhan et al. (2014), which allows for only a single instance of growth in the north American population prior to the establishment of the Hawaiian population. In each model, parameters are estimated jointly and correspond to the following: Nu = population size at a particular time, T = time, m = migration rate between populations and s = founding population size. Note that occasions of population size change depicted here are allowed to freely optimize to be either population growths or declines.

TABLE 1 Number of samples remaining after filtering, Tajima's *D*, observed heterozygosity (H_O) and nucleotide diversity (π).

Population	No. of samples	Tajima's *D*	H_O	π
North America (NAM)	83	−1.92	0.055	0.064
Hawaii (HAW)	9	−0.211	0.048	0.055
Guam (GUA)	19	0.091	0.031	0.032
Rota (ROT)	16	0.388	0.035	0.038
Saipan (SAI)	4	0.326	0.022	0.025
Queensland (QLD)	15	0.349	0.042	0.044
New South Wales (NSW)	5	0.433	0.038	0.042
Victoria (VIC)	2	0.899	0.041	0.043

Note: Populations from Guam, Rota and Saipan are all part of the Mariana Islands archipelago. Samples from Maui and Oahu are pooled into a single Hawaiian population. Queensland, New South Wales and Victoria are all within the Australian continent. Populations from Samoa, New Caledonia, Fiji and New Zealand are not shown because no samples remained from these populations after filtering (but see Figure 2).

To optimize the models we fit during the analysis, we used a variation of the dadi_pipeline, the sequential step-down parameter permutation approach described by Portik et al. (2017). Unlike this method, however, we set the starting parameters for each sequential run via weighting the parameters from each run in the previous iteration by the relative Akaike information criterion (AIC) score of that iteration, such that all but the worst runs contribute in some degree to the starting parameters for the next step. The number of runs and iterations per step are listed in Table S3. Individual optimization runs were killed if they took longer than 48 h to complete; these runs tended to take far longer to finish and often included integration errors due to extremely small population sizes, resulting in extremely large amounts of genetic drift. Most runs completed in under 48 h and are included in the results.

2.9 | Parameter estimation

To extract meaningful parameter units from the results, we assumed 0.3 years per generation and used the per-base mutation rate of 8.4×10^{-9} reported from *Drosophila melanogaster* (Haag-Liautard et al., 2007). We use these values to match those used by Zhan et al. (2014) for ease of comparison. We also used a potentially more realistic generation time of seven generations per year and the slower mutation rate reported for the more closely related *Heliconius melpomene* of 2.9×10^{-9} (Keightley et al., 2015). To determine the length of the considered genomic region, we multiplied the total number of bases sequenced after quality filtering (but not SNP *p*-value filtering so as to count nonpolymorphic sites) by the ratio of SNPs in the final allele frequency spectrum to the total number of called SNPs.
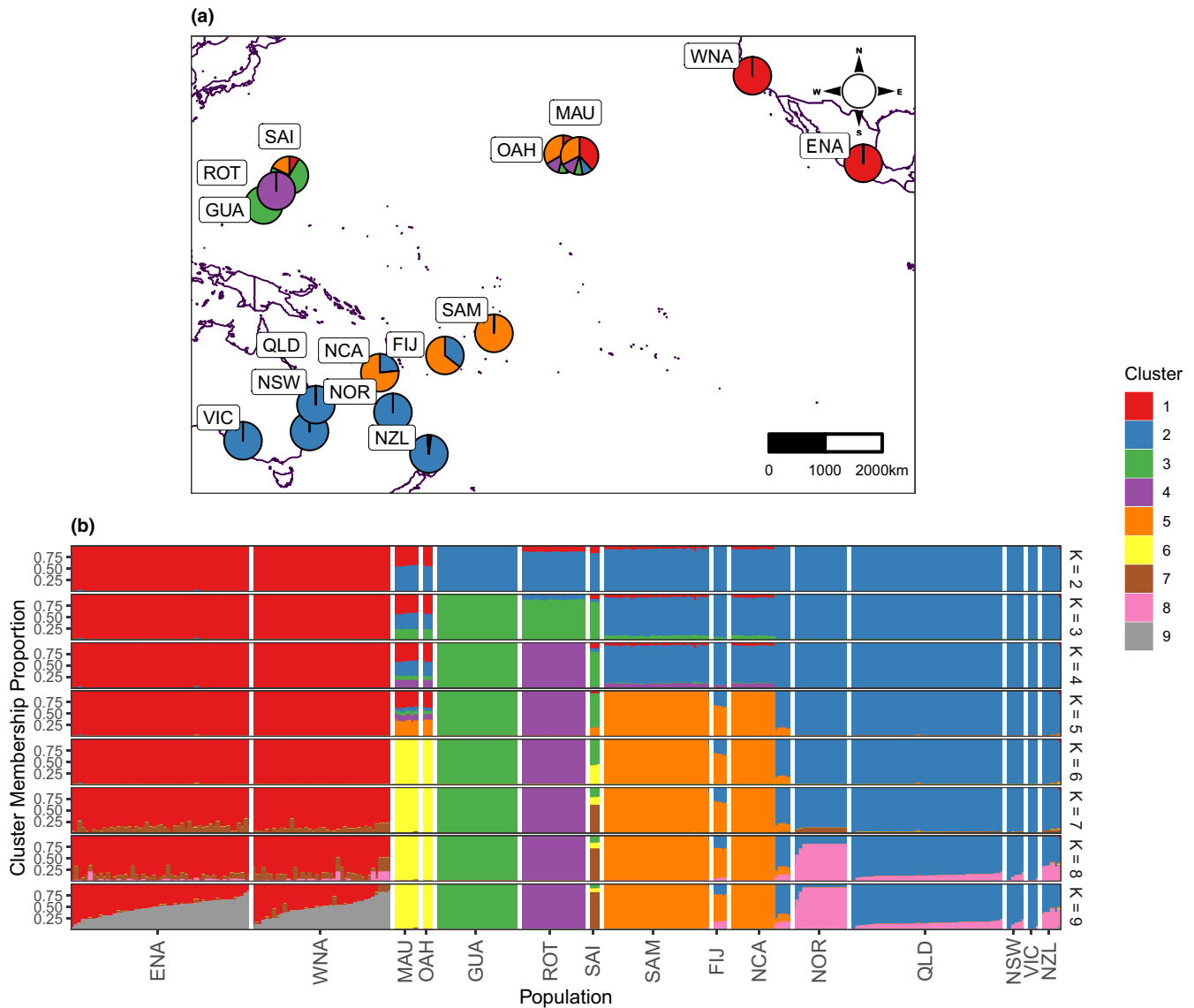
FIGURE 2 Relatedness among sampled populations. (a) Map of sampled populations, with pie charts reflecting average population results from NGSadmix ($k = 5$). Note that the geographical positions of the Mariana Islands (ROT, GUA and SAI) and the Hawaiian islands (OAH and MAU) are shifted slightly for readability. (b) NGSadmix plots showing the proportion of ancestry across clustering values between $k = 2$ and $k = 9$. At $k = 5$, Hawaii reflects a mixture of ancestry comprising north American, Mariana Islands and southwestern Pacific samples. At $k = 6$, Hawaii becomes its own cluster. At values beyond $k = 6$, populations are subdivided.

3 | RESULTS

3.1 | Sequencing results

After paralogue filtering, we were able to genotype 2,159,978 sites in at least 50% of individuals. In total, 541,899 of these sites were polymorphic, and 71,157 had a minor allele frequency above 0.05. For data set 1, we retained 11,384 loci after removing loci within 10 kb of each other. Data sets 2 and 3 had 70,878 and 413,271 sites, respectively. The number of samples from each population after filtering is shown in Table 1. Note that some populations had no samples that passed filtering, and so are not included in Table 1.

3.2 | Overall patterns of relatedness

PCA separated North American, Hawaiian, Mariana Islands and southwest Pacific samples along two axes of expansion (Figure S1B). NGSadmix results showed a similar result (Figure 2b), splitting the North American samples from the Pacific samples at $K = 2$, the Mariana Islands at $K = 3$, Rota from the other Mariana Islands at $K = 4$, Samoa, Fiji and New Caledonia at $K = 5$, Hawaii at $K = 6$, Saipan from Guam at $K = 7$, and Norfolk Island at $K = 8$. Notably, the Hawaiian samples were consistently classed as having ancestry partially from all other clusters until $K = 6$, consistent with an initial introduction into the archipelago. At $K = 9$, nearly all North American samples were assigned to two genetic clusters with ancestry proportions

unrelated to their geographical sampling locations, which can be interpreted as the presence of a fictive cluster with no biological significance (e.g., Chen et al., 2007; Guillot et al., 2005). $K = 2$ and $K = 5$ had the highest ΔK values (Evanno et al., 2005), although we were not able to estimate ΔK for $K = 3$ due to very low likelihood variance between runs at $K = 2$ and $K = 3$, thereby producing an undefined ΔK (Figure S2). Eastern and Western North America were never split. Directionality index (ψ) scores indicated westward establishments (Figure S4). Genetic diversity (π, H_O and Het/Hom) was highest in the ancestral North American populations, followed by Hawaii, Australia and then the remaining Pacific Island populations (Table 1; Figure S4). Tajima's D was positive in all sites except North America and Hawaii (Table 1). F_{ST} results also reflect the patterns we observed in the PCA and NGSadmix results (Figure 2b; Table S2, Figure S1A). These results varied slightly in the heavily filtered data set, but followed the same general trends (with a higher diversity in North America, Hawaii and Australia, and very similar Tajima's D and F_{ST} results, Tables S5 and S6). These results did have a higher relative diversity in the Pacific populations than in the full data set, however, and, as a natural consequence of removing loci with low minor allele frequencies, much higher overall diversity estimates.

3.3 | Patterns of differentiation within expansion populations

Samples from the Mariana Islands (especially the well-sampled Guam and Rota populations) appear to form highly distinct populations, despite their close physical proximity (Figure 2a,b; Figure S1A,B). By contrast, populations within Hawaii (Maui and Oahu) and Australia (Queensland, New South Wales and Victoria) do not show strong patterns of differentiation (Figure 2a,b). Norfolk Island, the other previously unsampled population in our data set, groups closely with samples from Australia and New Zealand (Figure 2a,b; Figure S1A,B). IBD patterns were significant within the Mariana Island ($p = .001$, $r = .723$), North American ($p = .001$, $r = .036$) and Hawaiian ($p = .005$, $r = .456$) samples, and were present but not significant within Australian samples ($p = .109$, $r = .122$). In the heavily filtered data set, IBD patterns were significant within all except Hawaii ($p = .001$, $r = .593$ in the Mariana Islands; $p = .002$, $r = .054$ in North America; $p = .609$, $r = .131$ in Hawaii; and $p = .014$, $r = .238$ in Australia).

3.4 | Timing of establishment and patterns of ongoing gene flow

Among the large set of possible demographic models, the simple *Found and Grow* scenario (Figure 1b), which had a constant ancestral population size in North America, Hawaii colonization and then population growth in both sites, produced the lowest AIC scores on the final pass of the pipeline (Figure S5). However, the *Two Epoch* model, which had a single admixture event but no consistent migration (Figure 1c), had the lowest AIC score across all passes of the

pipeline. The new *Three Epoch* model (Figure 1a), which involved multiple rounds of demographic expansion in the ancestral North American population, followed by colonization and growth in Hawaii, had a lower AIC score than *Found and Grow* across all passes and a lower AIC score than the *Two Epoch* model on the final pass (Figure S5). These were the top three models across all passes and in the final pass alone (Figure S5, Table S4). The *Three Epoch* model is a more complex version of the model specified in Zhan et al. (2014, hereafter *Zhan*). We therefore report only the results for the *Found and Grow*, *Three Epoch*, *Two Epoch* and *Zhan* models here.

These four models gave highly variable estimates of establishment timing and founding population size, with the *Three Epoch* model generally producing much broader estimates for these parameters. For example, while the *Three Epoch* model suggested establishment times that ranged between ~10^2 and 10^5 years ago, the other models suggested times between 10^4 and 10^5 years ago (Figure 3). Similarly, the latter models were more consistent in predicting a large founding population of 10^3–10^6 individuals, while the *Three Epoch* models suggested a broader founding population size of between 10 and 10^6 individuals (Figure 3). These models also differed in their estimates of the contemporary N_e for the Hawaiian population, with the *Found and Grow* and *Two Epoch* models suggesting a large N_e of around 10^6 and the *Three Epoch* and *Zhan* models generally producing estimates of Hawaiian N_e between 10^2 and 10^7 (Figure 3). For all of the models, the major discrepancy between the observed and simulated SFS tended to be that the models underestimated the number of rare derived alleles in the North American but not the Hawaiian populations (Figure 4; Figures S6–S9). This pattern may indicate that the models did not optimize for strong enough founder effects in Hawaii, which would have caused a more drastic loss of rare alleles.

For other parameters, the four models generated similar estimates. Each of these models suggest very low levels of contemporary migration between North America and Hawaii, with the *Found and Grow* and *Two Epoch* models converging near 0 for both directions and the *Three Epoch* model generally suggesting migration rates of $<5 \times 10^{-4}$ and $<2.5 \times 10^{-4}$ for individuals per generation from North America to Hawaii and from Hawaii to North America, respectively, and the *Zhan* model suggesting migration rates of $<2.5 \times 10^{-4}$ in either direction (Figure 3). Using a more accurate generation time and mutation rate than the values used by Zhan et al. (2014) produced a result with slightly more distant divergence times and larger effective sizes, but not to a substantial degree.

4 | DISCUSSION

Many geographical range expansions occur via serial stepwise dispersal, and we found strong evidence for this pattern in Pacific monarch butterflies, consistent with a previous study (Pierce, Zalucki, et al., 2014a). Monarchs in the Mariana Islands are the product of a distinct expansion event within the Pacific. Summary statistics support a scenario of directional dispersal from North America to Hawaii,
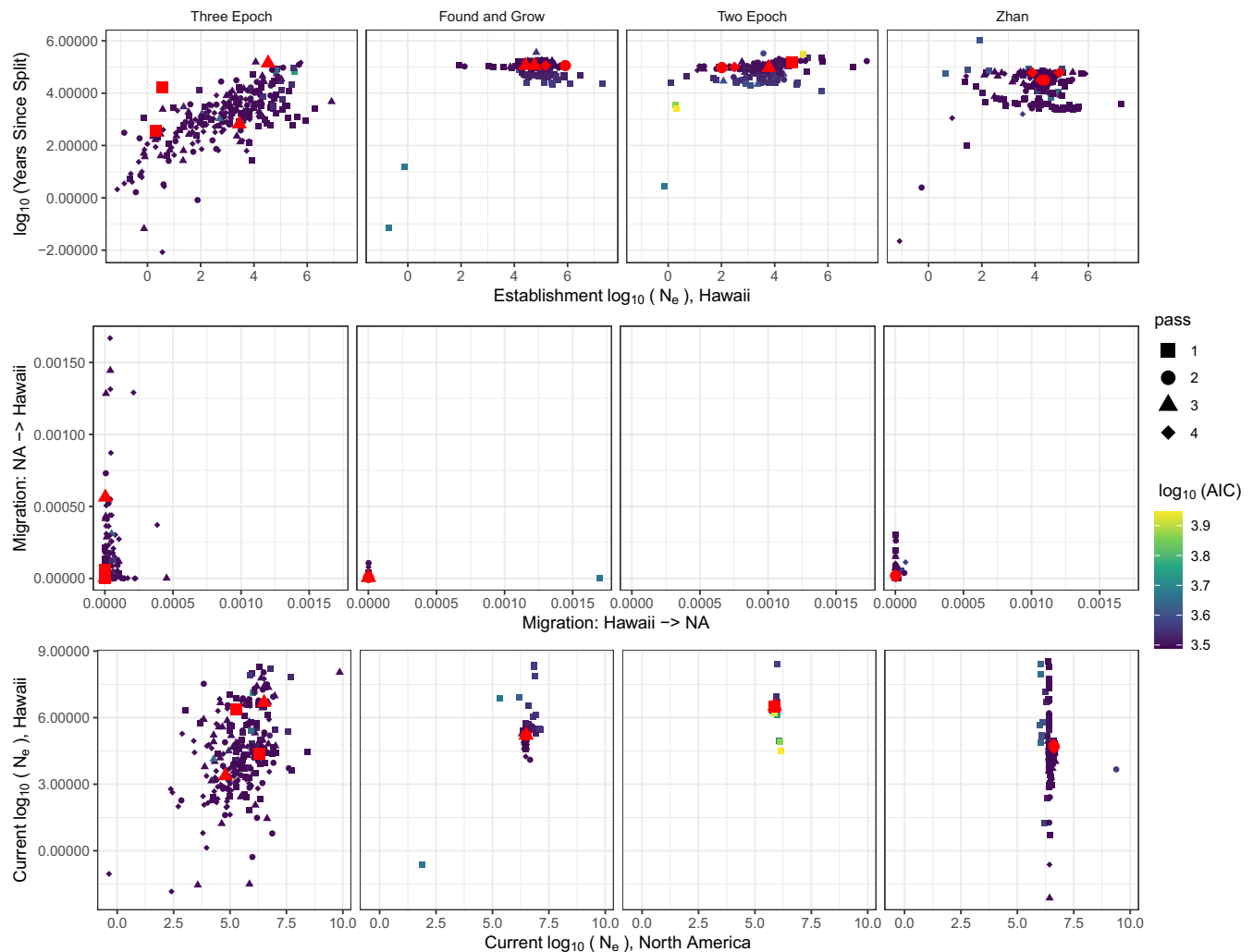
FIGURE 3 Results of the DADI optimization runs for the (left to right) *three epoch*, *found and grow*, *two epoch* and *Zhan* demographic models. In the key at right, "pass" refers to each iteration of the sequential step-down permutation approach (see methods: DADI model selection), and the $\log_{10}$(AIC) score corresponds to the likelihood of observing a particular set of parameter values. (top row) estimated effective size of the founding population in Hawaii vs. years since establishment. (middle row) estimated migration rates from North America to Hawaii and vice versa. (bottom row) estimated current effective population sizes in North America and Hawaii. Red dots mark the runs with the lowest AIC scores in each quadrant of the respective parameter space; these runs correspond to the heatmaps in Figure 4 and Figures S6 and S7 and the residuals in Figure S8. Note that the *two epoch* model does not include a constant migration rate between Hawaii and North America, and so is blank for the middle panel.

from Hawaii to Guam, and from Hawaii to Australia. This pattern is reflected in both the positive directionality index measures (0.07, 0.08 and 0.05, respectively) (Peter & Slatkin, 2013) and other summary statistics, such as the general increase in Tajima's *D* across the Pacific, which is consistent with stronger or more recent population bottlenecks during successive colonization. Interestingly, monarch populations in Hawaii and Australia seem to maintain relatively high levels of genetic diversity, despite the apparent bottlenecks associated with establishment. This is especially striking in the Australian population, which was itself probably founded by individuals from a much smaller population in New Caledonia (Clarke & Zalucki, 2004). The retention of genetic diversity in Hawaii and Australia may reflect rapid population growth upon establishment, which could temper the loss of allelic diversity that might be predicted with a bottleneck

event, akin to the scenario described in Hawaiian *Drosophila* by Nei et al. (1975). The slightly negative Tajima's *D* value in Hawaii is consistent with population growth following a bottleneck and is consistent with this hypothesis.

Within the Mariana Islands, there was a strong pattern of differentiation between islands, especially between the nearby islands of Guam and Rota. This pattern is striking because of their close geographical proximity: these islands are separated by only 40 km of open ocean. By contrast, our samples from North America, despite coming from overwintering sites nearly 2000 km apart, formed a single genetically indistinguishable population. This pattern is apparent from the strong pattern of IBD observed within Mariana Islands samples compared to weak/no IBD within North America. Our results are similar to those of Dapporto et al. (2017)
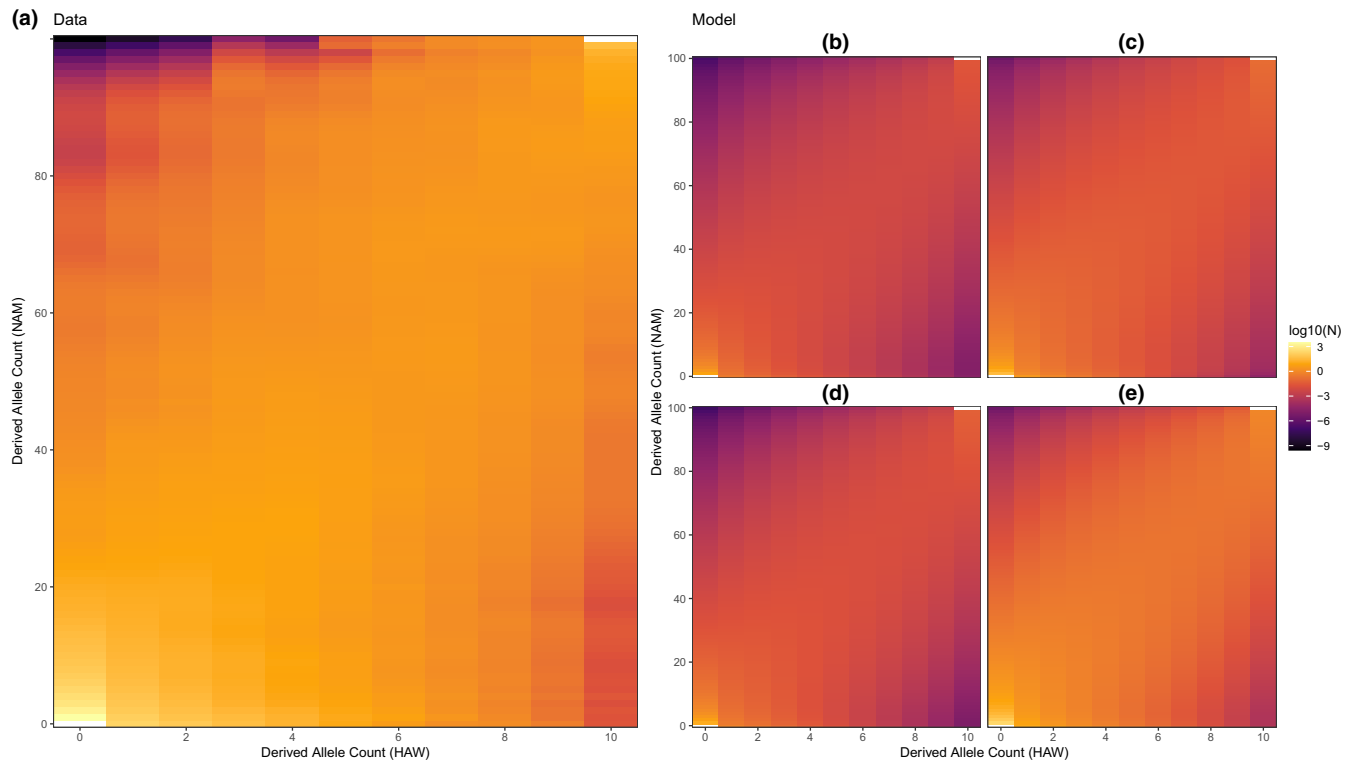
FIGURE 4 (a) observed data and (b-e) model-estimated derived site frequency spectra for the *three epoch* model. Cell brightness corresponds to the number of loci with derived allele frequencies in the given bin for both Hawaii (HAW) and North America (NAM). Note that spectra have been projected to a size of 100 for North America and 10 for Hawaii to match the approximate number of individuals sequenced from each population. Model-derived site frequency spectra correspond to the four points marked in red in Figure 3 for the *three epoch* model, which are parameter estimates from the runs with the lowest AIC score. Figures S6–S8 show comparable plots for the other candidate demographic models. BL, bottom left; BR, bottom right; TL, top left; TR, top right.

and Vodă et al. (2016), who also noted strong genetic differentiation between butterfly lineages even from nearby islands, as well as Alvial et al. (2018), who showed that a migratory dragonfly exhibits little genetic differentiation across its migratory Central and South American range but substantial genetic differentiation between nonmigratory populations on Pacific islands. However, our findings are unique because the observed differences in population structure between migratory and nonmigratory populations developed very recently, probably emerging over the past 150 years.

The lack of differentiation within North American monarchs corroborates other population genetic analyses of eastern and western North American monarchs (Brower & Boyce, 1991; Lyons et al., 2012; Shephard et al., 2002; Talla et al., 2020; Zhan et al., 2014) and is consistent with studies that have suggested movement of individuals between eastern and western North America (Billings, 2019; Dingle et al., 2005; Morris et al., 2015). The strong population genetic differentiation within the Mariana Islands but not at the scale of the entire North American continent highlights both (i) the pervasive role that long-distance migration in North America plays in collapsing any patterns of population structure that might otherwise develop, and (ii) the fact that many nonmigratory Pacific monarch populations probably have extremely small effective population sizes that are susceptible to very strong genetic drift.

In contrast to populations within the Mariana Islands, Hawaiian and Australian monarchs show only modest evidence for IBD that might be expected in nonmigratory monarch populations. Within Hawaii, our samples from Maui and Oahu formed a single genetic cluster, consistent with the results of Pierce, de Roode, et al. (2014b). Likewise, Australian samples from New South Wales and Victoria grouped with samples from Queensland. This result differs somewhat from the results of Hughes and Zalucki (1984), who reported considerable among-site genetic variation within Queensland, but it is consistent with similar later work (Zalucki et al., 1987).

For Hawaiian monarchs, it is not immediately clear why the islands of Maui and Oahu do not form clearly distinct populations. One possibility is that prevailing winds promote gene flow between islands in a way that differs from the Mariana Islands. Pacific monarchs are likely to be moved by wind patterns, similar to wind-driven movement patterns noted in migratory *Vanessa cardui* (Stefanescu et al., 2007), and it has been suggested that a tropical cyclone may have led to the monarch's establishment in Australia, following an "outbreak" of monarchs shortly after establishing in New Caledonia (Clarke & Zalucki, 2004). Another possibility is between-island movement of monarchs by butterfly breeders in Hawaii, who sell monarchs for release at weddings and celebrations (D. Loo-McDowell, pers. comm.). In the case of Australian monarchs, the lack of strong differentiation across the continent may be driven

by seasonal migration patterns akin to those seen in western North American monarchs (James, 1993; James & James, 2019). Australian monarchs retain migration-associated behaviours such as seasonal reproductive arrest and sustained directional flight—necessary although not sufficient conditions for long-distance migration—that further support the notion that they may undergo large-scale seasonal movements (Freedman et al., 2018; W. B. Hemstrom, M. P. Zalucki, M. G. Freedman, & M.R. Miller, in prep.; James, 1993). Thus, the lack of continent-wide population structure seen in migratory North American monarchs may be recapitulated, albeit to a lesser extent, in Australia.

Interpreting the results of our demographic models is somewhat more complicated than interpreting basic patterns of relatedness among populations. This is due to the conflicting inferences provided by the two best-performing model structures and the wide range of parameter estimates in the *Three Epoch* models. Although we present the results of both the simpler *Found and Grow* and *Two Epoch* models, and the more complicated *Three Epoch* model, we are inclined to place more confidence in the estimates produced by the *Three Epoch* model for two reasons: (i) the demographic scenario that it specifies—recent demographic expansion in the ancestral North American population prior to geographical expansion—has empirical support from other studies (Boyle et al., 2022; Pfeiler et al., 2017; Zhan et al., 2014) and accords with our understanding of past changes in climate, and (ii) this model structure produces parameter estimates that match our prior understanding for how and when monarch range expansion may have occurred. The latter point is related to the former: because a North American population expansion is not allowed until after the founding of Hawaii in the *Found and Grow* model, this model forces an ancient founding of the Hawaiian population in order to allow for the ancient growth of the North American population. As such, we focus our discussion on the estimates produced by the *Three Epoch* model.

In general, our demographic results do not exclude a recent founding of the Hawaiian population by North American monarchs (Figure 3d). While our model optimizations span several orders of magnitude for the time since establishment, many of the iterations settled on introduction estimates of less than 200 years ago for the *Three Epoch* model. Since the earliest historical records of monarchs on Hawaii date to roughly 200 years ago (1841) (Zalucki & Clarke, 2004), we are inclined to accept the results of iterations with shorter estimated divergence times. Other lines of evidence supporting recent (<200 years) Hawaiian establishment include: (i) the lack of noticeable phenotypic differentiation between North American and Pacific Island monarchs, especially relative to the pronounced phenotypic differences in nonmigratory populations from the Caribbean and South America (Freedman et al., 2020), which have historically been treated as separate subspecies (Ackery & Vane-Wright, 1984); (ii) the probable need for human-mediated transport of the monarch's host plants (some of which are native to subtropical Africa) as a precondition of monarch establishment in the Pacific; and (iii) recent genomic evidence showing that captive breeding of monarchs over short timescales is sufficient to generate

patterns of genetic divergence comparable to those observed between North American and Pacific populations (Tenger-Trolander et al., 2019). Notably, our re-implementation of the model used by Zhan et al. (2014) produced results similar to theirs, with the majority of model iterations supporting an introduction time >1000 years ago (Figure 3). This highlights the need to run a range of possible demographic models when attempting to infer demographic history, since failing to account for underlying complexity in population histories can result in very divergent parameter estimates.

One complication for interpreting our demographic models is that they consistently underestimated the number of rare, derived alleles present in North America but not Hawaii. During very strong bottlenecks, we would expect many rare alleles to be lost, suggesting that our models may be overestimating the founding population size, and thus probable establishment date. Since DADI can struggle to calculate SFS when population sizes are very small due to large amounts of drift, iterations that optimize to this segment of parameter space are more likely to have integration errors or very long processing times. These uncompleted runs were not included in our model results, and so this part of the parameter space may be inadequately explored. Since small founding population sizes correlated with recent introductions across our model results, this also suggests that it would be unwise to rule out a recent introduction with a very strong bottleneck based on a demographic analysis alone.

Demographic model results were also variable in their estimates of founding population sizes in Hawaii. Some models produced estimates as high as 10,000 founding individuals, which seems implausible given the long distance (>3500 km) between North America and Hawaii. The extremely wide range of parameter estimates for founding population size and timing may reflect that, in practice, it is difficult to distinguish between a very recent strong bottleneck vs. a more distant but less severe bottleneck. Our model results are consistent with this, since the model iterations with recent establishments also tended to have smaller establishment population sizes (Figure 3).

In contrast to variable estimates of establishment timing and founding population size, demographic models were consistent in suggesting very low contemporary migration rates (on the order of 0.0001 individuals per generation from North America to Hawaii and vice versa). Our results thus contrast with those of Pierce, Zalucki et al. (2014a), who inferred much higher migration rates (nearly 10 individuals/generations) between North America and Hawaii. We are more confident in our results due to (i) the much larger number of sampled loci, (ii) the more realistic demographic model that we used in our analysis and (iii) the absence of modern records of regular North America to Hawaii establishment events. However, we do note that monarchs appear to be capable of wind-aided dispersal over extremely long distances, as evidenced by occasional records from the UK that coincide with records of migratory North American birds blown by storms (Asher et al., 2001); thus, ongoing movement from North America to Hawaii is not entirely implausible.

Understanding how migratory and nonmigratory populations of monarchs differ genetically, phenotypically and ecologically has

important conservation implications. First, monarchs were recently evaluated by the U.S. Fish and Wildlife Service, who determined that a listing under the U.S. Endangered Species Act is "warranted but precluded" (USFWS, 2021). This decision will be re-evaluated annually, and the retention of genetic diversity in nonmigratory monarch populations from outlying US states/territories (Hawaii, American Samoa, the Mariana Islands, Puerto Rico, US Virgin Islands) may be important in decisions regarding the adaptive capacity of the species (Freedman et al., 2021). Second, some recent evidence suggests that climate warming and planting of nonnative milkweed species might tip the scales in favour of year-round breeding and loss of migratory behaviour in North American monarchs, particularly in western North America (Crone & Schultz, 2021; James, 2021; James et al., 2021). The increased prevalence of partial migration within North America, both along the US Gulf Coast and in California (James et al., 2021; Satterfield et al., 2016, 2018), may affect patterns of spatial genetic diversity: for example, if future sequencing of North American monarchs finds evidence for population structure within areas of their range where year-round breeding occurs, this would provide evidence that loss of migration is actively driving genetic differentiation. Finally, our results are also helpful for monarch conservation because they provide a relatively large sample of North American monarchs ($n = 90$) against which future sequencing efforts can be compared to look for evidence of contemporary losses of genetic diversity associated with potential population decline.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflicts of interest related to this work.

FUNDING INFORMATION

OPEN RESEARCH BADGES

DATA AVAILABILITY STATEMENT

ORCID

William B. Hemstrom https://orcid.org/0000-0002-2408-9535
Micah G. Freedman https://orcid.org/0000-0003-1739-1864

REFERENCES

Ackery, P. R., & Vane-Wright, R. I. (1984). *Milkweed butterflies: Their cladistics and biology.* Cornell University Press.

Adriaensen, F., & Dhondt, A. A. (1990). Population dynamics and partial migration of the European robin (*Erithacus rubecula*) in different habitats. *The Journal of Animal Ecology, 59*, 1077–1090.

Ali, O. A., O'Rourke, S. M., Amish, S. J., Meek, M. H., Luikart, G., Jeffres, C., & Miller, M. R. (2016). RAD capture (rapture): Flexible and efficient sequence-based genotyping. *Genetics, 202*, 389–400.

Altizer, S., & Davis, A. K. (2010). Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. *Evolution, 64*, 1018–1028.

Alvial, I. E., Vargas, H. A., Marinov, M., Esquivel, C., Araya, J., Araya-Donoso, R., Vila, I., & Véliz, D. (2018). Isolation on a remote Island: Genetic and morphological differentiation of a cosmopolitan odonate. *Heredity, 122*, 893–905.

Asher, J., Warren, M., Fox, R., Harding, P., Jeffcoate, G., & Jeffcoate, S. (Eds.). (2001). *The millennium atlas of butterflies in Britain and Ireland* (pp. 319–323). Oxford University Press.

Berg, J. E., Hebblewhite, M., St. Clair, C. C., & Merrill, E. H. (2019). Prevalence and mechanisms of partial migration in ungulates. *Frontiers in Ecology and Evolution., 7*, 325.

Bettin, O., Cornejo, C., Edwards, P. J., & Holderegger, R. (2007). Phylogeography of the high alpine plant *Senecio halleri* (Asteraceae) in the European Alps: In situ glacial survival with postglacial stepwise dispersal into peripheral areas. *Molecular Ecology, 16*, 2517–2524.

Billings, J. (2019). Opening a window on southwestern monarchs: Fall migrant monarch butterflies, *Danaus plexippus* (L.), tagged synchronously in southeastern Arizona migrate to overwintering regions in either southern California or Central Mexico. *Journal of the Lepidopterists' Society, 73*, 257–267.

Boyle, J. H., Strickler, S., Twyford, A., Ricono, A., Powell, A., Zhang, J., Xu, H., Dalgeish, H. J., Jander, G., Agrawal, A. A., & Puzey, J. R. (2022). Temporal matches and mismatches between monarch butterfly and milkweed population changes over the past 12,000 years. *biorXiv.* [PREPRINT]. https://doi.org/10.1101/2022.02.25.481796

Brower, A. V. Z., & Boyce, T. M. (1991). Mitochondrial DNA variation in monarch butterflies. *Evolution, 45*, 1281–1286.

Chapman, B. B., Brönmark, C., Nilsson, J.-Å., & Hansson, L.-A. (2011). The ecology and evolution of partial migration. *Oikos, 120*, 1764–1775.

Charles, K. L., Bell, R. C., Blackburn, D. C., Burger, M., Fujita, M. K., Gvoždík, V., Jongsma, G. F. M., Kouete, M. T., Leaché, A. D., & Portik,

D. M. (2018). Sky, sea, and forest islands: Diversification in the African leaf-folding frog *Afrixalus paradorsalis* (Anura: Hyperoliidae) of the lower Guineo-Congolian rain forest. *Journal of Biogeography*, *45*, 1781–1794.

Chen, C., Durand, E., Forbes, F., & François, O. (2007). Bayesian clustering algorithms ascertaining spatial population structure: A new computer program and a comparison study. *Molecular Ecology Notes*, *7*, 747–756.

Clarke, A. R., & Zalucki, M. P. (2004). Monarchs in Australia: On the winds of a storm? *Biological Invasions*, *6*, 123–127.

Cohen, E. B., Hostetler, J. A., Hallworth, M. T., Rushing, C. S., Sillett, T. S., & Marra, P. P. (2018). Quantifying the strength of migratory connectivity. *Methods in Ecology and Evolution*, *9*, 513–524.

Crone, E. E., & Schultz, C. B. (2021). Resilience or catastrophe? A possible state change for monarch butterflies in western North America. *Ecology Letters*, *24*, 1533–1538.

Dapporto, L., Cini, A., Menchetti, M., Vodă, R., Bonelli, S., Casacci, L. P., Dincă, V., Scalercio, S., Hinojosa, J. C., Biermann, H., Forbicioni, L., Mazzantini, U., Venturi, L., Zanichelli, F., Balletto, E., Shreeve, T. G., Dennis, R. L. H., & Vila, R. (2017). Rise and fall of Island butterfly diversity: Understanding genetic differentiation and extinction in a highly diverse archipelago. *Diversity and Distributions*, *23*, 1169–1181.

Dapporto, L., Cini, A., Vodă, R., Dincă, V., Wiemers, M., Menchetti, M., Magini, G., Talavera, G., Shreeve, T., Bonelli, S., Casacci, L. P., Balletto, E., Scalercio, S., & Vila, R. (2019). Integrating three comprehensive data sets shows that mitochondrial DNA variation is linked to species traits and paleogeographical events in European butterflies. *Molecular Ecology Resources*, *19*, 1623–1636.

Dingle, H. (2014). *Migration: The biology of life on the move*. Oxford University Press.

Dingle, H., Blakley, N. R., & Miller, E. R. (1980). Variation in body size and flight performance in milkweed bugs (*Oncopeltus*). *Evolution*, *34*, 371–385.

Dingle, H., Zalucki, M. P., Rochester, W. A., & Armijo-Prewitt, T. (2005). Distribution of the monarch butterfly, *Danaus plexippus* (L.) (lepidoptera: Nymphalidae), in western North America. *Biological Journal of the Linnean Society. Linnean Society of London*, *85*, 491–500.

Edwards, A. W. (1971). Distances between populations on the basis of gene frequencies. *Biometrics*, *27*, 873–881.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 481–501.

Fernández-Haeger, J., Jordano, D., & Zalucki, M. P. (2015). Monarchs across the Atlantic Ocean. In K. S. Oberhauser, K. R. Nail, & S. Altizer (Eds.), *Monarchs in a changing world: Biology and conservation of an iconic butterfly* (pp. 247–256). Cornell University Press.

Francis, R. M. (2017). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*, 27–32.

Freedman, M. G., de Roode, J. C., Forister, M. L., Kronforst, M. R., Pierce, A. A., Schultz, C. B., Taylor, O. R., & Crone, E. E. (2021). Are eastern and western monarch butterflies distinct populations? A review of evidence for ecological, phenotypic, and genetic differentiation and implications for conservation. *Conservation Science and Practice*, *3*(7), e432.

Freedman, M. G., Dingle, H., Strauss, S. Y., & Ramírez, S. R. (2020). 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*, *117*, 28887–28893.

Freedman, M. G., Dingle, H., Tabuloc, C. A., Chiu, J. C., Yang, L. H., & Zalucki, M. P. (2018). Non-migratory monarch butterflies, *Danaus plexippus* (L.), retain developmental plasticity and a navigational mechanism associated with migration. *Biological Journal of the Linnean Society. Linnean Society of London*, *123*, 265–278.

Gao, B., Hedlund, J., Reynolds, D. R., Zhai, B., Hu, G., & Chapman, J. W. (2020). The "migratory connectivity" concept, and its applicability to insect migrants. *Movement Ecology*, *8*, 48.

Gilbert, K. J., Andrew, R. L., Bock, D. G., Franklin, M. T., Kane, N. C., Moore, J.-S., & Vines, T. H. (2012). Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology*, *21*(20), 4925–4930.

Gómez-Bahamón, V., Márquez, R., Jahn, A. E., Miyaki, C. Y., Tuero, D. T., Laverde-R, O., Restrepo, S., & Cadena, C. D. (2020). Speciation associated with shifts in migratory behavior in an avian radiation. *Current Biology*, *30*, 1312–1321.e6.

Gong, X., Davenport, E. R., Wang, D., & Clark, A. G. (2019). Lack of spatial and temporal genetic structure of Japanese eel (*Anguilla japonica*) populations. *Conservation Genetics*, *20*, 467–475.

Guillot, G., Estoup, A., Mortier, F., & Cosson, J. F. (2005). A spatial statistical model for landscape genetics. *Genetics*, *170*, 1261–1280.

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*, e1000695.

Haag-Liautard, C., Dorris, M., Maside, X., Macaskill, S., Halligan, D. L., Houle, D., Charlesworth, B., & Keightley, P. D. (2007). Direct estimation of per nucleotide and genomic deleterious mutation rates in drosophila. *Nature*, *445*, 82–85.

Hemstrom, W., & Jones, M. (2021). *snpR: User friendly population genomics for SNP datasets with categorical metadata*. Authorea. https://doi.org/10.22541/au.161264719.94032617/v1

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society. Linnean Society of London*, *58*, 247–276.

Hughes, J. M., & Zalucki, M. P. (1984). Genetic variation in a continuously breeding population of *Danaus plexippus* L. (lepidoptera: Nymphalidae). *Heredity*, *52*, 1–7.

Ibrahim, K. M., Nichols, R. A., & Hewitt, G. M. (1996). Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, *77*, 282–291.

Irwin, D. E., Irwin, J. H., & Smith, T. B. (2011). Genetic variation and seasonal migratory connectivity in Wilson's warblers (*Wilsonia pusilla*): Species-level differences in nuclear DNA between western and eastern populations. *Molecular Ecology*, *20*, 3102–3115.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806.

James, D. G. (1993). Migration biology of the monarch butterfly in Australia. In S. B. Malcolm & M. P. Zalucki (Eds.), *Biology and conservation of the monarch butterfly* (pp. 189–200). Los Angeles County Museum of Natural History.

James, D. G. (2021). Western north American monarchs: Spiraling into oblivion or adapting to a changing environment? *Animal Migration*, *8*, 19–26.

James, D. G., & James, T. A. (2019). 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*, *73*, 177–190.

James, D. G., Schaefer, M. C., Krimmer, E. K., & Carl, A. (2021). First population study on winter breeding monarch butterflies, *Danaus plexippus* (lepidoptera: Nymphalidae) in the urban South Bay of San Francisco, California. *Insects*, *12*, 946.

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*(14), 3594–3602.

Keightley, P. D., Pinharanda, A., Ness, R. W., Simpson, F., Dasmahapatra, K. K., Mallet, J., Davey, J. W., & Jiggins, C. D. (2015). Estimation of the spontaneous mutation rate in Heliconius melpomene. *Molecular Biology and Evolution*, *32*, 239–243.

Kondo, B., Peters, J. L., Rosensteel, B. B., & Omland, K. E. (2008). Coalescent analyses of multiple loci support a new route to speciation in birds. *Evolution*, *62*, 1182–1191.

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*, 356.

Kuchta, S. R., & Tan, A.-M. (2005). Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. *Molecular Ecology*, *14*, 225–244.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Linderoth, T. P (2018). *Identifying population histories, adaptive genes, and genetic duplication from population-scale next generation sequencing*. Ph.D. thesis. https://escholarship.org/uc/item/5kp4q40k

Lyons, J. I., Pierce, A. A., Barribeau, S. M., Sternberg, E. D., Mongue, A. J., & De Roode, J. C. (2012). Lack of genetic differentiation between monarch butterflies with divergent migration destinations. *Molecular Ecology*, *21*, 3433–3444.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.

Menz, M. H. M., Reynolds, D. R., Gao, B., Hu, G., Chapman, J. W., & Wotton, K. R. (2019). Mechanisms and consequences of partial migration in insects. *Frontiers in Ecology and Evolution*, *7*, 403.

Morris, G. M., Kline, C., & Morris, S. M. (2015). Status of *Danaus plexippus* population in Arizona. *Journal of the Lepidopterists' Society*, *69*, 91–107.

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, *29*, 1–10.

Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528.

Peter, B. M., & Slatkin, M. (2013). Detecting range expansions from genetic data. *Evolution, 67*, 3274–3289.

Pfeiler, E., Nazario-Yepiz, N. O., Pérez-Gálvez, F., Chávez-Mora, C. A., Laclette, M. R. L., Rendón-Salinas, E., & Markow, T. A. (2017). Population genetics of overwintering monarch butterflies, *Danaus plexippus* (Linnaeus), from Central Mexico inferred from mitochondrial DNA and microsatellite markers. *The Journal of Heredity*, *108*, 163–175.

Pierce, A. A., Altizer, S., Chamberlain, N. L., Kronforst, M. R., & de Roode, J. C. (2015). Unraveling the mysteries of monarch migration and global dispersal through molecular genetic techniques. In K. S. Oberhauser, K. R. Nail, & S. Altizer (Eds.), *Monarchs in a changing world: Biology and conservation of an iconic butterfly* (pp. 257–267). Cornell University Press.

Pierce, A. A., de Roode, J. C., Altizer, S., & Bartel, R. A. (2014b). Extreme heterogeneity in parasitism despite low population genetic structure among monarch butterflies inhabiting the Hawaiian islands. *PLoS One*, *9*, e100061.

Pierce, A. A., Zalucki, M. P., Bangura, M., Udawatta, M., Kronforst, M. R., Altizer, S., Haeger, J. F., & de Roode, J. C. (2014a). Serial founder effects and genetic differentiation during worldwide range expansion of monarch butterflies. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20142230.

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., Rödel, M.-O., Blackburn, D. C., & Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*, 5245–5263.

Prince, D. J., O'Rourke, S. M., Thompson, T. Q., Ali, O. A., Lyman, H. S., Saglam, I. K., Hotaling, T. J., Spidle, A. P., & Miller, M. R. (2017). The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. *Science Advances*, *3*, e1603198.

Rousset, F. (2008). genepop'007: A complete re-implementation of the genepop software for windows and Linux. *Molecular Ecology Resources*, *8*, 103–106.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*, 406–425.

Samarasin, P., Shuter, B. J., & Rodd, F. H. (2017). After 100 years: Hydroelectric dam-induced life-history divergence and population genetic changes in sockeye salmon (*Oncorhynchus nerka*). *Conservation Genetics*, *18*(6), 1449–1462.

Satterfield, D. A., Maerz, J. C., Hunter, M. D., Flockhart, D. T. T., Hobson, K. A., Norris, D. R., Streit, H., de Roode, J. C., & Altizer, S. (2018). Migratory monarchs that encounter resident monarchs show life-history differences and higher rates of parasite infection. *Ecology Letters*, *21*, 1670–1680.

Satterfield, D. A., Villablanca, F. X., Maerz, J. C., & Altizer, S. (2016). Migratory monarchs wintering in California experience low infection risk compared to monarchs breeding year-round on non-native milkweed. *Integrative and Comparative Biology*, *56*, 343–352.

Shephard, J. M., Hughes, J. M., & Zalucki, M. P. (2002). Genetic differentiation between Australian and north American populations of the monarch butterfly *Danaus plexippus* (L.) (lepidoptera: Nymphalidae): An exploration using allozyme electrophoresis. *Biological Journal of the Linnean Society. Linnean Society of London*, *75*, 437–452.

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*, 693–702.

Slatkin, M., & Excoffier, L. (2012). Serial founder effects during range expansion: A spatial analog of genetic drift. *Genetics*, *191*, 171–181.

Stefanescu, C., Alarcón, M., & Avila, A. (2007). Migration of the painted lady butterfly, *Vanessa cardui*, to North-Eastern Spain is aided by African wind currents. *The Journal of Animal Ecology*, *76*, 888–898.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595.

Talla, V., Pierce, A. A., Adams, K. L., de Man, T. J. B., Nallu, S., Villablanca, F. X., Kronforst, M. R., & de Roode, J. C. (2020). Genomic evidence for gene flow between monarchs with divergent migratory phenotypes and flight performance. *Molecular Ecology*, *29*, 2567–2582.

Tenger-Trolander, A., Lu, W., Noyes, M., & Kronforst, M. R. (2019). Contemporary loss of migration in monarch butterflies. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 14671–14676.

Tiedemann, R., Paulus, K. B., Scheer, M., Von Kistowski, K. G., Skírnisson, K., Bloch, D., & Dam, M. (2004). Mitochondrial DNA and microsatellite variation in the eider duck (*Somateria mollissima*) indicate stepwise postglacial colonization of Europe and limited current long-distance dispersal. *Molecular Ecology*, *13*, 1481–1494.

USFWS (United States fish and wildlife service) (2021). *Assessing the status of the monarch butterfly*. https://www.fws.gov/savethemonarch/ssa.html

Vane-Wright, R. I. (1993). The Columbus hypothesis: An explanation for the dramatic 19th century range expansion of the monarch butterfly. In S. B. Malcolm & M. P. Zalucki (Eds.), *Biology and conservation of the monarch butterfly* (pp. 179–188). Los Angeles County Museum of Natural History.

Vodă, R., Dapporto, L., Dincă, V., Shreeve, T. G., Khaldi, M., Barech, G., ... Vila, R. (2016). Historical and contemporary factors generate unique butterfly communities on islands. *Scientific Reports*, 6, 28828.

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15, 1419–1439.

Waples, R. S., Teel, D. J., Myers, J. M., & Marshall, A. R. (2004). Life-history divergence in Chinook salmon: Historic contingency and parallel evolution. *Evolution*, 58, 386–403.

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005). A note on exact tests of hardy-Weinberg equilibrium. *American Journal of Human Genetics*, 76, 887–893.

Wise, K. A. J. (1980). Monarch butterfly dispersal in New Zealand. *Records of the Auckland Institute and Museum*, 17, 157–173.

Zalucki, M. P., & Clarke, A. R. (2004). Monarchs across the Pacific: The Columbus hypothesis revisited. *Biological Journal of the Linnean Society. Linnean Society of London*, 82, 111–121.

Zalucki, M. P., Hughes, J. M., & Carter, P. A. (1987). Genetic variation in *Danaus plexippus* L.: Habitat selection or differences in activity times? *Heredity*, 59, 213–221.

Zhan, S., & Reppert, S. M. (2013). MonarchBase: The monarch butterfly genome database. *Nucleic Acids Research*, 41, D758–D763.

Zhan, S., Zhang, W., Niitepõld, K., Hsu, J., Haeger, J. F., Zalucki, M. P., Altizer, S., de Roode, J. C., Reppert, S. M., & Kronforst, M. R. (2014). The genetics of monarch butterfly migration and warning colouration. *Nature*, 514, 317–321.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.