Genomic evidence for gene flow between monarchs with divergent migratory phenotypes and flight performance

Running title: gene flow in migratory monarch butterflies

Venkat Talla[a], Amanda A. Pierce[a], Kandis L. Adams[a], Tom J.B. de Man[a], Sumitha Nallu[c], Francis X. Villablanca[d], Marcus R. Kronforst[c] & Jacobus C. de Roode[a*]

[a] Department of Biology, Emory University, Atlanta, GA 30322, USA

[c] Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA
[d] Biological Sciences Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA

[*]Corresponding Author:
Jacobus de Roode
tel. 404-727-2340
email: jderood@emory.edu

Key words: monarch butterfly, migration, gene flow, conservation biology, genomics/proteomics, population genetics-empirical

Abstract

Monarch butterflies are known for their spectacular annual migration in eastern North America, with millions of monarchs flying up to 4,500 kilometers to overwintering sites in central Mexico. Monarchs also live west of the Rocky Mountains, where they travel shorter distances to overwinter along the Pacific Coast. It is often assumed that eastern and western monarchs form distinct evolutionary units, but genomic studies to support this notion are lacking. We used a tethered flight mill to show that migratory eastern monarchs have greater flight performance than western monarchs, consistent with their greater migratory distances. However, analyzing more than 20 million SNPs in 43 monarch genomes, we found no evidence for genomic differentiation between eastern and western monarchs. Genomic analysis also showed identical and low levels of genetic diversity, and demographic analyses indicated similar effective population sizes and ongoing gene flow between eastern and western monarchs. Gene expression analysis of a subset of candidate genes during active flight revealed differential gene expression related to non-muscular motor activity. Our results demonstrate that eastern and western monarchs maintain migratory differences despite ongoing gene flow, and suggest that migratory differences between eastern and western monarchs are not driven by select major-effects alleles. Instead, variation in migratory distance and destination may be driven by environmentally induced differential gene expression, or by many alleles of small effect.

Introduction

Seasonal migration is common in nature (Dingle, 2014) and allows many different animals to escape deteriorating habitats, escape predators and parasites, and benefit from seasonally available resources in multiple regions (Dingle, 1972; Alerstam, Hedenstrom & Åkesson, 2003; Alerstam, 2006; McKinnon *et al.*, 2010; Altizer, Bartel & Han, 2011; Fricke, Hencecroth & Hoerner, 2011; Dingle, 2014). Migration is likely to be polygenic (Dingle, 1991) and studies have demonstrated that genes involved in muscle development, energy metabolism and circadian rhythm tend to show genetic divergence or differential expression patterns between migratory and non-migratory individuals (McFarlan, Bonen & Guglielmo, 2009; O'Malley, Ford & Hard, 2010; Postel, Thompson, Barker, Viney & Morris, 2010; Trivedi, Kumar, Rani & Kumar, 2014). While it is clear that migration imposes selection for specific gene variants or transcription levels, the interplay between animal migration and genome evolution remain understudied. Genomes may be affected by migration in varying ways. Populations of the same species often vary in their

63 migratory propensity, with some populations migrating and others not, or with populations
64 migrating over different distances and to different destinations. This could result in spatial or
65 temporal separation between different migrants, and consequently reduced gene flow and
66 increased genome-wide genetic differentiation, as found in beluga whales and noctule bats
67 (O'Corry-Crowe, Suydam, Rosenberg, Frost & Dizon, 1997; Petit & Mayer, 2000). Alternatively,
68 the use of common breeding or overwintering grounds can result in a lack of genome-wide
69 differentiation, even if differential selection acts on individuals during part of the year (Dallimer &
70 Jones, 2002; Dallimer, Jones, Pemberton & Cheke, 2003). An extreme example occurs in Pacific
71 salmon, in which the genetic differentiation between early (premature) and late (typical, mature)
72 migrants is restricted to a single gene, *GREB1L* (Prince *et al.*, 2017); while selection acts on this
73 gene seasonally, large amounts of gene flow homogenize the remainder of the genome. Insights
74 into the genetic basis of animal migration thus require genome-wide studies, to identify genes that
75 are under selection against a potential background of variable gene flow (Bensch, Andersson &
76 Åkesson, 1999; Liedvogel, Åkesson & Bensch, 2011).
77
78 Eastern North American monarch butterflies undergo one of the most well-known and spectacular
79 migrations of the animal kingdom (Gustafsson, Agrawal, Lewenstein & Wolf, 2015; Reppert & de
80 Roode, 2018), with up to hundreds of millions of butterflies migrating up to 4,500 km to reach their
81 overwintering sites in central Mexico (Urquhart & Urquhart, 1978; Brower, 1995; Flockhart *et al.*,
82 2017). Monarch caterpillars are specialist feeders of milkweed host plants, which die back
83 seasonally in North America, thereby preventing monarchs from breeding throughout the year. In
84 the fall, developing monarch caterpillars respond to changing temperature, shortening day length
85 and senescing host plants to enter a state of reproductive diapause (Goehring & Oberhauser,
86 2002), which enables them to survive the 6-8 months that it takes to migrate south, overwinter,
87 and re-migrate north in the spring (Herman & Tatar, 2001). Prior to spring re-migration,
88 overwintering monarchs complete reproductive development and mate at the Mexican
89 overwintering sites or in their recolonized breeding areas (Herman, Brower & Calvert, 1989).
90 Monarchs recolonize the southern parts of the United States and lay eggs on re-emerging
91 milkweed, and 2-4 successive generations of reproductive monarchs recolonize their entire 4.5
92 million km^2 breeding range (Flockhart *et al.*, 2013).
93
94 While monarchs are best known for this long-distance migration from eastern North America to
95 Mexico, monarchs that inhabit breeding grounds west of the Rocky Mountains migrate shorter
96 distances to overwinter in groves of Eucalyptus and native conifers along California's Pacific

Coast (Nagano *et al.*, 1993; James *et al.*, 2018). Whereas eastern monarchs may fly over 4,500km to reach the Mexican overwintering sites, western monarchs reach the California Coast by flying as little as 500km, with the greatest recorded distances being 1,600km (Yang, Ostrovsky, Rogers & Welker, 2016). Whether these dramatic differences in migration distance are the result of differential selection, or plasticity from genotype by environment interaction remains unknown. Eastern and western North American butterflies have divergent wing morphology (Altizer & Davis, 2010; Freedman & Dingle, 2018), and it is often assumed that they form distinct genetic populations (Brower *et al.*, 1995; NatureServe, 2019). However, observational studies (Brower & Pyle, 2004) and limited allozyme and microsatellite studies (Shephard, Hughes & Zalucki, 2002; Lyons *et al.*, 2012) have indicated large amounts of genetic exchange between eastern and western monarchs. This lack of genome-wide genetic differentiation suggests that migratory differences may instead be driven by restricted loci or differential environment-induced gene expression (Liedvogel *et al.*, 2011). Here, we compare flight performance of eastern and western monarchs, carry out an analysis of 43 genomes (Fig. 1), and measure the expression of a small number of candidate genes in eastern and western monarchs during flight trials.

Materials and Methods

Flight trials

We collected eastern monarchs (n=32; 17 male, 15 female) from migratory stopover site St. Marks, FL in October 2016 and western monarchs (n=31; 16 male, 15 female) from an overwintering site near Oceano, CA in December 2016 to perform flight trials. All butterflies were housed in overwintering-like conditions in an incubator to ensure they were in the same overwintering state during flight trials in December 2016. We used two flight mills as described in Bradley et al 2005 (Bradley & Altizer, 2005) and an ASCO PS-2000 datalogger (Pasco Scientific, Roseville, CA, USA) to allow eastern and western monarchs to fly in continuous circles of 4.27 m circumference. We recorded the time elapsed between each rotation (to measure instantaneous speed), the cumulative flight time, and the body mass of the monarch pre- and post- flight trial.

Flight trials were performed in a laboratory space maintained at 25°C and controlled for light conditions. One day prior to flight trials, monarchs were removed in groups and a steel wire attachment (32 gauge, 9 cm long) was glued to the dorsal side of the thorax using rubber cement.

Following wire attachment, monarchs were held in cylindrical mesh flight cages (diameter= 0.38 m, height= 0.56 m) to allow acclimation to the wire and for free feeding. We calculated five measures of flight performance: flight duration, distance, loss of body mass relative to total distance flown, power, and speed. During the flight trials, we allowed monarchs to fly for 30 minutes with trials ending prior to the 30 minute maximum if monarchs suspended flight for more than 10 seconds on 3 separate occasions. Flight trials were considered unsuccessful if the monarch refused to fly at least one full rotation on the flight mill. We then measured flight performance for the 29 western monarchs (14 male, 15 female) and 27 eastern monarchs (14 male, 13 female) which successfully completed flight trials. Distance was measured as total distance in meters of a flight trial. Loss of body mass was calculated as the change in body mass ($mass_{initial}$ - $mass_{final}$) divided by distance flown (in m), then log-transformed. Power was calculated as ($1/2*mass*velocity^2$) divided by time (in s). Speed (m/s) was averaged across 2-minute intervals for the duration of each flight in order to calculate the average flight speed.

We also measured morphological traits relevant to flight, including wing size and wing shape. We measured these traits to determine if any differences in flight behavior were due to differences in wing morphology. Following existing protocols (Altizer & Davis, 2010; Li, Pierce & de Roode, 2016), forewings were scanned on a flatbed scanner and the Fovea Pro plugin (Reindeer Graphics, Inc., Asheville, NC) for Adobe Photoshop was used to measure forewing area, length, breadth and perimeter. From these measurements we calculated aspect ratio, by dividing length by breadth of the forewing, and roundness, by using the equation $4*\pi*area/ (perimeter)^2$ (Altizer & Davis, 2010). Using Principal Component Analysis (PCA), forewing area, length, and width were reduced into one variable (PC1) to measure forewing size, while forewing aspect ratio and roundness were reduced to a second variable (PC2) to measure forewing shape.

We used analysis of variance (ANOVA) in R 3.1.3 (R Development Core Team, 2012) to test for differences in PC1 (wing size) and PC2 (wing shape) between eastern and western monarchs. We used analysis of covariance to test for differences in flight duration, flight distance, loss of body mass relative to total distance flown, flight power, and flight speed between eastern and western monarchs. In these analyses, we included butterfly sex as an additional explanatory variable and included PC1 (wing size) and PC2 (wing shape) as covariates. Significance of terms in analyses of variance and covariance was assessed by model simplification followed by model comparison using the command “anova” (Crawley, 2007). Models were plotted to verify the assumptions of homogeneity of variance and normality of errors (Crawley, 2007).

Genome sequencing

We used publicly available re-sequencing data from 14 eastern monarchs (8 males, 6 females) from Zhan et al. 2014 (Zhan *et al.*, 2014), which were collected in 2007-2009, and 30 newly re-sequenced western monarchs (15 males, 15 females), which were collected in January 2015 (Table S1). These numbers far exceed the number of genomes required to provide high power estimates of genetic differentiation (Willing, Dreyer & Van Oosterhout, 2012). The eastern monarchs from Zhan et al. were collected among multiple stopover points along the east coast and the Mexican overwintering sites (Fig. 1). Western monarchs were sampled from three overwintering locations along the California Coast: Big Sur, Oceano, and Carpinteria (Fig. 1). The newly sequenced western samples were sequenced on an Illumina HiSeq2000 platform. Paired-end libraries were prepared using an Illumina paired-end library kit. We combined 10 samples in a sequencing lane (2 x 100 bp) to generate 12X depth of coverage on average (Table S1). Sequences have been submitted to European Nucleotide Archive (ENA) project PRJEB33413.

Read mapping and genotyping

Reads obtained from sequencing were trimmed to remove adapter sequences and bases with Phred quality less than 20 using Cutadapt v. 1.14 (Martin, 2011). Trimmed reads were then mapped to the publicly available monarch reference genome assembly (Zhan & Reppert, 2012), using BWA-MEM v. 0.7.12 (Li & Durbin, 2009). The resulting alignment files were then sorted using SAMtools v. 1.2 (Li *et al.*, 2009). Indel realignment, base recalibration and variant recalibration were performed using GATK v. 3.8.0 (McKenna *et al.*, 2010). Variants were called for each sample using the Haplotypecaller module in GATK v. 3.8.0 and were then genotyped using the GenotypeGVCFs module in GATK v. 3.8.0 (McKenna *et al.*, 2010). High confidence variants with variant quality score greater than 80 were selected to recalibrate variant quality scores using VQSR filtering in GATK v. 3.8.0 (McKenna *et al.*, 2010). Indels and variants within the repeat regions of the reference genome were then removed using Vcftools v. 0.1.15 (Danecek *et al.*, 2011). One sample (PL3) with the lowest mapping success and genome-wide depth was removed prior to downstream analysis. Our bioinformatic pipeline is visualized in Fig. S1.

Population structure analysis

SNPs that were covered in all the individuals were used to estimate the genetic structure of North American monarchs. Principal component analysis (PCA) of the SNPs was performed using SNPRelate (Zheng *et al.*, 2012) in the R/Bioconductor package. Cross validation error rates were

checked using ADMIXTURE (Alexander, Novembre & Lange, 2009) for 'K' values ranging from 1 to 5 to determine the total number of populations in the dataset.

Window-based population genetic analysis

To understand genetic differentiation in monarchs, we calculated various population genetic statistics using Vcftools v. 0.1.15 (Danecek *et al.*, 2011) for individual populations and for pairwise population comparisons. To reduce the number of false positives, we only considered SNPs that were covered in all individuals in the population for the population-based statistics and SNPs that were covered in all individuals in both populations in pairwise comparison statistics. Nucleotide diversity (θ_π) and Tajima's D (T_d) were calculated in windows of 10,000 base pairs (10kb) across the genome using Vcftools v. 0.1.15 (Danecek *et al.*, 2011). Western monarchs were down-sampled to match the number of eastern samples (8 males and 6 females) to calculate Tajima's D (T_d) and allele frequencies using Vcftools v. 0.1.15 (Danecek *et al.*, 2011). We calculated genetic differentiation (F_{ST}) for each site using Vcftools v. 0.1.15 (Danecek *et al.*, 2011) and averaged across the genome in windows of 10kb. To ensure that our conclusions were not driven by genomic window size, we also calculated genetic differentiation (F_{ST}) for different window sizes (100 bp, 500 bp and 5,000 bp) to verify our findings. Absolute divergence (D_{XY}) was calculated in windows of 10kb across the genome using the allele frequencies. Windows with less than 10% of total sites covered were filtered out to eliminate extremely high values. Fixed, shared and private polymorphisms were calculated between eastern and western monarchs using the allele frequencies. F_{ST} values were Z-transformed (F_{ST}^Z = (Window F_{ST} / Genome Average F_{ST}) / Standard deviation of Genome wide F_{ST}) to obtain the relative genetic differentiation in the windows to the genomic mean to identify outlier windows. The top 1% of the F_{ST}^Z values were selected as the genetic differentiation outliers and the bottom 1% Tajima's D values were selected as Tajima's D outliers.

Chromosome assignment

The current publicly available monarch genome used in this study consists of 5,397 scaffolds with an N50 value of 715.6 kB. These scaffolds were assigned to chromosomes using coverage-based chromosome assignments produced by Mongue *et al.* (Mongue, Nguyen, Volenikova & Walters, 2017). F_{ST}, θ_π, T_d and F_{ST}^Z outliers were calculated for autosomes, Z-chromosome, and neo-Z chromosomes separately.

Genome-wide phylogenetic relationships

We used SAGUARO (Zamani *et al.*, 2013) to investigate genome-wide phylogenetic relationships between samples, and to identify differentiated genomic regions and regions that may have introgressed between populations. We used the program VCF2HMM, implemented in SAGUARO, to convert VCF file to HMM format, and implemented this file to analyze phylogenetic patterns across autosomes and the Z chromosome.

Genetic diversity and demographic history

Genetic diversity (θ_π) was calculated for different site categories across the genome based on the publicly available genome annotation in MonarchBase (Zhan & Reppert, 2013). This was done separately in windows of 10kb for eastern and western monarchs. Genomic positions were categorized as intergenic, intronic, 1st, 2nd, 3rd codon positions and 4-fold degenerate sites (4D).

We used two different approaches to analyze demographic history. First, we used a diffusion approximation method of ∂a∂i (Gutenkunst, Hernandez, Williamson & Bustamante, 2010) to investigate the joint demographic history of eastern and western monarchs. For this analysis we only considered autosomal scaffolds as Z chromosomes have different effective population sizes and a low density of SNPs, which could affect the resulting site frequency spectrum. We generated a Two-Dimensional Joint Site Frequency Spectrum (2D-JSFS) of eastern and western monarchs using the "dadi.Misc.make_data_dict_vcf" function provided with ∂a∂i. We scanned for the likelihoods of a set of 15 demographic models to address the following questions (Fig. S9): (1) have eastern and western monarchs diverged in the past; (2) if there is divergence, is there migration between the two populations; (3) if there is migration, what is the rate of migration; (4) what is the most likely scenario to explain changes in effective population size? We simulated 4 two-population models provided with ∂a∂i (Fig. S9-L to S9-O) and 11 two-population models provided with dadi_pipeline (github.com/dportik/dadi_pipeline) (Portik *et al.*, 2017). We used log likelihoods to find the most likely model that can explain the joint site frequency spectrum of eastern and western monarchs. We performed parameter optimization by running 100 iterations performed in a total of 4 rounds (10, 20, 30 and 40 iterations for the first, second, third and fourth round, respectively). Parameters with the highest log-likelihood were used as starting parameters of the next round. We used the Broyden Fletcher Goldfarb Shanno (BFGS) algorithm to optimize the parameters. Results of all 16 optimized models were summarized using

"Summarize_Outputs.py" provided with "dadi_pipeline". This script extracts the iterations with the highest log-likelihood and compares them between models. The model with the highest log-likelihood score and the lowest Akaike Information Criterion (AIC) was considered the most likely model to explain the 2D-JSFS.

As a second approach to understand the demographic history of eastern and western monarchs, we analyzed the genome-wide patterns of Tajima's D (T_D) calculated in windows of 10 kb using Vcftools v. 0.1.15 (Danecek *et al.*, 2011) to determine if eastern and western monarchs have different demographic histories.

Coverage-based SNP filtering

As we used previously sequenced genomes as well as new ones, sequencing coverage in the dataset varied from 7X to 25X. To remove coverage bias in our analysis we only used positions that are covered in all individuals within a population to calculate site frequency spectrum and population genetic statistics. To calculate genetic differentiation (F_{ST}) and absolute divergence (D_{XY}) we used positions covered in all individuals considered in the comparison. Filtering was done equally on polymorphic and non-polymorphic sites. We calculated nucleotide diversity for different coverage filters for all individuals (1X to 7X) to verify that the genome-wide nucleotide diversities are unbiased to coverage (Fig. S2).

Gene expression analysis

Eastern (n=10; 5 males, 5 females) and western (n=10; 5 males, 5 females) monarchs were randomly selected following the flight trials described above. They were subsequently flown on the flight mill for an additional two minutes and then immediately frozen in liquid nitrogen. Tri Reagent (Sigma) was used to extract RNA from the thorax of frozen samples. cDNA was synthesized from 600ng of total RNA using High Capacity cDNA Reverse Transcription kit (Thermofisher) according to manufacturer instructions.

Gene expression of six candidate genes was quantified relative to two housekeeping genes: 18S and 28S (Pan *et al.*, 2015). The six candidate genes included two dynein genes (DPOGS201379, DPOGS211203) and a myosin gene associated with motor activity (DPOGS200868) that were related to monarch migration phenotypes in the genomic analysis by Zhan et al. (Zhan *et al.*, 2014). We also measured expression of a neurotransmitter gated ion channel (GABA receptor) gene (DPOGS202675) that is involved in the invertebrate neuromuscular system (Lummis, 1990;

Schuske, Beg & Jorgensen, 2004), and a putative protein (DPOGS211604) whose homolog was found to be expressed in the wing disc of the silkworm (Mita *et al.*, 2003). Finally, we quantified expression of a myosin heavy chain gene associated with non-muscular motor activity (DPOGS215054) that controls flight in fruit flies (Wells, Edwards & Bernstein, 1996). For each monarch, we carried out three replicate PCR reactions for 18S, 28S and each of the six candidate genes. Expression of candidate genes relative to 18S and 28S was calculated as:

$$relative\ expression = 2^{-(C_{\mathrm{T}},candidate\ gene-(C_{\mathrm{T,18S}}+C_{\mathrm{T,28S}})/2)}$$

Gene-specific primers were used in PCR reactions (20µl) containing 7µl of ddH_2O, 10µl of 2xSYBR Green MasterMix (Bio-Rad), 1µl of each specific primer (10mM), and 1µl of first-strand cDNA template. The qPCR program included an initial denaturation for 3 min at 95°C followed by 40 cycles of denaturation at 95°C for 10s, annealing for 30s at 55°C, and extension for 30s at 72°C. For melting curve analysis, a dissociation step cycle (55°C for 10s, and then 0.5°C for 10s until 95°C) was added. All primers were tested for efficiency prior to use, and only primer pairs with the same efficiency as the primers for housekeeping genes were used. Primers for 18S and 28S were obtained from Pan et al. (Pan *et al.*, 2015). Candidate gene-specific primers used were as follows (F, forward primer; R, reverse primer): *DPOGS201379-F*, 5'-CTGACCAGCACGAAGAGAAA-3'; *DPOGS201379-R*, 5'- GACAATATCCCGGCGAATAGAA-3'; *DPOGS211203-F*, 5'-GATGCGATTGCTGCATTGAATA-3'; *DPOGS211203-R*, 5'-ATACCGCTGCCATCACTAAC-3'; *DPOGS202675*-F, 5'-CTCCCTTGTCGTGATGTTGT-3'; *DPOGS202675*-R, 5'-GTCGGCTCTCAATCCAGTAAA -3'; *DPOGS200868*-F, 5'-TCGGAACAGGAGGAGTATCT-3'; *DPOGS200868*-R, 5'-GCCTCTATGCCTCTCTTCTATG-3'; *DPOGS215054*-F, 5'-GTCGCTGACTTCTCCATCATAC-3'; *DPOGS215054*-R; 5'-GTTCTCGTTCAGAGGATCCATATT-3'; *DPOGS211604*-F, 5'-CAACGAGGAAGCCAGACTAAA-3'; *DPOGS211604*-R, 5'-TGTGGCATTGGTCTTCCATAA-3'.

Due to homogeneity-of-variance and error normality assumptions, we could not use linear models to analyze gene expression. Instead, we used one-tailed Mann-Whitney U tests in R 3.1.3 (R Development Core Team, 2012) to determine whether gene expression was higher in eastern than western monarchs.

Results

Flight performance

When testing monarchs on a tethered flight mill (Fig. 2A), eastern monarchs flew longer ($F_{1,54}$=8.81, P=0.004) and thereby realized greater flight distances than western monarchs (Fig. 2B, C; $F_{1,54}$=4.56, P=0.037). In contrast, western monarchs flew with greater power than eastern monarchs (Fig. 2D; $F_{1,54}$=6.00, P=0.018), which is expected in butterfly populations adapted to shorter flight distances (McKay, Ezenwa & Altizer, 2016). Female and male butterflies did not differ in flight duration ($F_{1,51}$=0.052, P=0.82), flight distance ($F_{1,51}$=0.10, P=0.75), or power ($F_{1,51}$=0.007, P=0.93). These differences in flight performance between eastern and western migratory monarchs correspond with their drastically different migration distances.

We also found significant differences in wing morphology of eastern and western monarchs. Wing morphology measurements showed that eastern monarchs had larger wings than western monarchs (PC1: $F_{1,54}$=11.0, P=0.002), confirming previous studies (Altizer & Davis, 2010; Freedman & Dingle, 2018). Wing morphology is an important determinant of migration in many winged animals (Altizer & Davis, 2010) and migratory monarchs have larger forewings than non-migratory monarchs (Dockx, 2007; Altizer & Davis, 2010; Dockx, 2012; Li *et al.*, 2016; Yang *et al.*, 2016; Flockhart *et al.*, 2017). In contrast, eastern and western monarchs did not differ in wing shape (PC2: $F_{1,54}$=2.05, P=0.16). Additionally, wing size and shape did not significantly affect flight duration (PC1: $F_{1,53}$=1.40, P=0.24; PC2: $F_{1,52}$=0.34, P=0.56), flight distance (PC1: $F_{1,53}$=1.73, P=0.19; PC2: $F_{1,52}$=0.82, P=0.36) or flight power (PC1: $F_{1,53}$=0.47, P=0.50; PC2: $F_{1,52}$=1.68, P=0.20). We also found no differences in wing size (PC1: $F_{1,53}$=0.042, P=0.84) and shape (PC2: $F_{1,53}$=1.42, P=0.24) between females and males (within location).

Average speed did not vary significantly between eastern and western monarchs ($F_{1,54}$=0.001, P=0.98) or between male and female butterflies ($F_{1,51}$=0.20, P=0.66); moreover, wing size and wing shape did not affect average speed (PC1: $F_{1,53}$=1.05, P=0.31; PC2: $F_{1,52}$=1.81, P=0.19). Similarly, average weight loss during flight trials did not vary between eastern and western monarchs ($F_{1,54}$=0.0007, P=0.98), and also was not affected by sex, wing size, and wing shape (Sex: $F_{1,51}$=0.84, P=0.36; PC1: $F_{1,53}$=0.002, P=0.97; PC2: $F_{1,52}$=1.42, P=0.24).

Population structure and genetic differentiation

The observed differences in flight performance may suggest strong population differentiation, but our genomic analyses revealed that this is not the case. The GATK genotyping pipeline resulted in a total of 32.2 million SNPs. A total of 25.9 million SNPs passed quality control filtering to be used for further analysis. A total of 20.9 million SNPs that were covered in all 43 samples were used for Principal Component Analysis (PCA) and ADMIXTURE (Alexander *et al.*, 2009) analysis to determine the genetic structure in the dataset. The PCA generated using the SNPrealte (Zheng *et al.*, 2012) package showed no evidence for clustering of eastern and western monarchs (Fig. 3), and determined that all samples in the data set belong to one population with K=1 showing the lowest error rate (Fig. 3; Table S2). While eastern and western monarchs could not be separated, three samples from Big Sur, California, seemed to cluster.

Genome-wide genetic differentiation (F_{ST}) and absolute divergence (D_{XY}) between eastern and western monarchs, calculated in windows of 10kb, were extremely low (Table 1, Fig. 4). Genetic differentiation between monarchs from the three western overwintering sites was likewise low (Tables 2, S3; Fig. S3). The genome-wide differentiation landscapes of eastern and western monarch comparisons were highly correlated with the genome-wide differentiation landscapes of the comparisons between western monarchs from different overwintering sites (Fig. S4). The maximum values of F_{ST} in all comparisons were also extremely low (Table S3), as was the genetic differentiation within genes (eastern vs. western $F_{ST\,(\mathrm{Genes})} = 0.0008 \pm 0.0004$). F_{ST}^{Z} window outliers were calculated separately for autosomes, Z-chromosome and neo-Z-chromosome, and were very low (Table 3). The extremely low maximum measures of genetic differentiation and the top 1% F_{ST}^{Z} window outliers suggest that there are no regions in the genome with reduced gene flow. This is in contrast with many other species, where islands of differentiation appear to be common (Jiggins, Naisbit, Coe & Mallet, 2001; Martin *et al.*, 2013; Cruickshank & Hahn, 2014; Nadeau *et al.*, 2014; Talla, Kalsoom, Shipilina, Marova & Backström, 2017; Irwin *et al.*, 2018). These results were not driven by window size, as low differentiation was also found for window sizes of 100, 500 and 5,000bp (Fig. S5, Table S4).

In line with these results we also did not identify any fixed nucleotide differences between eastern and western monarchs (Fig. 5). The majority of polymorphisms are shared between eastern and western monarchs, while similar proportions of private polymorphisms are observed in both

eastern and western monarchs (Fig. 5). This supports the conclusion that there are no regions in the genome with restricted gene flow or islands of genetic differentiation.

Genome-wide phylogenetic relationships

SAGUARO (Zamani *et al.*, 2013) analysis resulted in a total of 11 possible phylogenetic relationships across the genome (Fig. S6). None of the phylogenetic relationship matrixes could separate eastern and western monarchs, consistent with the lack of differentiation based on PCA and ADMIXTURE analysis. The Z-chromosome showed a different pattern of phylogenetic relationships compared to the autosomes (Fig. S7, Table S5), with an over-representation of certain cacti (Cactus 0,1,7,8, and 10). As with our ADMIXTURE results, while eastern and western monarchs could not be separated, three monarchs from Big Sur, California, did appear to cluster (Fig. S6).

Genetic diversity and demographic history

Levels of genetic diversity were calculated in 10kb windows separately for eastern and western monarchs. Genomic windows were classified into autosomes, Z chromosome and neo-Z chromosome. Levels of genome-wide genetic diversity of eastern and western monarchs were essentially identical (Fig. 6A, S8, Tables 1, S6), with the genome-wide genetic diversity landscape of eastern and western monarchs showing an almost perfect correlation (Fig. S8). This provides further evidence for a lack of genome-wide genetic differentiation. The Z chromosome had a lower nucleotide diversity than the autosomes, and the neo-Z chromosome had an intermediate nucleotide diversity, reflective of their effective population sizes (Table 1). The neo-Z chromosome was identified to be an ancestral autosome but fused to the Z chromosome in the monarch butterfly (Gu *et al.*, 2019). The dosage compensation of the neo-Z chromosome segment looks more similar to the autosomes than the ancestral Z chromosome (Gu *et al.*, 2019). Consistent with this finding, we found that the genetic diversity of the neo-Z chromosome segment is higher than the ancestral Z-chromosome (Table 1). In line with genetic diversity (θ_π), genome-wide Tajima's D was also highly similar in eastern and western monarchs (Table 1; Fig. 4, 6B) suggesting that eastern and western monarchs have a similar demographic history. The negative genome-wide Tajima's D in both eastern and western monarchs indicates a recent recovery from a past bottleneck. We found a total of 128 common Tajima's D outlier windows between eastern (out of 223 windows) and western (out of 202 windows) monarchs (Table S7). We did not identify any common windows between the Tajima's D outliers and genetic differentiation outliers,

suggesting that the low genome-wide genetic differentiation is the effect of consistent gene flow between these two groups rather than random genetic drift.

We used ∂a∂i to scan the likelihoods of 15 models in total to find the best demographic model to explain the 2D-SFS of eastern and western monarchs (Fig. S9). A summary of the likelihood scores of the optimized models is given in Table S8. An extended table with the best 5 iterations in each model and their optimized parameters is given in Table S9. The model "bottlegrowth_split_mig" showed the highest log-likelihood score and the lowest Akaike Information Criterion (AIC) (Fig. 7, S9, Table S9). This model assumes a bottleneck before divergence between the two populations, followed by an exponential growth in population size with migration. Top 5 iterations of "bottlegrowth_split_mig" gave consistently higher likelihood scores than all other models. Both "no divergence" and "no migration" models had low likelihood scores in the optimization. Models with bottleneck and exponential growth before the split showed the highest likelihood scores, giving the weight to this scenario (Tables S8, S9). The model "bottlegrowth" which considers a bottleneck followed by exponential growth without a split into two populations had low likelihood (Fig. S9, Tables S8, S9). We visually verified the most likely model by plotting data and model outputs using the "Plot_2D" function in ∂a∂i (Fig. S10). We used the parameters of the iteration with the highest log-likelihood score for the model "bottlegrowth_split_mig" to calculate effective population sizes, migration rate and time of the bottleneck (Appendix 1 in Supplementary Materials). According to these parameter estimates, eastern and western monarchs experienced a bottleneck about 412 thousand years, underwent exponential growth, then diverged about 112 thousand years ago with a migration rate of $7.33*10^{-07}$ per generation (Fig. 7). ∂a∂I analysis also showed that eastern and western monarchs have similar effective population sizes with a symmetric migration (Fig. 7, S9, Table S6, Appendix 1). This finding was also confirmed by the Tajima's D values calculated across the genome in windows of 10kb. Tajima's D was similar for eastern and western monarchs (Fig. 4, 6B, Table 1).

Gene expression

Eastern and western monarchs were randomly selected, flown on the flight mill for two minutes, and immediately frozen in liquid nitrogen prior to RNA extraction. Overall, the expression of the six candidate genes varied widely between individual butterflies, and expression of several genes tended to be higher in western monarchs (Fig. 8). Mann-Whitney U tests showed that western

monarchs had significantly higher expression of the myosin heavy chain gene associated with non-muscular motor activity (Fig. 8; DPOGS215054; P=0.039).

Discussion

Our analysis of more than 20 million SNPs shows that eastern and western North American monarchs have extremely low genome-wide genetic differentiation. We did not detect any fixed nucleotide differences between eastern and western monarchs, and even the smallest window (100bp) size analyses indicated low maximum F_{ST} values of 0.06, indicating a lack of genomic islands of differentiation. The windows with maximum genetic differentiation between eastern and western monarchs were low compared to the genome-wide average genetic differentiation between subpopulations in other butterflye species (Nadeau *et al.*, 2013; Talla *et al.*, 2019; Martin *et al.*, 2020). We found an almost perfect genome-wide correlation between nucleotide diversity in eastern and western monarchs, and genome-wide phylogenetic analyses indicated no clustering of eastern and western monarchs. Both ∂a∂i and Tajima's D results suggest that eastern and western monarchs have a similar effective population size. Importantly, both methods contrast with population census data, which show much smaller numbers of western than eastern monarchs (Schultz, Brown, Pelton & Crone, 2017; Malcolm, 2018; Pelton, Schultz, Jepsen, Black & Crone, 2019), and support the notion of frequent genetic exchange between these monarchs.

Our results are in line with previous studies based on more limited genetic markers, including allozymes (Shephard *et al.*, 2002) and microsatellites (Lyons *et al.*, 2012). These findings also support observational, geographical, and tagging studies that have suggested regular interchange between eastern and western monarchs (Brower & Pyle, 2004; Dingle, Zalucki, Rochester & Armijo-Prewitt, 2005; Morris, Kline & Morris, 2015). Because the Rocky Mountains form a dispersal barrier for monarchs, the high levels of interchange between eastern and western monarchs indicated by our genomic analyses most likely occur during both the spring migration – when north-flying monarchs from Mexico could end up west of the Rocky Mountains (Brower & Pyle, 2004) – and autumn migration, when monarchs from western North America can end up migrating south to Mexico (Morris *et al.*, 2015; Billings, 2019). Indeed, population genetic analyses of microsatellites are consistent with a radial dispersal of monarchs from Mexico, including northward dispersal to western North America (Pierce, Altizer, Chamberlain, Kronforst & de Roode, 2015). Further tagging studies will be necessary to map the migration routes of western monarchs (Dingle *et al.*, 2005; James *et al.*, 2018) and to determine where actual genetic exchanges

between eastern and western monarchs are occurring. It is interesting to note that in our study, three samples from Big Sur, California, appeared to cluster in both the ADMIXTURE and SAGUARO analyses. This suggests potential genetic sub-structuring in western North America, consistent with the non-random migratory pathways of western overwintering monarchs inferred by tagging and stable isotope studies (Nagano *et al.*, 1993; Yang *et al.*, 2016).

As with many migratory species, monarch migration has a genetic basis, and genome comparisons between migratory and non-migratory populations have revealed strong evidence for the existence of migration-related genes (Zhan *et al.*, 2014). While migration *per se* is genetically determined and associated with large-effect alleles, the lack of genomic divergence observed here suggests that differences in migration routes, distances and destinations for migratory monarchs are not. Our flight trials clearly demonstrated differences in flight performance between eastern and western monarchs, and these phenotypic differences may be driven by either a large number of small-effect alleles or by differential gene expression (Liedvogel *et al.*, 2011) induced by environmental triggers in eastern and western North America. To conclusively discern between these two alternatives, one would ideally carry out genetic crosses between eastern and western monarchs, and then release both parental genotypes and cross-progeny offspring on both sides of the Rocky Mountains to study migratory behavior, for example through the use of radio-tracking tagged monarchs (Wilcox *et al*., 2020). However, current United States Department of Agriculture regulations prohibit the transfer and release of monarchs across the Rocky Mountains – partly based on the assumption that eastern and western monarchs are genetically distinct populations – preventing such a definitive study design.

However, some progress could be made by comparing the transcriptomes of eastern and western monarchs throughout the year. Our preliminary gene expression studies showed a trend for differential gene expression, and one gene related to non-muscular motor activity was significantly differentially expressed. While these results show that eastern and western monarchs may differentially express migration-related genes during active flight, it is likely that many other genes are differentially expressed both during active flight, and in the developmental stages leading up to migration. Previous studies have shown different transcriptome profiles between breeding and migratory monarchs in eastern North America, including differences in expression of genes related to juvenile hormone production (Zhu, Gegear, Casselman, Kanginakudru & Reppert, 2009). Other studies have shown that non-migratory monarchs in Australia, which evolved from migratory monarchs in North America (Zhan *et al.*, 2014), have retained the ability to enter

reproductive diapause (Freedman *et al.*, 2017), and that exposing eastern North American monarchs to artificial light and temperature conditions disrupts migration orientation behavior (Tenger-Trolander, Lu, Noyes & Kronforst, 2019). Our study further suggests that environmental variation on the east and west of the Rocky Mountains triggers monarchs to follow different pathways to develop into eastern and western migrants. Such factors may include the different species of milkweeds that monarchs use in eastern and western North America (Woodson, 1954; Dilts *et al.*, 2019), as recent work shows that milkweeds can significantly affect wing morphology (Davis & de Roode, 2018; Freedman & Dingle, 2018; Decker, Soule, de Roode & Hunter, 2019). A transcriptomics study examining differences during development and flight between eastern and western monarchs would be important in uncovering gene regulatory networks involved in migration ability, and further shine light on how these highly similar genomes can give rise to divergent migratory behavior.

Ultimately, determining the genetic or epigenetic basis of differential migration in eastern and western monarchs will not only advance our understanding of migration genetics, but may also have relevance for conservation biology. The population size of eastern migrating monarchs has dwindled over the last three decades (Vidal & Rendón-Salinas, 2014; Malcolm, 2018; Boyle, Dalgleish & Puzey, 2019), with some estimates indicating a decline over 80% from a high in 1996 (Semmens *et al.*, 2016). While studies disagree on the primary cause, an emerging picture is that monarch population decline is due to a combination of illegal logging at the Mexican overwintering sites, climate change, agriculture-induced loss of milkweed host plants in North America, and reduced availability of nectar sources along the fall migration flyways (Pleasants & Oberhauser, 2013; Vidal, López-García & Rendón-Salinas, 2014; Inamine, Ellner, Springer & Agrawal, 2016; Thogmartin *et al.*, 2017; Boyle *et al.*, 2019; Saunders *et al.*, 2019; Wilcox, Flockhart, Newman & Norris, 2019). Western monarch population size has also declined (Espeset *et al.*, 2016; Schultz *et al.*, 2017), reaching critically low levels in the 2018-2019 migrating season (Pelton, 2018; Pelton *et al.*, 2019). Monarch migration has been coined an endangered phenomenon (Brower *et al.*, 2012), and the population decline has led a group of organizations and scientists to petition the US Fish and Wildlife Service to protect monarchs under the Endangered Species Act (Center for Biological Diversity, Center for Food Safety, Xerces Society & Brower, 2014). Following recent advances in merging evolutionary biology with conservation biology (Hendry *et al.*, 2011; Lankau, Jørgensen, Harris & Sih, 2011; Sgro, Lowe & Hoffmann, 2011; Smith, Kinnison, Strauss, Fuller & Carroll, 2014), a crucial aspect of this process is to determine the adaptive capacity of monarch butterflies. This includes asking how much adaptive genetic variation monarch populations

harbor, and which populations must be preserved to allow the species to adapt to changing conditions and to preserve the processes that allow evolution to occur. If future studies reveal that differential eastern and western migration is driven by gene expression rather than by genetic differentiation, then this would suggest that preservation of eastern monarchs could potentially rescue western migration and *vice versa*. Future studies, ideally involving reciprocal translocation experiments, will be needed to address this important question.

Acknowledgements

We thank the Altizer lab for the use of flight mills, C. Gowler and A. Mongue for assistance with flight trials, and N. Barton and an anonymous reviewer for constructive feedback on a previous version of this manuscript. This work was supported by Emory University, National Science Foundation grants IOS-1557724, DEB-1754431 and IOS-1922720 to J.C.d.R., and National Science Foundation grant IOS-1452648 and National Institutes of Health grant GM108626 to M.R.K.

Author contributions

A.A.P., M.R.K. and J.C.d.R. conceived of the study; V.T. carried out population genomic analyses; A.A.P. and S.N. carried out sequencing; K.L.A. performed gene expression studies; T.J.B.d.M. assisted with bioinformatics; F.X.V. led sampling of western monarchs; A.A.P. and J.C.d.R. carried out and analyzed flight performance experiments; A.A.P., V.T., M.R.K. and J.C.d.R. wrote the manuscript and all other authors edited the manuscript.

Data availability

The sequencing reads have been submitted to European Nucleotide Archive (ENA) with the project accession number PRJEB33413. The scripts used for the genomic analysis can be found at https://github.com/venta380/Monarch_genomics. Flight, wing morphology and gene expression data have been uploaded to Dryad Digital Repository, doi:10.5061/dryad.hh4j1f6.

References

Alerstam, T. (2006) Conflicting evidence about long-distance animal navigation. *Science,* 313**,** 791-794.

Alerstam, T., Hedenstrom, A. & Åkesson, S. (2003) Long-distance migration: evolution and determinants. *Oikos,* 103**,** 247-260.

595 Alexander, D.H., Novembre, J. & Lange, K. (2009) Fast model-based estimation of ancestry in
596 unrelated individuals. *Genome Research,* 19**,** 1655-1664.
597 Altizer, S., Bartel, R. & Han, B.A. (2011) Animal migration and infectious disease risk. *Science,*
598 331**,** 296-302.
599 Altizer, S. & Davis, A.K. (2010) Populations of monarch butterflies with different migratory
600 behaviors show divergence in wing morphology. *Evolution,* 64**,** 1018-1028.
601 Bensch, S., Andersson, T. & Åkesson, S. (1999) Morphological and molecular variation across
602 a migratory divide in willow warblers, *Phylloscopus trochilus*. *Evolution,* 53**,** 1925-1935.
603 Billings, J. (2019) Opening a window on Southwestern monarchs: fall migrant monarch
604 butterflies, *Danaus plexippus* (L.), tagged synchronously in Southeastern Arizona
605 migrate to overwintering regions in rither Southern California or Central Mexico. *Journal*
606 *of the Lepidopterists' Society,* 73**,** 257-267.
607 Boyle, J.H., Dalgleish, H.J. & Puzey, J. (2019) Monarch butterfly and milkweed declines
608 substantially predate the use of genetically modified crops. *Proceedings of the National*
609 *Academy of Sciences,* 116**,** 3006-3011.
610 Bradley, C.A. & Altizer, S. (2005) Parasites hinder monarch butterfly flight: implications for
611 disease spread in migratory hosts. *Ecology Letters,* 8**,** 290-300.
612 Brower, L.P. (1995) Understanding and misunderstanding the migration of the monarch
613 butterfly (Nymphalidae) in North America: 1857-1995. *Journal of the Lepidopterists'*
614 *Society,* 49**,** 304-385.
615 Brower, L.P., Fink, L.S., Brower, A.V., Leong, K., Oberhauser, K., Altizer, S., . . . Zalucki, M.P.
616 (1995) On the dangers of interpopulational transfers of monarch butterflies. *Bioscience,*
617 45**,** 540-544.
618 Brower, L.P. & Pyle, R.M. (2004) The interchange of migratory monarchs between Mexico and
619 the western United States, and the importance of floral corridors to the fall and spring
620 migrations. In *Conserving Migratory Pollinators and Nectar Corridors in Western North*
621 *America* (ed. G.P. Nabhan). University of Arizona Press, pp. 144-166.
622 Brower, L.P., Taylor, O.R., Williams, E.H., Slayback, D.A., Zubieta, R.R. & Ramirez, M.I.
623 (2012) Decline of monarch butterflies overwintering in Mexico: is the migratory
624 phenomenon at risk? *Insect Conservation and Diversity,* 5**,** 95-100.
625 Center for Biological Diversity, Center for Food Safety, Xerces Society & Brower, L.P. (2014)
626 Petition to protect the Monarch Butterfly (*Danaus plexippus plexippus*) under the
627 Endangered Species Act. pp.
628 Crawley, M.J. (2007) *The R Book*. Chichester, John Wiley & Sons.
629 Cruickshank, T.E. & Hahn, M.W. (2014) Reanalysis suggests that genomic islands of
630 speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology,* 23**,**
631 3133-3157.
632 Dallimer, M. & Jones, P.J. (2002) Migration orientation behaviour of the red-billed quelea
633 *Quelea quelea*. *Journal of Avian Biology,* 33**,** 89-94.
634 Dallimer, M., Jones, P.J., Pemberton, J.M. & Cheke, R.A. (2003) Lack of genetic and plumage
635 differentiation in the red-billed quelea *Quelea quelea* across a migratory divide in
636 southern Africa. *Molecular Ecology,* 12**,** 345-353.
637 Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., . . . Sherry, S.T.
638 (2011) The variant call format and VCFtools. *Bioinformatics,* 27**,** 2156-2158.
639 Davis, A.K. & de Roode, J.C. (2018) Effects of the parasite, *Ophryocystis elektroscirrha*, on
640 wing characteristics important for migration in the monarch butterfly. *Animal Migration,* 5**,**
641 84-93.
642 Decker, L.E., Soule, A.J., de Roode, J.C. & Hunter, M.D. (2019) Phytochemical changes in
643 milkweed induced by elevated CO2 alter wing morphology but not toxin sequestration in
644 monarch butterflies. *Functional Ecology,* 33**,** 411-421.

Dilts, T., Steele, M., Engler, J.D., Pelton, E.M., Jepsen, S.J., McKnight, S., . . . Cruz, E.E. (2019) Host plants and climate structure habitat associations of the western monarch butterfly. *Frontiers in Ecology and Evolution,* 7**,** 188.

Dingle, H. (1972) Migration strategies of insects. *Science,* 175**,** 1327-1335.

Dingle, H. (1991) Evolutionary genetics of animal migration. *American Zoologist,* 31**,** 253-264.

Dingle, H. (2014) *Migration: The Biology of Life on the Move*. Oxford, Oxford University Press.

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,* 85**,** 491-500.

Dockx, C. (2007) 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,* 92**,** 605-616.

Dockx, C. (2012) Differences in phenotypic traits and migratory strategies between eastern North American monarch butterflies, *Danaus plexippus* (L.). *Biological Journal of the Linnean Society,* 106**,** 717-736.

Espeset, A.E., Harrison, J.G., Shapiro, A.M., Nice, C.C., Thorne, J.H., Waetjen, D.P., . . . Forister, M.L. (2016) Understanding a migratory species in a changing world: climatic effects and demographic declines in the western monarch revealed by four decades of intensive monitoring. *Oecologia,* 181**,** 819-830.

Flockhart, D.T.T., Fitz-gerald, B., Brower, L.P., Derbyshire, R., Altizer, S., Hobson, K.A., . . . Norris, D.R. (2017) Migration distance as a selective episode for wing morphology in a migratory insect. *Movement ecology,* 5**,** 7.

Flockhart, D.T.T., Wassenaar, L.I., Martin, T.G., Hobson, K.A., Wunder, M.B. & Norris, D.R. (2013) Tracking multi-generational colonization of the breeding grounds by monarch butterflies in eastern North America. *Proceedings of the Royal Society B-Biological Sciences,* 280**,** 20131087.

Freedman, M.G. & Dingle, H. (2018) Wing morphology in migratory North American monarchs: characterizing sources of variation and understanding changes through time. *Animal Migration,* 5**,** 61-73.

Freedman, M.G., Dingle, H., Tabuloc, C.A., Chiu, J.C., Yang, L.H. & Zalucki, M.P. (2017) Non-migratory monarch butterflies, *Danaus plexippus* (L.), retain developmental plasticity and a navigational mechanism associated with migration. *Biological Journal of the Linnean Society*.

Fricke, H.C., Hencecroth, J. & Hoerner, M.E. (2011) Lowland-upland migration of sauropod dinosaurs during the late Jurassic epoch. *Nature,* 480**,** 513–515.

Goehring, L. & Oberhauser, K.S. (2002) Effects of photoperiod, temperature, and host plant age on induction of reproductive diapause and development time in *Danaus plexippus*. *Ecological Entomology,* 27**,** 674-685.

Gu, L., Reilly, P.F., Lewis, J.J., Reed, R.D., Andolfatto, P. & Walters, J.R. (2019) Dichotomy of dosage compensation along the neo Z chromosome of the monarch butterfly. *Current Biology,* 29**,** 4071-4077.

Gustafsson, K.M., Agrawal, A.A., Lewenstein, B.V. & Wolf, S.A. (2015) The monarch butterfly through time and space: the social construction of an icon. *Bioscience,* 65**,** 612-622.

Gutenkunst, R., Hernandez, R., Williamson, S. & Bustamante, C. (2010) Diffusion approximations for demographic inference: DaDi. *Nature Precedings,* 1**,** https://doi.org/10.1038/npre.2010.4594.1031.

Hendry, A.P., Kinnison, M.T., Heino, M., Day, T., Smith, T.B., Fitt, G., . . . Zalucki, M.P. (2011) Evolutionary principles and their practical application. *Evolutionary Applications,* 4**,** 159-183.

694 Herman, W.S., Brower, L.P. & Calvert, W.H. (1989) Reproductive tract development in
695 monarch butterflies overwintering in California and Mexico. *Journal of the Lepidopterists'*
696 *Society,* 43**,** 50-58.
697 Herman, W.S. & Tatar, M. (2001) Juvenile hormone regulation of longevity in the migratory
698 monarch butterfly. *Proceedings of the Royal Society B-Biological Sciences,* 268**,** 2509-
699 2514.
700 Inamine, H., Ellner, S.P., Springer, J.P. & Agrawal, A.A. (2016) Linking the continental
701 migratory cycle of the monarch butterfly to understand its population decline. *Oikos,* 125**,**
702 1081-1091.
703 Irwin, D.E., Milá, B., Toews, D.P., Brelsford, A., Kenyon, H.L., Porter, A.N., . . . Irwin, J.H.
704 (2018) A comparison of genomic islands of differentiation across three young avian
705 species pairs. *Molecular Ecology,* 27**,** 4839-4855.
706 James, D.G., James, T.S., Seymour, L., Kappen, L., Russell, T., Harryman, B. & Bly, C.
707 (2018) Citizen scientist tagging reveals destinations of migrating monarch butterflies,
708 *Danaus plexippus* (L.) from the Pacific Northwest. *The Journal of the Lepidopterists'*
709 *Society,* 72**,** 127-144.
710 Jiggins, C.D., Naisbit, R.E., Coe, R.L. & Mallet, J. (2001) Reproductive isolation caused by
711 colour pattern mimicry. *Nature,* 411**,** 302-305.
712 Keightley, P.D., Pinharanda, A., Ness, R.W., Simpson, F., Dasmahapatra, K.K., Mallet, J., . . .
713 Jiggins, C.D. (2015) Estimation of the spontaneous mutation rate in *Heliconius*
714 *melpomene*. *Molecular Biology and Evolution,* 32**,** 239-243.
715 Lankau, R., Jørgensen, P.S., Harris, D.J. & Sih, A. (2011) Incorporating evolutionary principles
716 into environmental management and policy. *Evolutionary Applications,* 4**,** 315-325.
717 Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows–Wheeler
718 transform. *Bioinformatics,* 25**,** 1754-1760.
719 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009) The
720 sequence alignment/map format and SAMtools. *Bioinformatics,* 25**,** 2078-2079.
721 Li, Y., Pierce, A.A. & de Roode, J.C. (2016) Variation in forewing size linked to migratory
722 status in monarch butterflies. *Animal Migration,* 3**,** 27-34.
723 Liedvogel, M., Åkesson, S. & Bensch, S. (2011) The genetics of migration on the move.
724 *Trends in Ecology & Evolution,* 26**,** 561-569.
725 Lummis, S.C. (1990) GABA receptors in insects. *Comparative Biochemistry and Physiology*
726 *Part C: Comparative Pharmacology,* 95**,** 1-8.
727 Lyons, J.I., Pierce, A.A., Barribeau, S.M., Sternberg, E.D., Mongue, A.J. & de Roode, J.C.
728 (2012) Lack of genetic differentiation between monarch butterflies with divergent
729 migration destinations. *Molecular Ecology,* 21**,** 3433-3444.
730 Malcolm, S.B. (2018) Anthropogenic impacts on mortality and population viability of the
731 monarch butterfly. *Annual Review of Entomology,* 63**,** 277-302.
732 Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing
733 reads. *EMBnet. journal,* 17**,** 10-12.
734 Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., . . .
735 Jiggins, C.D. (2013) Genome-wide evidence for speciation with gene flow in *Heliconius*
736 butterflies. *Genome Research,* 23**,** 1817-1828.
737 Martin, S.H., Singh, K.S., Gordon, I.J., Omufwoko, K.S., Collins, S., Warren, I.A., . . . Martins,
738 D. (2020) Whole-chromosome hitchhiking driven by a male-killing endosymbiont. *PLoS*
739 *Biology,* 18**,** e3000610.
740 McFarlan, J.T., Bonen, A. & Guglielmo, C.G. (2009) Seasonal upregulation of fatty acid
741 transporters in flight muscles of migratory white-throated sparrows (*Zonotrichia*
742 *albicollis*). *Journal of Experimental Biology,* 212**,** 2934-2940.

McKay, A.F., Ezenwa, V.O. & Altizer, S. (2016) Unravelling the costs of flight for immune defenses in the migratory monarch butterfly. *Integrative and Comparative Biology,* 56**,** 278-289.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . Daly, M. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research,* 20**,** 1297-1303.

McKinnon, L., Smith, P.A., Nol, E., Martin, J.L., Doyle, F.I., Abraham, K.F., . . . Bêty, J. (2010) Lower predation risk for migratory birds at high latitudes. *Science,* 327**,** 326-327.

Mita, K., Morimyo, M., Okano, K., Koike, Y., Nohata, J., Kawasaki, H., . . . Shimada, T. (2003) The construction of an EST database for *Bombyx mori* and its application. *Proceedings of the National Academy of Sciences,* 100**,** 14121-14126.

Mongue, A.J., Nguyen, P., Volenikova, A. & Walters, J.R. (2017) Neo-sex chromosomes in the Monarch butterfly, *Danaus plexippus*. *G3: Genes, Genomes, Genetics,* 7**,** 3281-3294.

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

Nadeau, N.J., Martin, S.H., Kozak, K.M., Salazar, C., Dasmahapatra, K.K., Davey, J.W., . . . Jiggins, C.D. (2013) Genome-wide patterns of divergence and gene flow across a butterfly radiation. *Molecular Ecology,* 22**,** 814-826.

Nadeau, N.J., Ruiz, M., Salazar, P., Counterman, B., Medina, J.A., Ortiz-Zuazaga, H., . . . Papa, R. (2014) Population genomics of parallel hybrid zones in the mimetic butterflies, *H. melpomene* and *H. erato*. *Genome Research,* 24**,** 1316-1333.

Nagano, C.D., Sakai, W.H., Malcolm, S.B., Cockrell, B.J., Donahue, J.P. & Brower, L.P. (1993) Spring migration of monarch butterflies in California. In *Biology and conservation of the monarch butterfly* (ed. M.P. Zalucki). Los Angeles CA Natural History Museum of Los Angeles County, pp. 217-232.

NatureServe (2019) NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. http://explorer.natureserve.org Arlington, Virginia NatureServe, pp.

O'Corry-Crowe, G.M., Suydam, R.S., Rosenberg, A., Frost, K.J. & Dizon, A.E. (1997) Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Nearctic revealed by mitochondrial DNA. *Molecular Ecology,* 6**,** 955-970.

O'Malley, K.G., Ford, M.J. & Hard, J.J. (2010) Clock polymorphism in Pacific salmon: evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B-Biological Sciences,* 277**,** 3703-3714.

Pan, H., Yang, X., Bidne, K., Hellmich, R.L., Siegfried, B.D. & Zhou, X. (2015) Selection of reference genes for RT-qPCR analysis in the monarch butterfly, *Danaus plexippus* (L.), a migrating bio-indicator. *PLoS One,* 10**,** e0129482.

Pelton, E. (2018) Early Thanksgiving Counts Show a Critically Low Monarch Population in California. Xerces Society, pp. https://xerces.org/2018/2011/2029/critically-low-monarch-population-in-california/.

Pelton, E.M., Schultz, C.B., Jepsen, S.J., Black, S.H. & Crone, E.E. (2019) Western monarch population plummets: status, probable causes, and recommended conservation actions. *Frontiers in Ecology and Evolution,* 7**,** 258.

Petit, E. & Mayer, F. (2000) A population genetic analysis of migration: the case of the noctule bat (*Nyctalus noctula*). *Molecular Ecology,* 9**,** 683-690.

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 *Monarchs in a Changing World: Biology and Conservation of an Iconic Insect* (eds K.O. Oberhauser, S. Altizer & K. Nail). Ithaca NY Cornell University Press, pp. 257-267.

Pleasants, J.M. & Oberhauser, K.S. (2013) Milkweed loss in agricultural fields because of herbicide use: effect on the monarch butterfly population. *Insect Conservation and Diversity,* 6**,** 135-144.
Portik, D.M., Leaché, A.D., Rivera, D., Barej, M.F., Burger, M., Hirschfeld, M., . . . Fujita, M.K. (2017) Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology,* 26**,** 5245-5263.
Postel, U., Thompson, F., Barker, G., Viney, M. & Morris, S. (2010) Migration-related changes in gene expression in leg muscle of the Christmas Island red crab *Gecarcoidea natalis*: seasonal preparation for long-distance walking. *Journal of Experimental Biology,* 213**,** 1740-1750.
Prince, D.J., O'Rourke, S.M., Thompson, T.Q., Ali, O.A., Lyman, H.S., Saglam, I.K., . . . 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.
R Development Core Team (2012).
Reppert, S.M. & de Roode, J.C. (2018) Demystifying monarch butterfly migration. *Current Biology,* 28**,** R1009–R1022.
Saunders, S.P., Ries, L., Neupane, N., Ramirez, M.I., García-Serrano, E., Rendón-Salinas, E. & Zipkin, E.F. (2019) Multiscale seasonal factors drive the size of winter monarch colonies. *Proceedings of the National Academy of Sciences,* 116**,** 8609-8614.
Schultz, C.B., Brown, L.M., Pelton, E. & Crone, E.E. (2017) Citizen science monitoring demonstrates dramatic declines of monarch butterflies in western North America. *Biological Conservation,* 214**,** 343-346.
Schuske, K., Beg, A.A. & Jorgensen, E.M. (2004) The GABA nervous system in *C. elegans*. *Trends in Neurosciences,* 27**,** 407-414.
Semmens, B.X., Semmens, D.J., Thogmartin, W.E., Wiederholt, R., López-Hoffman, L., Diffendorfer, J.E., . . . Taylor, O.R. (2016) Quasi-extinction risk and population targets for the Eastern, migratory population of monarch butterflies (*Danaus plexippus*). *Scientific Reports,* 6**,** 23265.
Sgro, C.M., Lowe, A.J. & Hoffmann, A.A. (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications,* 4**,** 326-337.
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,* 75**,** 437-452.
Smith, T.B., Kinnison, M.T., Strauss, S.Y., Fuller, T.L. & Carroll, S.P. (2014) Prescriptive evolution to conserve and manage biodiversity. *Annual Review of Ecology, Evolution, and Systematics,* 45**,** 1-22.
Talla, V., Johansson, A., Dincă, V., Vila, R., Friberg, M., Wiklund, C. & Backström, N. (2019) Lack of gene flow: narrow and dispersed differentiation islands in a triplet of *Leptidea* butterfly species. *Molecular Ecology,* 28**,** 3756-3770.
Talla, V., Kalsoom, F., Shipilina, D., Marova, I. & Backström, N. (2017) Heterogeneous patterns of genetic diversity and differentiation in European and Siberian Chiffchaff (*Phylloscopus collybita abietinus/P. tristis*). *G3: Genes, Genomes, Genetics,* 7**,** 3983-3998.
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***,** 201904690.
Thogmartin, W.E., Wiederholt, R., Oberhauser, K., Drum, R.G., Diffendorfer, J.E., Altizer, S., . . . Semmens, B. (2017) Monarch butterfly population decline in North America: identifying the threatening processes. *Royal Society open science,* 4**,** 170760.

Trivedi, A.K., Kumar, J., Rani, S. & Kumar, V. (2014) Annual life history-dependent gene expression in the hypothalamus and liver of a migratory songbird: insights into the molecular regulation of seasonal metabolism. *Journal of Biological Rhythms,* 29**,** 332-345.

Urquhart, F.A. & Urquhart, N.R. (1978) Autumnal migration routes of the eastern population of the monarch butterfly (*Danaus p. plexippus* L.; Danaidae; Lepidoptera) in North America to the overwintering site in the Neovolcanic Plateau of Mexico. *Canadian Journal of Zoology,* 56**,** 1759-1764.

Vidal, O., López-García, J. & Rendón-Salinas, E. (2014) Trends in deforestation and forest degradation after a decade of monitoring in the Monarch Butterfly Biosphere Reserve in Mexico. *Conservation Biology,* 28**,** 177-186.

Vidal, O. & Rendón-Salinas, E. (2014) Dynamics and trends of overwintering colonies of the monarch butterfly in Mexico. *Biological Conservation,* 180**,** 165-175.

Wells, L., Edwards, K.A. & Bernstein, S.I. (1996) Myosin heavy chain isoforms regulate muscle function but not myofibril assembly. *The EMBO journal,* 15**,** 4454-4459.

Wilcox, A.A., Flockhart, D., Newman, A.E. & Norris, D.R. (2019) An evaluation of studies on the potential threats contributing to the decline of eastern migratory North American monarch butterflies (*Danaus plexippus*). *Frontiers in Ecology and Evolution,* 7**,** 99.

Willing, E.-M., Dreyer, C. & Van Oosterhout, C. (2012) Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLoS One,* 7**,** e42649.

Woodson, R.E. (1954) The North American species of *Asclepias* L. *Annals of the Missouri Botanical Garden,* 41**,** 1-211.

Yang, L.H., Ostrovsky, D., Rogers, M.C. & Welker, J.M. (2016) Intra-population variation in the natal origins and wing morphology of overwintering western monarch butterflies *Danaus plexippus*. *Ecography,* 39**,** 998-1007.

Zamani, N., Russell, P., Lantz, H., Hoeppner, M.P., Meadows, J.R., Vijay, N., . . . Grabherr, M.G. (2013) Unsupervised genome-wide recognition of local relationship patterns. *BMC Genomics,* 14**,** 347.

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., . . . Kronforst, M.R. (2014) The genetics of monarch butterfly migration and warning colouration. *Nature,* 514**,** 317-321.

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C. & Weir, B.S. (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics,* 28**,** 3326-3328.

Zhu, H.S., Gegear, R.J., Casselman, A., Kanginakudru, S. & Reppert, S.M. (2009) Defining behavioral and molecular differences between summer and migratory monarch butterflies. *BMC Biology,* 7**,** 14.

Tables

Table 1

Genetic diversity (θ_π), Tajima's D (T_D) and absolute divergence (D_{XY}) calculated for eastern and western monarchs separately for autosomes, Z-chromosome and neo-Z chromosome.

	Auto	Z	neo Z	Genome wide
θ_π (East)	0.0118 ± 0.0042	0.0072 ± 0.0033	0.0097 ± 0.0034	0.0115 ± 0.0043
θ_π (West)	0.0115 ± 0.0040	0.0070 ± 0.0030	0.0094 ± 0.0031	0.0112 ± 0.0041
T_d (East)	-1.1694 ± 0.3180	-1.0388 ± 0.4054	-0.9418 ± 0.2616	-1.1595 ± 0.3255
T_d (West)	-1.1901 ± 0.3205	-1.0201 ± 0.4015	-0.8975 ± 0.2491	-1.1772 ± 0.3288
D_{XY}	0.0082 ± 0.0026	0.0049 ± 0.0020	0.0063 ± 0.0020	0.0079 ± 0.0027

Table 2

Genome wide genetic differentiation between monarchs collected in eastern North America and those collected at different overwintering sites in western North America.

	Big Sur, CA	Oceano, CA	Carpinteria, CA	Eastern
Big Sur, CA	0			
Oceano, CA	0.001237 ± 0.000516	0		
Carpinteria, CA	0.001332 ± 0.000544	0.000908 ± 0.000381	0	
Eastern	0.001425 ± 0.000550	0.001009 ± 0.000391	0.001115 ± 0.000422	0

Table 3

Levels of genetic differentiation (F_{ST}) across the genome calculated for autosomes, Z-chromosome and neo-Z chromosome. F_{ST} was also calculated within 1% F_{ST}^Z window outliers and genes.

	Auto	Z	neo Z	Genome wide
All windows	0.001071 ± 0.000358	0.000711 ± 0.000292	0.000867 ± 0.000292	0.001048 ± 0.000364
1% F_{ST}^Z outliers	0.001930 ± 0.000132	0.001481 ± 0.000058	0.001691 ± 0.000002	0.001901 ± 0.000168
Genes	0.000863 ± 0.000445	0.000493 ± 0.000289	0.000629 ± 0.000351	0.000840 ± 0.000446

Figures

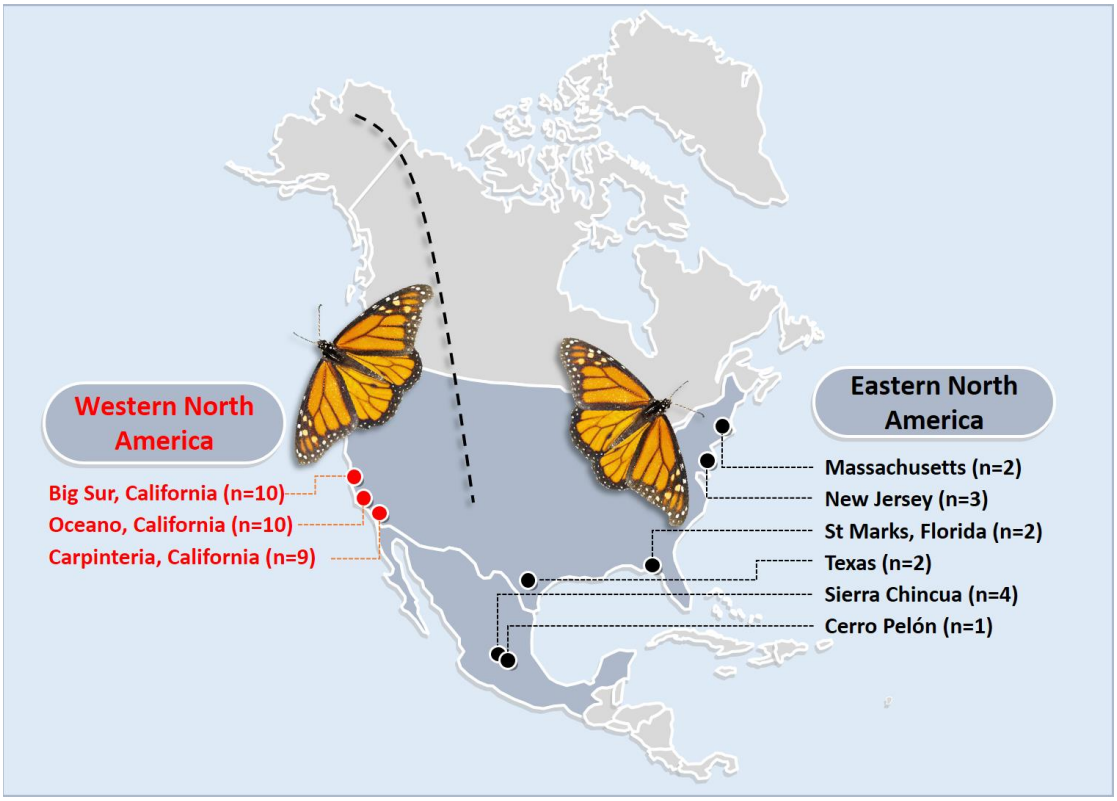


Figure 1. Map showing the sampling locations (and sample sizes) of monarchs used for genomic analyses. The Rocky Mountains (indicated by the schematic black dashed line) are generally believed to form a geographic barrier between eastern and western North American monarchs. Eastern monarchs (indicated in black) migrate to Mexico, where they overwinter at Oyamel forest sites including Sierra Chincua and Cerro Pelón. Western monarchs (indicated in red) migrate to eucalyptus and Monterey Pine groves along the Pacific Coast in California at many sites, including in Big Sur, Oceano, and Carpinteria.

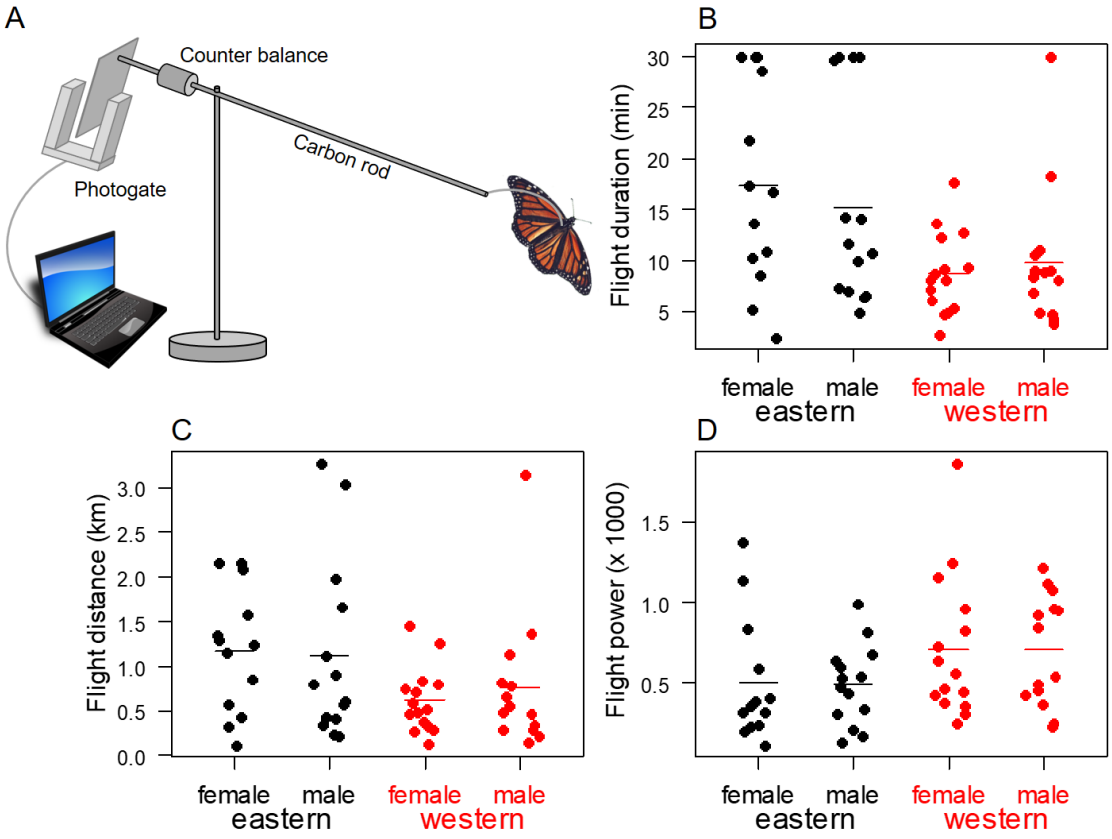
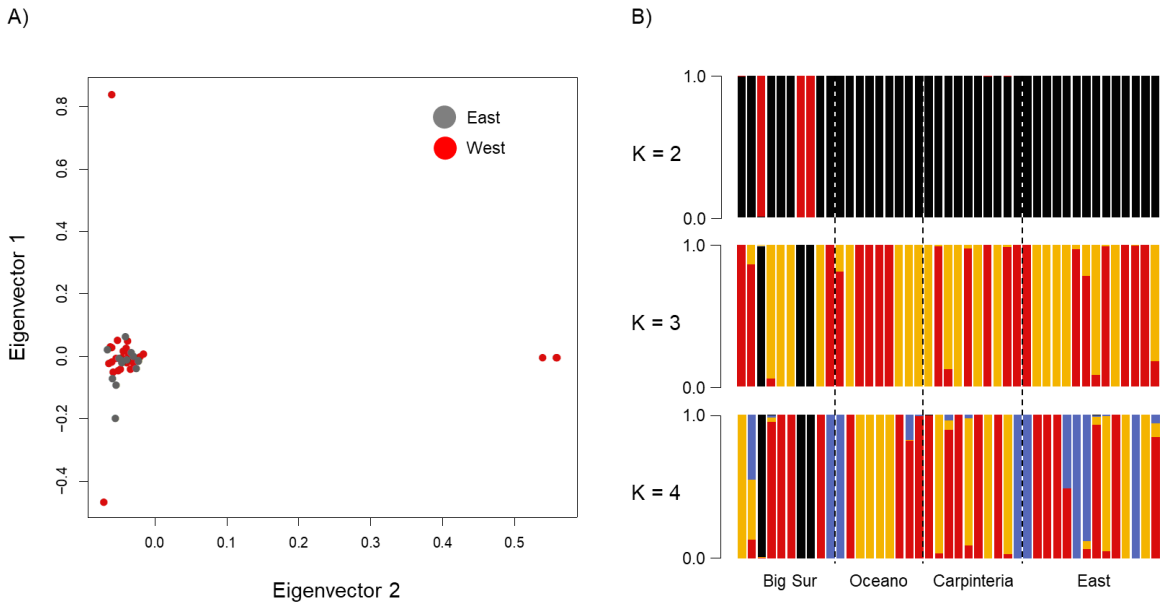


Figure 2. Flight performance of eastern and western North American monarchs. Butterflies were collected from St. Marks, a stopover of eastern monarchs on their way to Mexico, and Pismo Beach near Oceano, an overwintering site of western monarchs in California. Butterflies were placed on a tethered flight mill (A), and flight time (B), distance (C), and power were recorded during standardized flight trials. Data points show individual butterflies, while horizontal lines indicate means. Eastern butterflies flew significantly longer (P=0.004) and greater distances (P=0.04) than eastern monarchs, while western monarchs had higher powered flight (P=0.02).



933

934 **Figure 3. Lack of differentiation between eastern and western monarchs.** (A) Genetic
935 clustering was based on principal component analysis (PCA) of SNPs covered in all the samples
936 of eastern and western monarchs, generated using SNPrealte (Zheng *et al.*, 2012). Each point
937 represents one sample in the data set (black: eastern monarchs; red: western monarchs).
938 Although there are a few outliers from the western population, we did not observe any clear
939 clustering patterns of eastern and western monarchs (top 5 Principle components are given in
940 Figure S5). (B) Admixture plot for eastern and western monarchs with values of 'K' set to 2, 3 and
941 4. No specific pattern of clustering was observed between eastern and western monarchs.

942

943

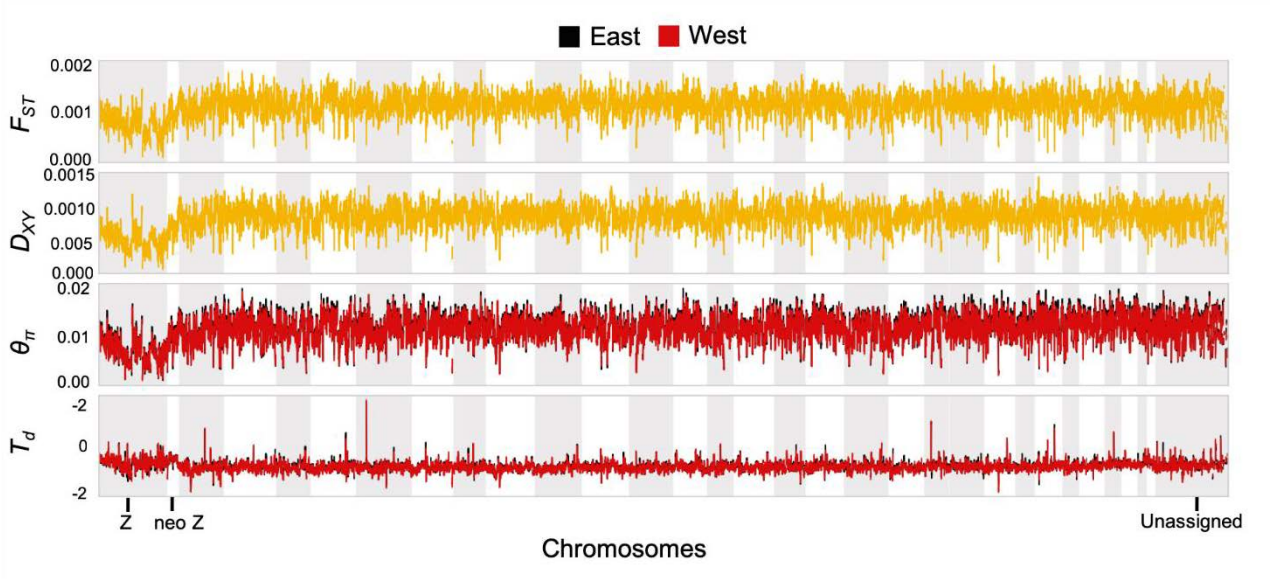


Figure 4. Genome-wide genetic differentiation (F_{ST}), regional variation in absolute divergence (D_{XY}), Tajima's D (T_D) and nucleotide diversity (θ_π) in eastern and western monarchs. These summary statistics were calculated in non-overlapping windows of 10kb across the genome. The alternating gray blocks represent different chromosomes in the genome. F_{ST} and D_{XY} are between-group comparisons, and a single yellow line is shown for these measures. In contrast, Tajima's D (T_d) and nucleotide diversity (θ_π) are group-specific measures, and eastern and western monarchs are indicated with black and red respectively; due to genome-wide overlap, the black (eastern) data are mostly hidden behind the red (western) data.

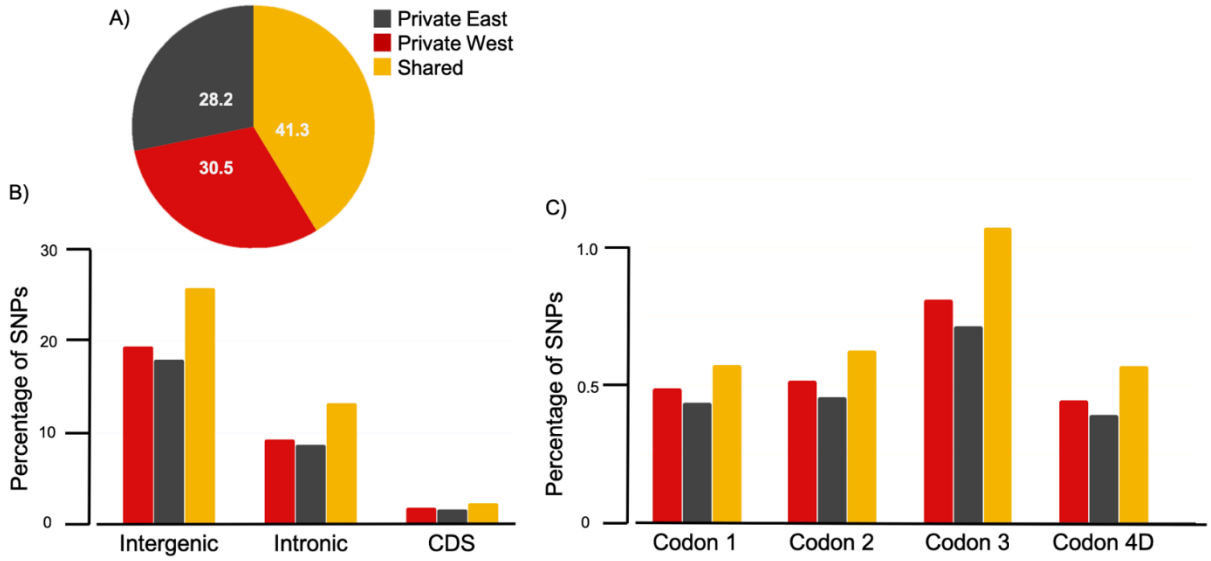
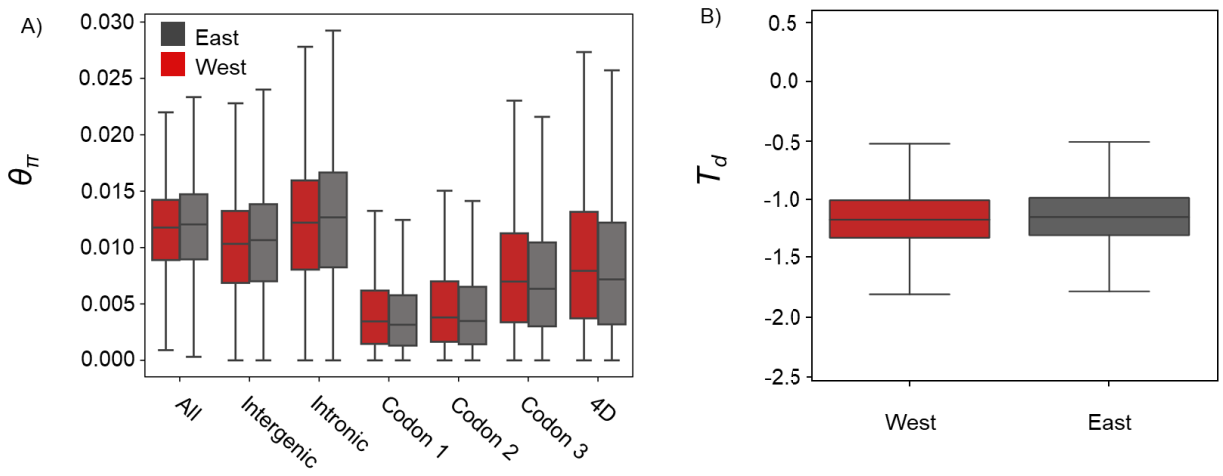


Figure 5. Allele sharing in eastern and western monarchs. (A) A visual illustration of percentages of SNPs that fall under private east, private west and shared between east and west. The division of these SNPs that fall under specific site categories such as intergenic, intronic, protein coding sequences (CDS) and different codon positions and 4-fold degenerate sites (Codon 4D) are shown given in panels B and C.



966
967

968 **Figure 6. Similarity in genetic diversity, Tajima's D and population demography in eastern**
969 **and western monarchs.** Panels A and B illustrate the similarity in the genome-wide genetic
970 diversity (θ_π) and Tajima's D (T_d) in eastern and western monarchs.

971
972
973

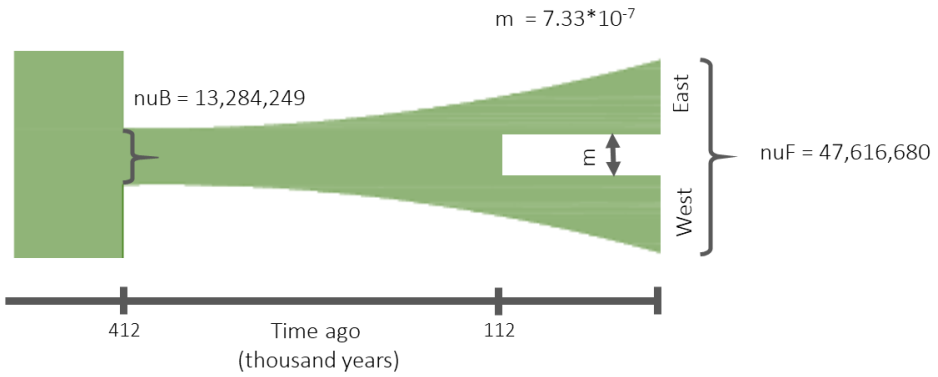


Figure 7. Visual representation of the most likely model in the demographic history analysis conducted using ∂a∂i. According to the model, eastern and western monarchs diverged 112 thousand years ago with a migration rate of $7.33*10^{-7}$. The populations went through a pre-divergence bottleneck about 412 thousand years ago. The effective population size during the bottleneck was estimated to be nuB 13,284,249 and the joint effective population size of eastern and western monarchs was estimated to be 47,616,680.

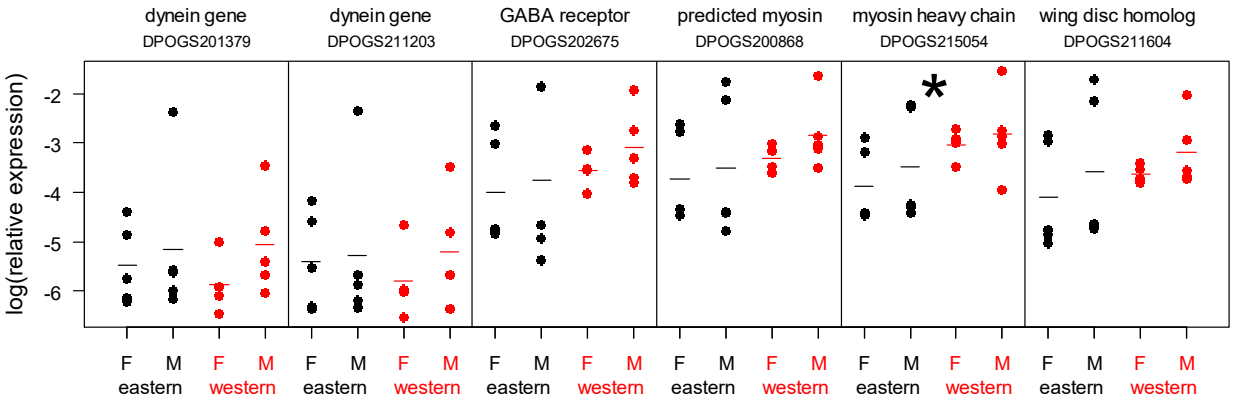


Figure 8. Analysis of gene expression of six genes involved in flight metabolism. Monarchs were subjected to two-minute flight trials and then flash-frozen. For each gene, the expression levels were calculated relative to 18S and 28S. Data points show individual butterflies, while horizontal lines indicate means. The asterisk indicates that the expression of the myosin heavy chain gene was significantly higher ($P<0.05$) in western monarchs.

Supporting information:

Genomic evidence for gene flow between monarchs with divergent migratory phenotypes and flight performance

Venkat Talla[a], Amanda A. Pierce[a], Kandis L. Adams[a], Tom J.B. de Man[a], Sumitha Nallu[c], Francis X. Villablanca[d], Marcus R. Kronforst[c] & Jacobus C. de Roode[a*]

[a] Department of Biology, Emory University, Atlanta, GA 30322, USA

[c] Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA
[d] Biological Sciences Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA

[*]Corresponding Author:
Jacobus de Roode
tel. 404-727-2340
email: jderood@emory.edu

Table S1

A summary of sample ID, collection site, collection date, sex, and mapping success of the samples to the monarch reference genome.

ID	Collection site	Date	Sex	Mapped %	Depth (X)	Accession	Population
HH1	Halcyon Hill, Oceano, CA, USA	January 2015	M	98.55	9.60	ENA: ERS3567792	West
HH2	Halcyon Hill, Oceano, CA, USA	January 2015	M	98.57	23.84	ENA: ERS3567793	West
HH3	Halcyon Hill, Oceano, CA, USA	January 2015	M	98.61	23.94	ENA: ERS3567794	West
HH4	Halcyon Hill, Oceano, CA, USA	January 2015	M	98.54	11.58	ENA: ERS3567795	West
HH5	Halcyon Hill, Oceano, CA, USA	January 2015	M	98.64	23.95	ENA: ERS3567796	West
HH6	Halcyon Hill, Oceano, CA, USA	January 2015	F	98.46	11.95	ENA: ERS3567797	West
HH7	Halcyon Hill, Oceano, CA, USA	January 2015	F	98.51	11.32	ENA: ERS3567798	West
HH8	Halcyon Hill, Oceano, CA, USA	January 2015	F	98.58	11.66	ENA: ERS3567799	West
HH9	Halcyon Hill, Oceano, CA, USA	January 2015	F	98.26	24.20	ENA: ERS3567800	West
HH10	Halcyon Hill, Oceano, CA, USA	January 2015	F	98.61	10.27	ENA: ERS3567801	West
PL1	Carpinteria, CA, USA	January 2015	M	98.56	22.46	ENA: ERS3567813	West

PL2	Carpinteria, CA, USA	January 2015	M	98.36	10.58	ENA: ERS3567814	West
PL3	Carpinteria, CA, USA	January 2015	M	34.66	2.84	NA	West
PL4	Carpinteria, CA, USA	January 2015	M	98.50	23.26	ENA: ERS3567815	West
PL5	Carpinteria, CA, USA	January 2015	M	98.16	21.97	ENA: ERS3567816	West
PL6	Carpinteria, CA, USA	January 2015	F	98.12	22.08	ENA: ERS3567817	West
PL7	Carpinteria, CA, USA	January 2015	F	98.57	23.94	ENA: ERS3567818	West
PL8	Carpinteria, CA, USA	January 2015	F	98.57	11.89	ENA: ERS3567819	West
PL9	Carpinteria, CA, USA	January 2015	F	98.32	11.87	ENA: ERS3567820	West
PL10	Carpinteria, CA, USA	January 2015	F	98.59	24.16	ENA: ERS3567821	West
ESB1	Esalen, Big Sur, California, USA	January 2015	M	98.55	23.23	ENA: ERS3567802	West
ESB2	Esalen, Big Sur, California, USA	January 2015	M	98.58	11.97	ENA: ERS3567803	West
ESB3	Esalen, Big Sur, California, USA	January 2015	F	98.42	8.35	ENA: ERS3567804	West
ESB4	Esalen, Big Sur, California, USA	January 2015	F	98.42	11.87	ENA: ERS3567805	West
ESB5	Esalen, Big Sur, California, USA	January 2015	F	98.46	24.09	ENA: ERS3567806	West
ESB6	Esalen, Big Sur, California, USA	January 2015	M	98.60	23.82	ENA: ERS3567807	West
ESB7	Esalen, Big Sur, California, USA	January 2015	F	98.62	23.86	ENA: ERS3567808	West
ESB8	Esalen, Big Sur, California, USA	January 2015	F	98.13	9.56	ENA: ERS3567809	West
ESB9	Esalen, Big Sur, California, USA	January 2015	M	98.40	10.77	ENA: ERS3567810	West
ESB10	Esalen, Big Sur, California, USA	January 2015	M	98.33	8.93	ENA: ERS3567811	West
stm163	St. Marks, FL, USA	October 2009	F	71.10	6.71	SRA: SRR155222	East
stm146	St. Marks, FL, USA	October 2009	M	98.12	22.98	SRA: SRR155222	East
T9	Texas, USA	March 30, 2009	M	87.91	7.93	SRA: SRR154952	East
T14	Texas, USA	March 28, 2009	M	97.29	9.71	SRA: SRR154952	East
NJ203	New Jersey, USA	Oct 1, 2009	F	95.55	10.94	SRA: SRR154857	East
NJ116	New Jersey, USA	September 30, 2009	M	86.38	14.81	SRA: SRR154857	East
NJ1	New Jersey, USA	September 30, 2009	F	95.91	10.04	SRA: SRR154857	East
HI023	Massachusetts, USA	July 3, 2008	M	99.13	11.95	SRA: SRR154857	East
HI033	Massachusetts, USA	August 10, 2009	F	97.86	7.02	SRA: SRR154857	East
mex986	Sierra Chincua, Mexico	January 9, 2007	M	97.42	9.47	SRA: SRR155220	East
mex919	Sierra Chincua, Mexico	January 9, 2007	M	98.43	10.77	SRA: SRR155220	East
mex915	Sierra Chincua, Mexico	January 9, 2007	F	98.23	9.80	SRA: SRR155220	East
mex536	Sierra Chincua, Mexico	February 11, 2008	M	98.01	9.64	SRA: SRR155220	East
mex1527	Cerro Pelon, Mexico	February 13, 2008	F	98.16	7.29	SRA: SRR155220	East

Table S2

Cross validation (CV) error rates calculated for SNPs to show the potential number of populations in a dataset.

K	CV error
1	0.35338
2	0.40578
3	0.4481
4	NA

Table S3

Maximum value for F_{ST} calculated in windows of 10kb across the genome for different population comparisons.

Comparisons	Max value of F_{ST}
East vs West	0.002525
Carpinteria vs Oceano	0.002752
Big_Sur vs Oceano	0.003849
Big_Sur vs Oceano	0.003849
Big_Sur vs Carpinteria	0.003693
Carpinteria vs East	0.003281
Oceano vs East	0.002958
Big_Sur vs East	0.004363

Table S4

Genome wide average F_{ST} and maximum F_{ST} calculated in varying window sizes of 100 bp, 500 bp, 5,000 bp and 10,000 bp.

Window size	Genome wide Average FST	Max value of FST
100	0.001213 ± 0.0012334	0.06025
500	0.001113 ± 0.0007157	0.01634
5000	0.001036 ± 0.0004367	0.00297
10000	0.001048 ± 0.000364	0.00253

Table S5

Summary of the genome-wide phylogenetic relationships (Cacti) calculated using Saguaro. The table shows the number of occurrences of each cactus on the Z chromosome and Autosomes.

Cacti	Z chromosome	Autosome
Cactus 0	33	3
Cactus 1	376	7
Cactus 2	3917	10398
Cactus 3	126	384
Cactus 4	9	14
Cactus 5	3	16
Cactus 6	1	7
Cactus 7	17	2
Cactus 8	47	33
Cactus 9	7	24
Cactus 10	764	381

Table S6

Summary table showing the genetic diversity ($\theta\pi$) calculated for each site class. These values were calculated across the genome in non-overlapping windows of 10kb.

	West ($\theta\pi$)	East ($\theta\pi$)
Intergenic	0.0105 ± 0.0133	0.0110 ± 0.0136
Intron	0.0121 ± 0.0058	0.0126 ± 0.0062
Codon 1	0.0047 ± 0.0057	0.0044 ± 0.0057
Codon 2	0.0052 ± 0.0072	0.0049 ± 0.0071
Codon 3	0.0081 ± 0.0071	0.0075 ± 0.0068
4D	0.0096 ± 0.0089	0.0089 ± 0.0087
All	0.0112 ± 0.0041	0.0115 ± 0.0043

Table S7

Total number of Tajima's D outliers (bottom 1%) in eastern and western monarchs and the windows shared between them in Autosomes, Z chromosome and Neo-Z chromosome.

	Autosome	Z	neo-Z
T_d Outliers East	188	33	2
T_d Outliers West	188	12	2
Common	120	7	1

Table S8

Best replicate of each of the optimized demographic models using ∂a∂i. These replicates were ordered as-per their log-likelihoods. The model with the highest log-likelihood was considered as the most likely model.

Model	Replicate	log-likelihood	AIC	ΔAIC	chi-squared	theta
bottlegrowth_split_mig	Round_4_Replicate_26	-3306.97	6623.94	0	4637.08	980321.63
bottlegrowth_split	Round_4_Replicate_22	-3815.46	7638.92	1014.98	5637.83	1011527.18
IM_pre	Round_3_Replicate_1	-12679.39	25374.78	18750.84	23631.5	2099707.76
sec_contact_asym_mig	Round_4_Replicate_17	-13172.32	26356.64	19732.7	24304.28	1107235.68
asym_mig	Round_4_Replicate_9	-13801.03	27612.06	20988.12	25491.42	1111105.32
asym_mig_size	Round_4_Replicate_8	-16014.33	32044.66	25420.72	30223.73	1110507.6
anc_asym_mig	Round_4_Replicate_5	-17059.42	34130.84	27506.9	31580.5	1068577.35
sec_contact_sym_mig	Round_3_Replicate_6	-19130.62	38271.24	31647.3	36124.26	1118371.68
bottlegrowth	Round_4_Replicate_39	-23066.45	46138.9	39514.96	44883.45	1032181.21
sym_mig_size	Round_2_Replicate_7	-34416.96	68847.92	62223.98	65460.71	611412.62
IM	Round_2_Replicate_13	-187843.39	375698.78	369074.84	355437.54	1081438.58
no_mig	Round_4_Replicate_16	-416097.42	832200.84	825576.9	902416.8	1646813.81
anc_sym_mig	Round_1_Replicate_5	-516780.88	1033571.76	1026947.82	1684798.33	161024.88
no_mig_size	Round_4_Replicate_4	-542602.75	1085217.5	1078593.56	1154052.67	1749776.67
no_divergence	Round_1_Replicate_1	-2224842.77	4449687.54	4443063.6	4738269.74	2614739.86

Table S9

An extended table of demographic models optimized using ∂a∂i. The table shows the top 5 replicates scored using their log-likelihoods for each model.

Model	Replicate	log-likelihood	AIC	chi-squared	theta
bottlegrowth_split_mig	Round_4_Replicate_26	-3306.97	6623.94	4637.08	980321.63
bottlegrowth_split_mig	Round_4_Replicate_6	-3307.1	6624.2	4637.82	980516.77
bottlegrowth_split_mig	Round_4_Replicate_35	-3307.2	6624.4	4637.67	981590.2
bottlegrowth_split_mig	Round_4_Replicate_10	-3307.24	6624.48	4637.31	976952.93
bottlegrowth_split_mig	Round_4_Replicate_14	-3307.25	6624.5	4637.62	981768.24
bottlegrowth_split	Round_4_Replicate_22	-3815.46	7638.92	5637.83	1011527.18
bottlegrowth_split	Round_4_Replicate_24	-3815.49	7638.98	5638.32	1010968.27
bottlegrowth_split	Round_4_Replicate_16	-3815.5	7639	5637.61	1011563.04
bottlegrowth_split	Round_4_Replicate_35	-3815.51	7639.02	5637.92	1011315.46
bottlegrowth_split	Round_4_Replicate_29	-3815.61	7639.22	5638.46	1011865.68
IM_pre	Round_3_Replicate_1	-12679.39	25374.78	23631.5	2099707.76
sec_contact_asym_mig	Round_4_Replicate_17	-13172.32	26356.64	24304.28	1107235.68
sec_contact_asym_mig	Round_4_Replicate_2	-13356.99	26725.98	24611.98	1109510.92
sec_contact_asym_mig	Round_4_Replicate_11	-13361.61	26735.22	24592.56	1107911.33
sec_contact_asym_mig	Round_4_Replicate_7	-13393.61	26799.22	24693.85	1110689.89
sec_contact_asym_mig	Round_4_Replicate_9	-13475.91	26963.82	24856.75	1109637.06
asym_mig	Round_4_Replicate_9	-13801.03	27612.06	25491.42	1111105.32
asym_mig	Round_4_Replicate_7	-13807.43	27624.86	25503.46	1110938.26
asym_mig	Round_4_Replicate_16	-13808.25	27626.5	25502.65	1111667.29
asym_mig	Round_4_Replicate_10	-13808.82	27627.64	25503.3	1111333.75
asym_mig	Round_4_Replicate_12	-13810.49	27630.98	25522.11	1111051.37
asym_mig_size	Round_4_Replicate_8	-16014.33	32044.66	30223.73	1110507.6
anc_asym_mig	Round_4_Replicate_5	-17059.42	34130.84	31580.5	1068577.35
anc_asym_mig	Round_4_Replicate_11	-17423.88	34859.76	32369.49	1053833.33
sec_contact_sym_mig	Round_3_Replicate_6	-19130.62	38271.24	36124.26	1118371.68
sec_contact_sym_mig	Round_3_Replicate_5	-19604.78	39219.56	37119.29	1123856.01
sec_contact_sym_mig	Round_3_Replicate_4	-19799.93	39609.86	37461.01	1126446.22
sec_contact_sym_mig	Round_3_Replicate_7	-19909.45	39828.9	37745.65	1124771.33
anc_asym_mig	Round_3_Replicate_6	-20130.37	40272.74	37429.51	1132598.63
sec_contact_sym_mig	Round_3_Replicate_8	-20159.65	40329.3	38238.38	1127388.28
anc_asym_mig	Round_4_Replicate_2	-20392.34	40796.68	37750.08	1133222.63
asym_mig_size	Round_4_Replicate_12	-20441.68	40899.36	37996.03	1044196.14
asym_mig_size	Round_4_Replicate_22	-21628.74	43273.48	40787.82	1005417.89

anc_asym_mig	Round_4_Replicate_10	-21688.56	43389.12	42336.13	1036632.6
bottlegrowth	Round_4_Replicate_39	-23066.45	46138.9	44883.45	1032181.21
bottlegrowth	Round_4_Replicate_10	-23066.5	46139	44888.02	1032788.39
bottlegrowth	Round_4_Replicate_31	-23066.74	46139.48	44882.23	1032340.6
bottlegrowth	Round_4_Replicate_17	-23067.26	46140.52	44891.81	1033863.61
bottlegrowth	Round_4_Replicate_33	-23069.3	46144.6	44897.06	1034226.22
asym_mig_size	Round_4_Replicate_3	-24520.58	49057.16	46558.6	1196443.42
asym_mig_size	Round_4_Replicate_7	-24730.61	49477.22	46682.75	989573.67
sym_mig_size	Round_2_Replicate_7	-34416.96	68847.92	65460.71	611412.62
IM_pre	Round_1_Replicate_5	-47027.58	94071.16	89313.16	1561174.95
IM_pre	Round_2_Replicate_10	-58432.45	116880.9	114099.3	2631910.67
sym_mig_size	Round_2_Replicate_14	-89038.18	178090.36	169280.62	1150049.25
IM_pre	Round_1_Replicate_9	-183087	366190	396547.83	650656.34
IM	Round_2_Replicate_13	-187843.39	375698.78	355437.54	1081438.58
sym_mig_size	Round_2_Replicate_16	-199810.02	399634.04	394830.23	675074.86
IM_pre	Round_3_Replicate_18	-208553.05	417122.1	426057.68	1101831.59
IM	Round_1_Replicate_3	-247348.65	494709.3	481068.01	1436670.21
sym_mig_size	Round_3_Replicate_2	-289788.82	579591.64	574974.54	1026012.07
sym_mig_size	Round_3_Replicate_1	-312280.77	624575.54	626270.96	654776.05
no_mig	Round_4_Replicate_16	-416097.42	832200.84	902416.8	1646813.81
no_mig	Round_4_Replicate_1	-416099.35	832204.7	902412.34	1646618.45
no_mig	Round_2_Replicate_10	-416104.44	832214.88	902265.68	1647331.94
no_mig	Round_4_Replicate_3	-416108.28	832222.56	902937.69	1646652.8
no_mig	Round_4_Replicate_7	-416111.53	832229.06	901563.18	1649035.21
anc_sym_mig	Round_1_Replicate_5	-516780.88	1033571.76	1684798.33	161024.88
no_mig_size	Round_4_Replicate_4	-542602.75	1085217.5	1154052.67	1749776.67
no_mig_size	Round_4_Replicate_28	-550320.14	1100652.28	1171633.89	1750530.39
IM	Round_1_Replicate_8	-573259.32	1146530.64	1169150.97	1484066.85
no_mig_size	Round_4_Replicate_9	-638987.74	1277987.48	1569683.36	1518467.52
no_mig_size	Round_4_Replicate_38	-658551.6	1317115.2	1392713.67	1779673.4
no_mig_size	Round_4_Replicate_34	-674082.67	1348177.34	1392012.18	1824827.44
anc_sym_mig	Round_1_Replicate_1	-912880.54	1825771.08	1710370.22	377100.26
anc_sym_mig	Round_1_Replicate_4	-1033490.47	2066990.94	2009989.38	1098132.31
IM	Round_1_Replicate_10	-1129007.21	2258026.42	2652043.88	3531478.47
IM	Round_1_Replicate_5	-1586466.39	3172944.78	3012671.12	1882015.03
anc_sym_mig	Round_1_Replicate_2	-2019707.72	4039425.44	4423788.46	553655.17
no_divergence	Round_1_Replicate_1	-2224842.77	4449687.54	4738269.74	2614739.86
no_divergence	Round_1_Replicate_2	-2224842.77	4449687.54	4738269.74	2614739.86
no_divergence	Round_1_Replicate_3	-2224842.77	4449687.54	4738269.74	2614739.86
no_divergence	Round_1_Replicate_4	-2224842.77	4449687.54	4738269.74	2614739.86

no_divergence	Round_1_Replicate_5	-2224842.77	4449687.54	4738269.74	2614739.86
anc_sym_mig	Round_1_Replicate_3	-11021112.9	22042235.8	72389626.2	1669358.18

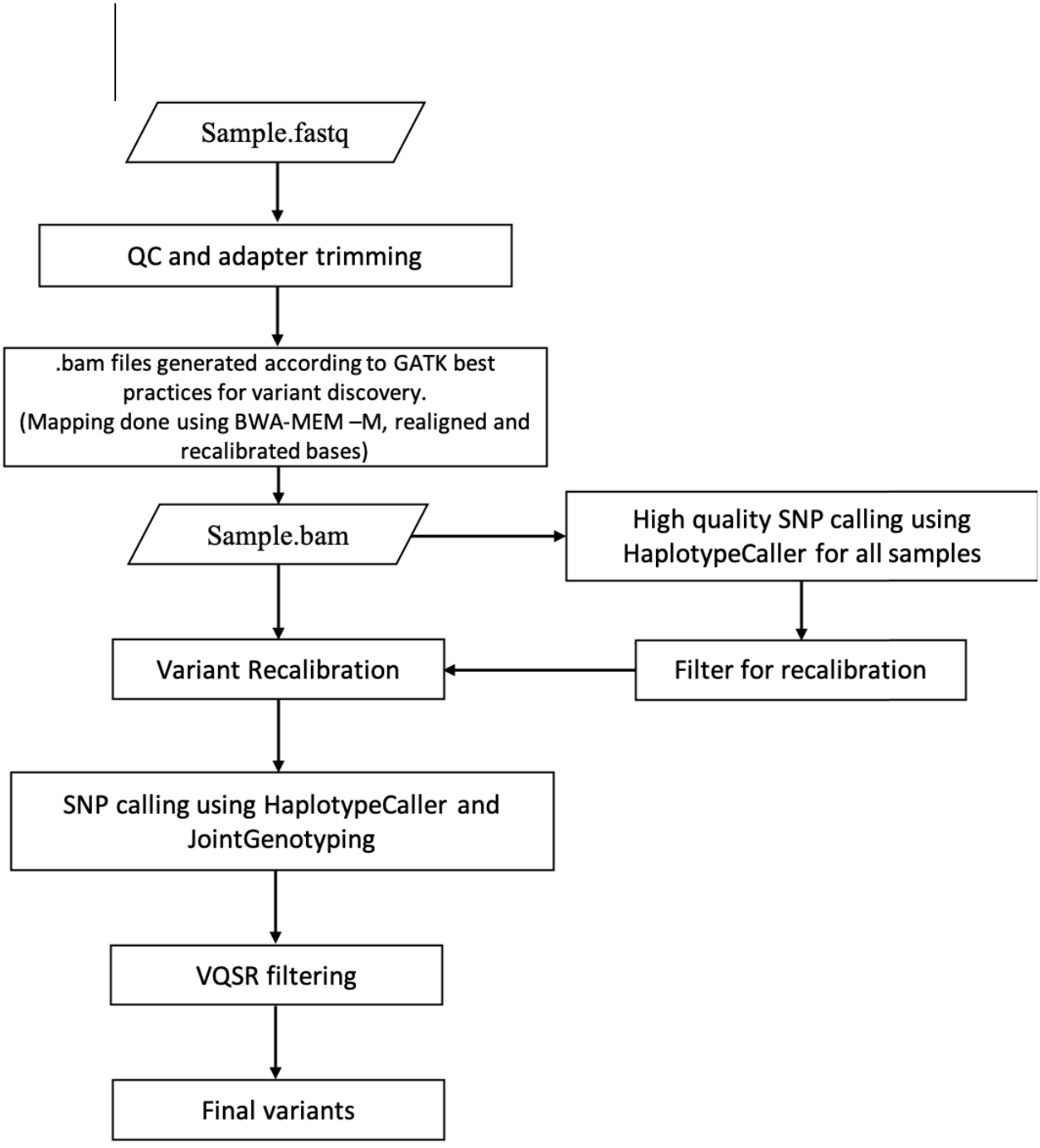


Figure S1

Schematic illustration of the pipeline used for genotyping the samples.

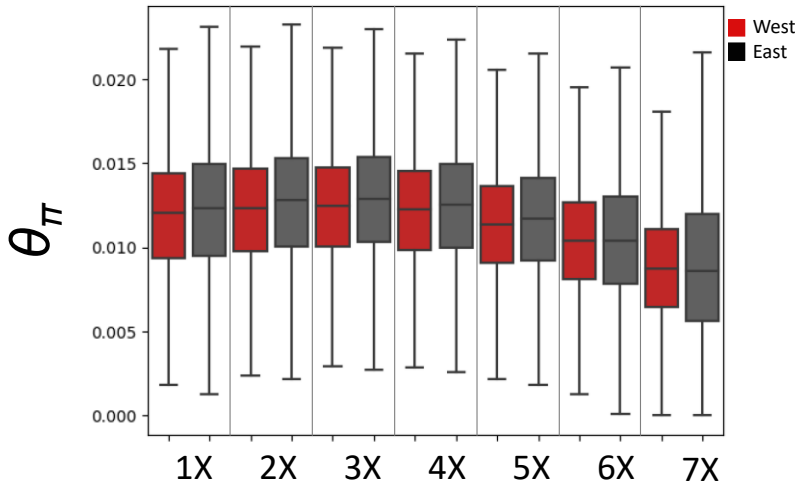


Figure S2

Nucleotide diversity (θ_π) estimates calculated in windows of 10 kb across the genome with coverage filters ranging from 1X to 7X. These filters were applied on polymorphic and non-polymorphic sites respectively.

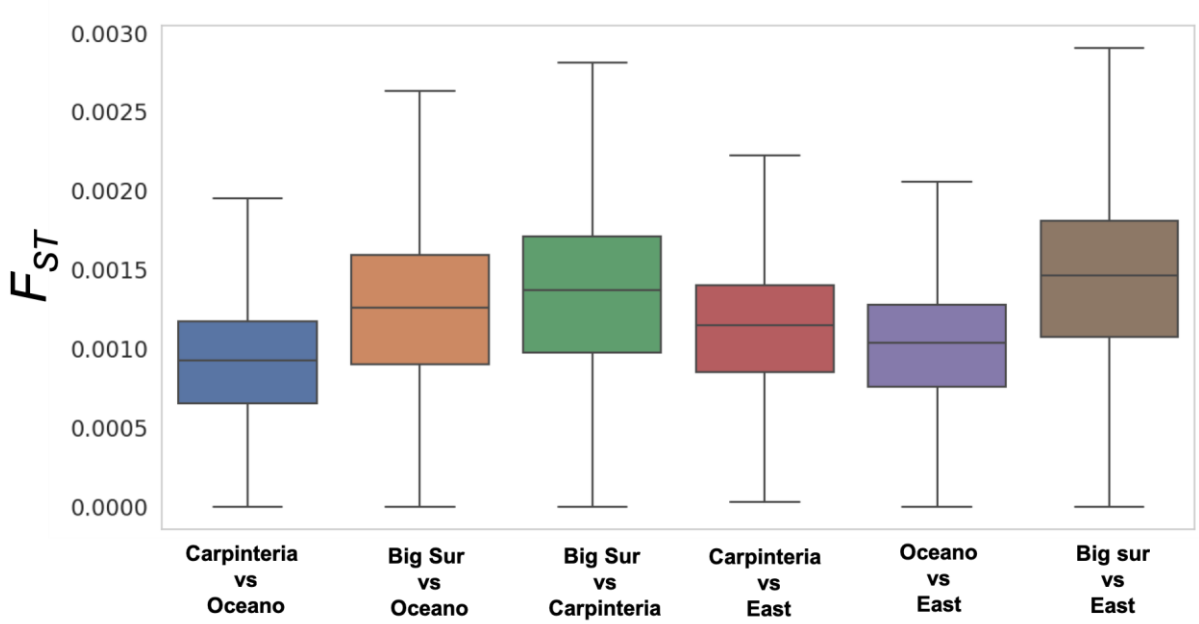


Figure S3

Boxplot illustrating the lack of genetic differentiation between monarchs collected in eastern North America and those collected at different overwintering sites in western North America.

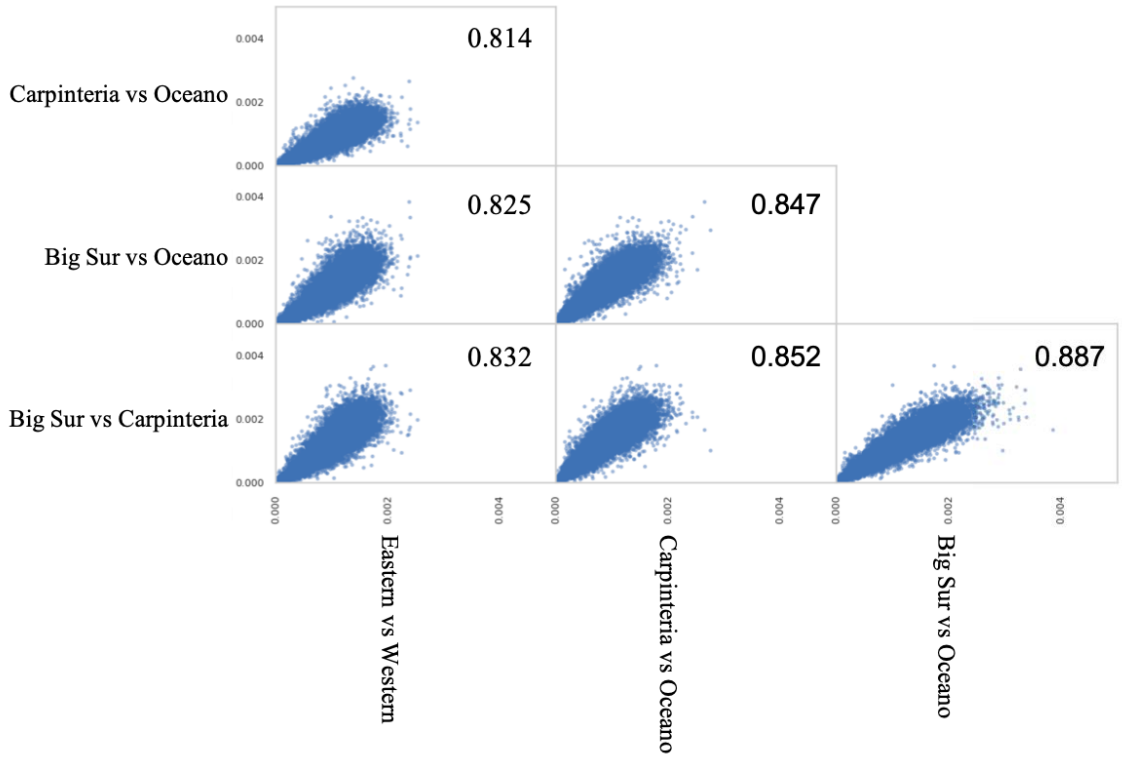


Figure S4

Correlation map of the genome wide genetic differentiation (F_{ST}) landscapes of different sampling locations.

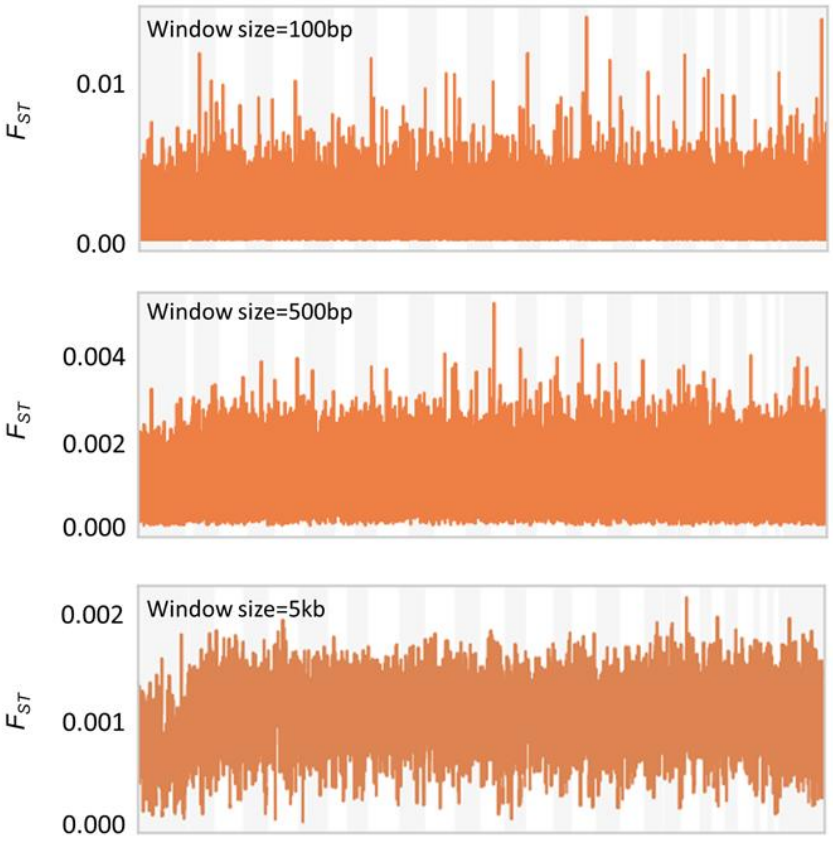


Figure S5

Genetic differentiation (F_{ST}) calculated in non-overlapping windows of 100bp, 500bp and 5,000bp across the genome.

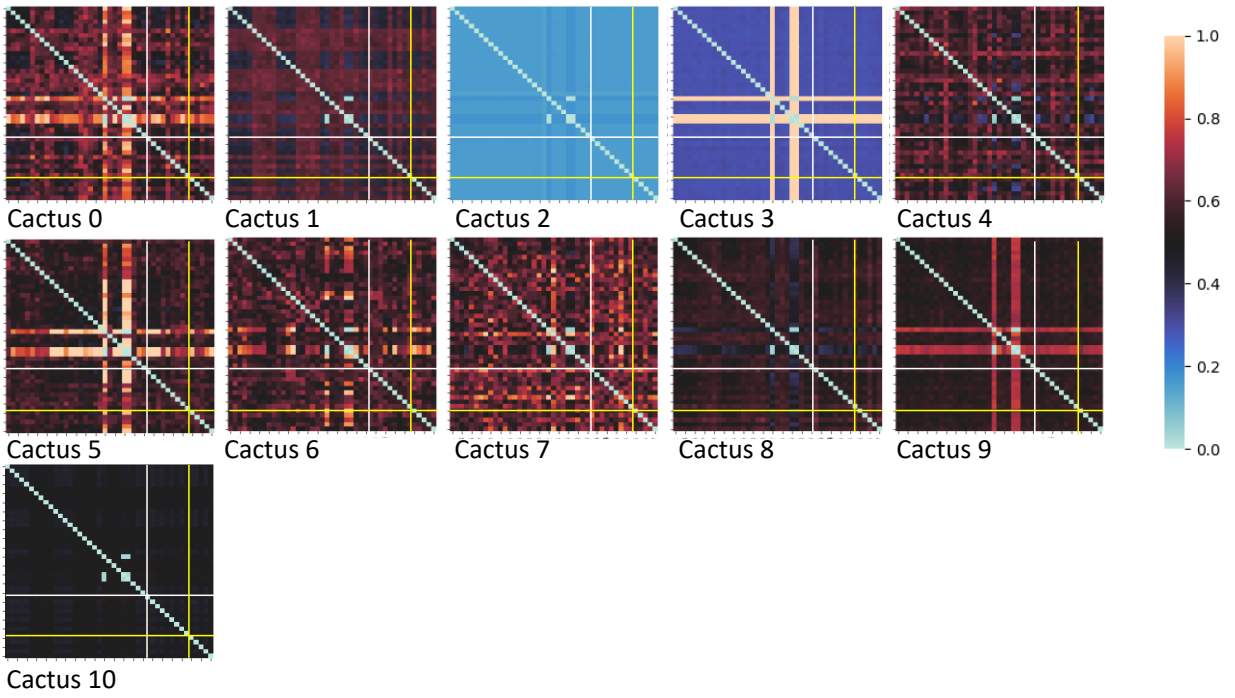


Figure S6

Illustration of phylogenetic relationship matrixes (Cacti) obtained from Saguaro. Both horizontally and vertically the first samples are from the west and then followed by the eastern samples. The horizontal and vertical white lines separate the eastern from the western samples. The yellow line separates the samples from Mexico from other samples. While none of the relationships separated eastern and western monarchs, three samples from Big Sur (visible as the differently shaded bands) appeared to cluster together.

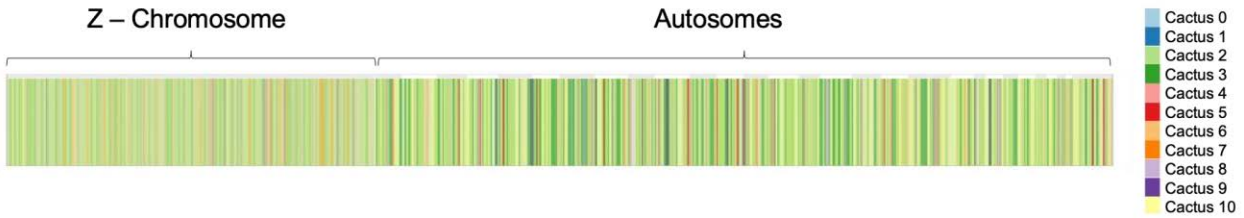


Figure S7

Genome-wide representation of the phylogenetic relationship matrixes (Cacti) obtained from Saguaro. These occurrences are summarized in Table S8. Note: The cacti are plotted based on their occurrences and length was not accounted for.

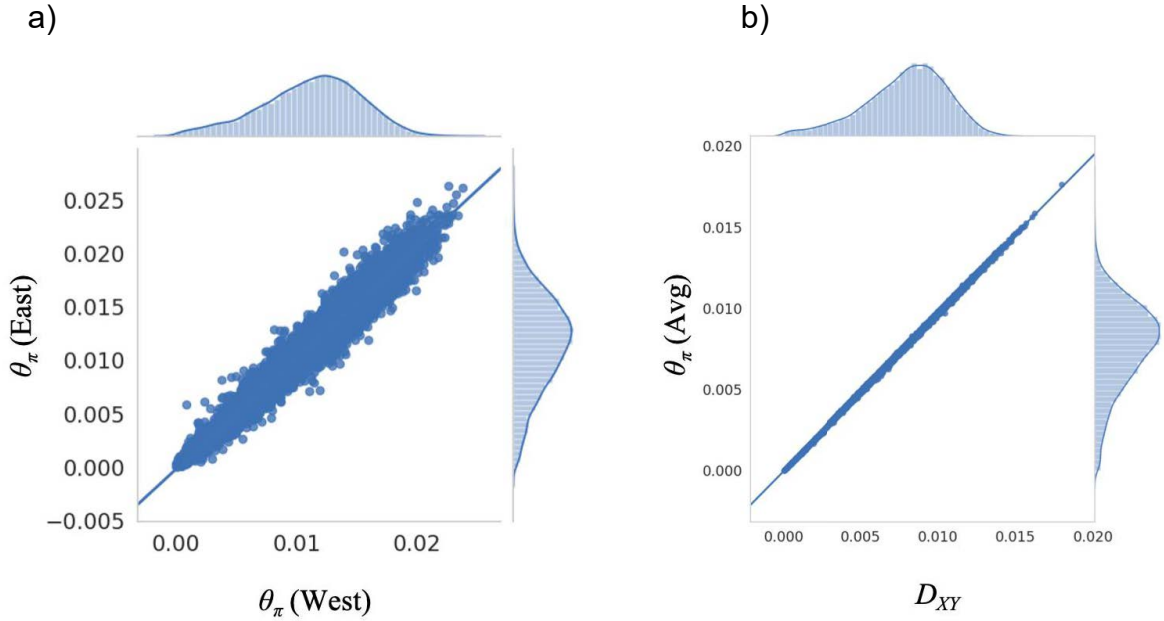


Figure S8

a) Correlation of genome-wide genetic diversity (θ_π) of eastern and western monarchs (Pearson's *r* -value = 0.98, P-value < 0.001). **b**) Linear relationship between average genetic diversity (θ_π) of eastern and western monarchs and the absolute divergence (D_{XY}; Pearson's *r* -value = 1.0 , P-value < 0.001) .

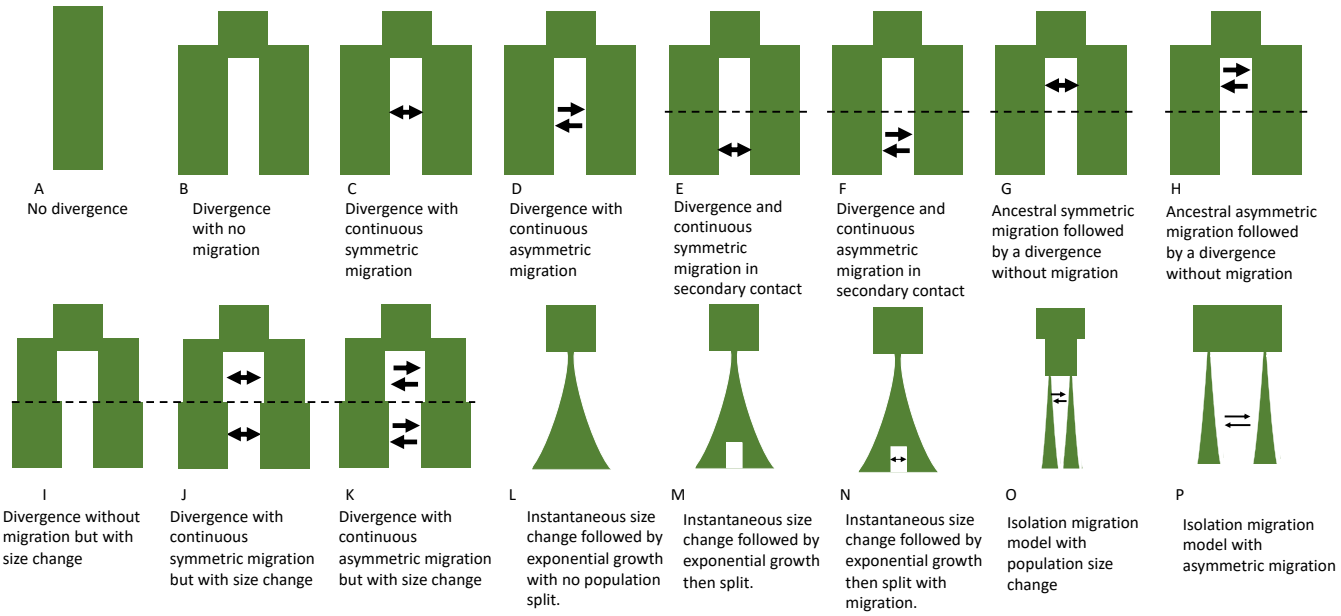


Figure S9

Visual illustrations of models screened using ∂a∂i.

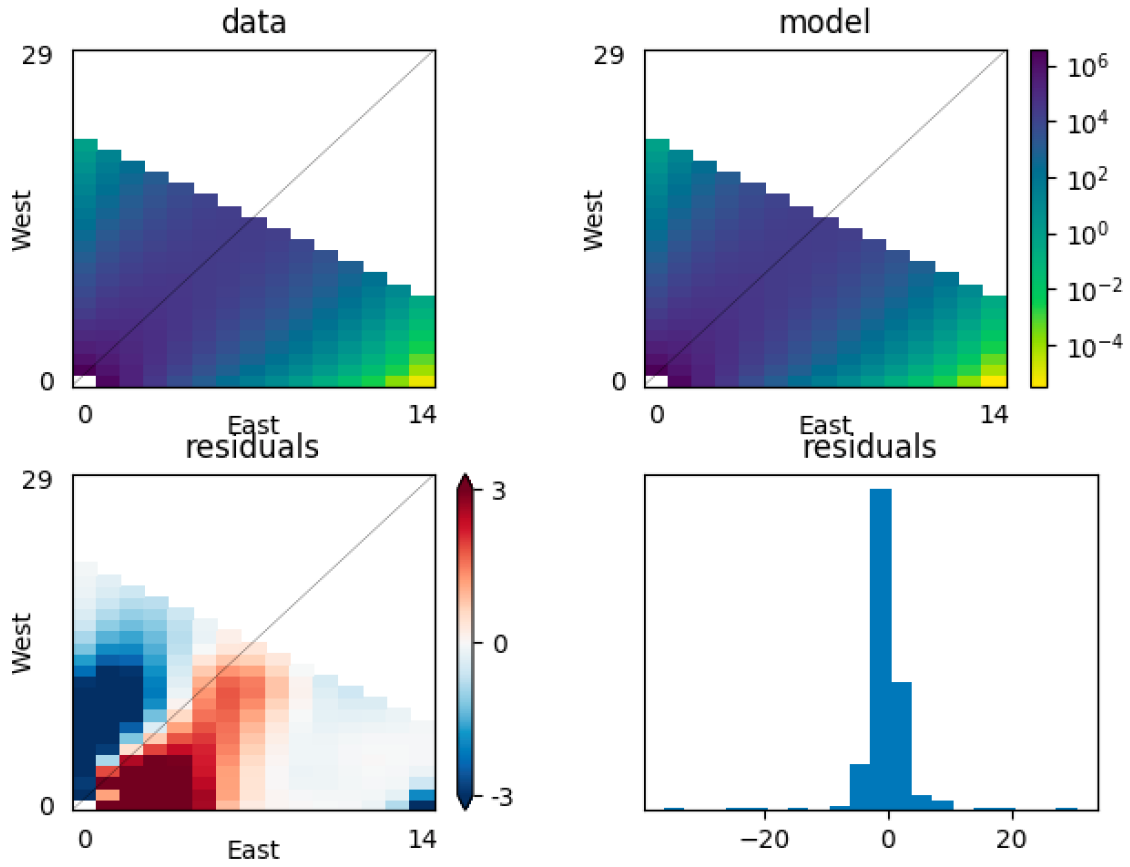


Figure S10

Visual illustration of the 2D-site frequency spectrum (data) of eastern and western monarchs with the model site frequency spectrum obtained from the best optimized model ('Bottlegrowth_split_mig') using ∂a∂i. The bottom graph shows the distribution of residuals of the model.

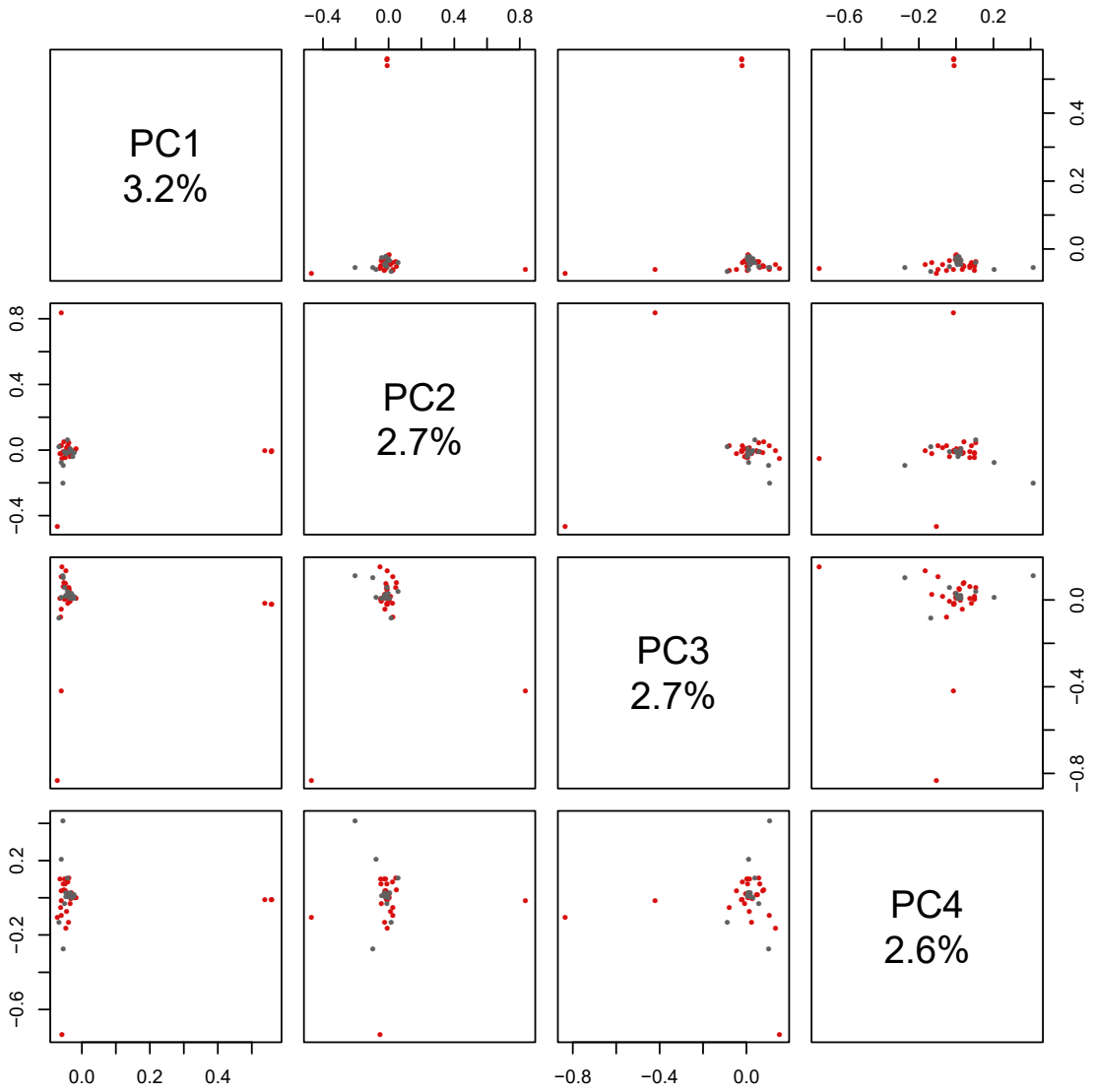


Figure S11

Principal component analysis of top four Eigenvectors for 20.9 million SNPs in the dataset.

Appendix 1

Here we describe the calculations to convert the optimized parameters from the most likely model from demographic history analysis using ∂a∂i.

Obtained parameters from the model.

nuB	nuF	m	T	Ts
1.7745	6.3606	10.9812	0.9189	0.2505

Other stats from the model

θ = 980321.63

L = sum of all the sites used in the site frequency spectrum = 11288846.4

u = mutation rate = $2.9*10^{-9}$ (from *Heliconius melpomene*)

Generation time =0.3

Conversion of parameters into time, effective population size and migration rate

$\theta = 4*Ne_{[ref]}*\mu*L$

$Ne_{[ref]}$ = θ /(4*u*L)= 7,486,193

$Ne_{[ref]}$ is the measure of effective population size that is used to scale the magnitude of the bottleneck.

nuB = nuB (from model parameters) * $Ne_{[ref]}$ = 1.7745*7,486,193= 13,284,249

nuF = nuF (from model parameters) * $Ne_{[ref]}$ = 6.3606*7,486,193= 47,616,680

m = 10.9812/(2*$Ne_{[ref]}$) = 10.9812/(2*7,486,193) = $7.33*10^{-7}$

T = 0.9189*2*$Ne_{[ref]}$*Generation time = 0.9189*2*7,486,193*0.3 = 4127437

Ts = 0.2505*2*$Ne_{[ref]}$* Generation time = 0.2505*2*7,486,193*0.3 = 1125174