

# Isotope geolocation and population genomics in *Vanessa cardui*: Short- and long-distance migrants are genetically undifferentiated

Megan S. Reich <sup>a,\*</sup>, Daria Shipilina <sup>b,\*</sup>, Venkat Talla <sup>b</sup>, Farid Bahleman<sup>c</sup>, Khadim Kébé <sup>d</sup>, Johanna L. Berger <sup>e,f</sup>, Niclas Backström <sup>b</sup>, Gerard Talavera <sup>f,2</sup> and Clément P. Bataille <sup>g,2</sup>

<sup>a</sup>Department of Biology, University of Ottawa, Ottawa, Canada K1N 7N9

<sup>b</sup>Evolutionary Biology Program, Department of Ecology and Genetics, Uppsala University, 752 36 Uppsala, Sweden

<sup>c</sup>SOS Savane—NGO, Tanguéta 145, Benin

<sup>d</sup>LADB, Higher School of Industrial and Biological Engineering, 11000 Dakar, Senegal

<sup>e</sup>Ecological Networks, Technische Universität Darmstadt, 64287 Darmstadt, Germany

<sup>f</sup>Institut Botànic de Barcelona (IBB), CSIC-CMNCB, 08038 Barcelona, Catalonia, Spain

<sup>g</sup>Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, Canada K1N 7N9

\*To whom correspondence should be addressed: Email: [daria.shipilina@ebc.uu.se](mailto:daria.shipilina@ebc.uu.se) (D.S.); Email: [megan.reich@uottawa.ca](mailto:megan.reich@uottawa.ca) (M.S.R.)

<sup>1</sup>M.S.R. and D.S. contributed equally to this work.

<sup>2</sup>Joint senior authors.

Edited By Zhonghe Zhou

## Abstract

The painted lady butterfly *Vanessa cardui* is renowned for its virtually cosmopolitan distribution and the remarkable long-distance migrations as part of its annual, multigenerational migratory cycle. In winter, *V. cardui* individuals inhabit breeding grounds north and south of the Sahara, suggesting distinct migratory behaviors within the species as individuals migrate southward from Europe in the autumn. However, the evolutionary and ecological factors shaping these differences in migratory behavior remain largely unexplored. Here, we performed whole-genome resequencing and analyzed the hydrogen and strontium isotopes of 40 *V. cardui* individuals simultaneously collected in the autumn from regions both north and south of the Sahara. Our investigation revealed two main migratory groups: (i) short-distance migrants, journeying from temperate Europe to the circum-Mediterranean region and (ii) long-distance migrants, originating from Europe, crossing the Mediterranean Sea and Sahara, and reaching West Africa, covering up to over 4,000 km. Despite these stark differences in migration distance, a genome-wide analysis revealed that short- and long-distance migrants belong to a single intercontinental panmictic population extending from northern Europe to sub-Saharan Africa. Contrary to common biogeographic patterns, the Sahara is not a catalyst for population structuring in this species. No significant genetic differentiation or signs of adaptation and selection were observed between the two migratory phenotypes. Nonetheless, two individuals, who were early arrivals to West Africa covering longer migration distances, exhibited some genetic differentiation. The lack of genetic structure between short- and long-distance migrants suggests that migration distance in *V. cardui* is a plastic response to environmental conditions.

**Keywords:** painted lady butterfly *Vanessa cardui*, isotope geolocation, insect migration, whole-genome resequencing, panmixia

## Significance Statement

Although migratory insects dominate living biomass fluxes and impact agriculture, ecosystems, and human communities, little is known about the controls of their migratory behavior. Our study develops an interdisciplinary framework, applied to the migratory butterfly *Vanessa cardui*, to combine isotope-based phenotyping with whole-genome resequencing data to explore the genetic basis of variation in insect migration behavior. We leverage new-generation isotope geolocation techniques to uncover striking differences in butterfly behavior, with some individuals migrating short distances to the circum-Mediterranean region and others migrating thousands of kilometers across the Mediterranean Sea and Sahara. This major difference does not coincide with genetic differentiation or population structure and is likely a plastic response to environmental cues.

## Introduction

Migration has evolved multiple times across taxa (1) and is often viewed as a complex adaptation allowing animals to move to

favorable environmental conditions, exploit seasonal resources, and escape parasites and predators (2). Many migratory animals show intraspecific variation in migratory behavior that results in

**Competing Interest:** The authors declare no competing interests.

**Received:** June 13, 2024. **Accepted:** December 16, 2024

© The Author(s) 2025. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

distinctive spatiotemporal movement patterns. These patterns are often associated with phenotypic traits, including migratory propensity, timing, orientation, and morphological features such as wing shape and size (3). Many studies, mainly focused on birds, have contributed to our understanding of the genetic basis of migration by comparing populations with different migratory phenotypes (4, 5). For example, intraspecific variation in migratory strategies manifests in regional genomic divergence in multiple bird species, such as the European blackcap *Sylvia atricapilla*, Swainson's thrush *Catharus ustulatus*, and willow warbler *Phylloscopus trochilus* (6–9). Similar patterns of variation in migratory phenotypes can be observed in insects, but whether these differences result in similar patterns of adaptation and genetic differentiation remains an open question.

Although the environmental factors and genetic underpinnings driving intraspecific variation in insect migratory behavior are only beginning to be understood (10), studies thus far indicate that these factors may differ from those of birds. A recent study on the monarch butterfly *Danaus plexippus* found a lack of genetic differentiation associated with geographically separated migration routes (11), indicating that intraspecific migratory diversity in insects does not necessarily lead to population structure or differentiation. It is estimated that at least 600 species of butterflies display some form of migratory behavior (12), and, extrapolating to include other insect taxa, the total number of migratory insect species is likely in the tens of thousands. However, few migratory insect species have been studied in detail, as research on insect migration has lagged behind that of birds and mammals due to the challenges of identifying migratory phenotypes in insects (13). Extrinsic markers such as biologgers commonly record migratory trajectories for birds and mammals, but few studies have used these techniques on insects (14–16) because insects are generally too small, short-lived, and numerous for these techniques to be applied on a large scale. Instead, studies on insects have mainly relied on observations of physiological state (e.g. wing wear scores; 17), behavior (e.g. insect-monitoring radar; 18), or regional field monitoring (e.g. spatiotemporal location; 19).

Fortunately, naturally occurring isotopes are intrinsic markers that have proven increasingly useful for estimating the geographic origin of wild-caught insect migrants and have the potential to identify migratory phenotypes. Isotope geolocation relies on the observation that insect larvae ingest the local environmental isotopic composition through their diet and incorporate this isotopic composition into their developing tissues (20, 21). Insect wings are relatively inactive tissues and largely preserve the isotopic composition of the natal environment (22, 23). Thus, the isotopic composition of the wing of a wild-caught insect can be measured and then compared with a spatial model of isotopic variation (i.e. an isoscape) to estimate the insect's natal origin, a process known as geographic assignment. However, single isotope-based geographic assignment generally yields broad and unspecific regions of natal origin, making it challenging to calculate downstream estimates, such as migration distance. More spatially constrained estimates of natal origin can be obtained by combining multiple isotopes with complementary patterns (24). Dual isotope-based geographic assignment combining hydrogen isotope values ( $\delta^2\text{H}$ ), driven by climatic variables (e.g. precipitation; 25, 21), and strontium isotope ratios ( $^{87}\text{Sr}/^{86}\text{Sr}$ ), primarily driven by geological processes (e.g. bedrock type and age; 26), have previously been used to classify migratory behaviors in birds and bats (e.g. 27, 28). Likewise, isotopes can be used to phenotype insect migratory behavior, which is often challenging to observe directly. When combined with genetic techniques, this approach provides

an effective tool for investigating the potential genomic basis of variation in insect migratory behavior.

Here, we combined isotope-based phenotyping of migration distance with whole-genome resequencing data to investigate intraspecific variation in migratory behavior for the painted lady butterfly *Vanessa cardui*. *Vanessa cardui* is renowned for its obligate, annual, multigenerational migratory cycles throughout Africa, Asia, Europe, and North America (29, 30). Since this species follows the oogenesis-flight syndrome, wherein there is an energetic trade-off between reproduction and migration (31–33), migration occurs during the prereproductive phase. Recently, crossings of the Sahara by *V. cardui* have been evidenced through isotope geolocation, pollen metabarcoding analysis, ecological niche models, and monitoring data, revealing a migratory cycle covering between 8 and 10 generations and involving both the Palearctic and Afrotropical regions. These recent findings show that the winter breeding range of *V. cardui* in the western Afro-Palearctic is made up of two regions separated by the Sahara: sub-Saharan Africa and the circum-Mediterranean region (19, 34). While some autumn butterflies migrate southward from temperate Europe to breed in the circum-Mediterranean region (34, 35), a more substantial portion migrate across the Mediterranean Sea and Sahara to subsequently breed in sub-Saharan Africa (19, 36, 37). However, our understanding of this extreme migratory journey across the Sahara is incomplete without genomic data. The substantial distance between these two wintering regions suggests the species engages in two distinct migratory strategies, leading us to explore different hypothetical scenarios. First, individuals wintering north and south of the Sahara could belong to a single panmictic population. In this scenario, nongenetic factors, such as plastic responses to environmental cues, may explain intraspecific differences in migratory distance. Alternatively, if individuals wintering north and south of Sahara constitute distinct groups with limited gene flow, some level of genetic structure might be observed across the Sahara, mainly as a consequence of (i) adaptive processes (e.g. adaptation to long-distance or minimal migration) and/or (ii) neutral processes (e.g. genetic drift, gene flow, isolation, and subsequent divergence in isolated populations).

In this study, we estimate the natal origin and migration distance of 40 adult *V. cardui* synchronically collected from north and south of the Sahara in the autumn by coupling  $\delta^2\text{H}$  and  $^{87}\text{Sr}/^{86}\text{Sr}$  into a dual geographic assignment framework. Next, a genome-wide resequencing approach is used to investigate potential genetic structuring associated with contrasting migratory phenotypes. By understanding the underlying factors shaping the remarkable trans-Saharan migratory journeys of *V. cardui*, this study seeks to contribute to our understanding of the ecology and evolution of insect migration.

## Results

### Isotope-based phenotyping reveals differences in migration distance

We sampled 40 *V. cardui* individuals from sites within the disjunct winter breeding grounds: sub-Saharan Africa (i.e. Benin, Senegal) and the circum-Mediterranean region (i.e. Morocco, Spain, Portugal, and Malta). Specimens were sampled during the time when monitoring data indicates that immigrants are arriving to these previously unoccupied regions: from August to October in sub-Saharan Africa and late September/early October to November in the circum-Mediterranean region (19). The sampled

butterflies exhibited reproductive behavior (i.e. hilltopping, oviposition; see Data Availability), indicating that they had likely completed their migratory phase (33). Next, we estimated the natal origins of the 40 *V. cardui* using  $\delta^2\text{H}$  and  $^{87}\text{Sr}/^{86}\text{Sr}$  following the geographic assignment framework of Reich et al. (24) (Table S1; Fig. 1D). From the posterior probability surfaces, we then calculated a conservative estimate of migration distance (i.e. minimum distance; Figs. S1 and 1C).

Based on the estimates of migration distance, all individuals were grouped into three categories: (i) long-distance migrants ( $n=13$ , 33%; with minimum migration distances ranging from  $>500$  to  $>4,000$  km), for which we could confidently infer long-distance migration; (ii) short-distance migrants ( $n=3$ , 7%; with minimum migration distances ranging from  $>140$  to  $>240$  km), for which we could confidently infer short-distance migration; and (iii) individuals of putatively local origin ( $n=24$ , 60%; having potentially travelled  $<100$  km), for whom our estimates could not definitively distinguish between **local and short-distance or long-distance migration**. Isotopic compositions corresponding to long-distance migrations were found exclusively in individuals captured south of the Sahara, specifically in Senegal and Benin. Based on the summarized posterior probability surfaces, these individuals seemed to originate either from northwestern Africa or Europe, notably from Portugal, Spain, France, or Sweden (Fig. 1B). However, some of the individuals ( $n=6$ , representing 32% of the samples from south of the Sahara; Fig. 1E) had a high probability of originating from areas near their capture points ( $<100$  km) and were designated as “putative locals” (i.e. having an isotopic composition compatible with being the offspring of early immigrants or of regional breeders, which can be present at low abundances). Notably, no individuals from south of the Sahara were classified in the short-distance migration category.

Among the samples collected north of the Sahara, some belonged to the short-distance migration category ( $n=3$ , 14%; Fig. 1A). However, the majority of these samples exhibited isotopic compositions consistent with local origin ( $n=18$ , 86%; Fig. 1E). The isotope-based analysis does not clarify whether this local composition indicates (i) premigratory, locally sourced offspring of early immigrants to the region or (ii) immigrants from a distant region with an isotopic composition similar to their capture location. We also considered natural history observations to interpret these patterns. The majority of individuals collected in the circum-Mediterranean region during October and November are likely the descendants of earlier arrivals to the collection area. This inference is substantiated by the favorable breeding conditions in the region during late September and October (19, 34, 37), as well as the observed presence of eggs, larvae, and recently emerged adults (i.e. expelling meconium) at the time of collection in Morocco, Malta, and Portugal. Thus, we categorize any samples with a high probability of originating near their capture location as putative locals.

To assess the genomic differentiation between migratory phenotypes, it is crucial to integrate the isotopic data with spatiotemporal information from individual samples. Consequently, for subsequent analysis, we grouped individuals that travelled long distances from Senegal and Benin, explicitly excluding individuals with putatively local origins, into the “long-distance” group. To balance sample sizes, we grouped individuals from north of the Sahara, encompassing both the short-distance ( $n=3$ ) and putatively local categories ( $n=18$ ), into the “short-distance” group, operating under the assumption that these putative locals are the next-generation offspring of short-distance migrants.

## Wing morphology and wear are uncorrelated with migration distance

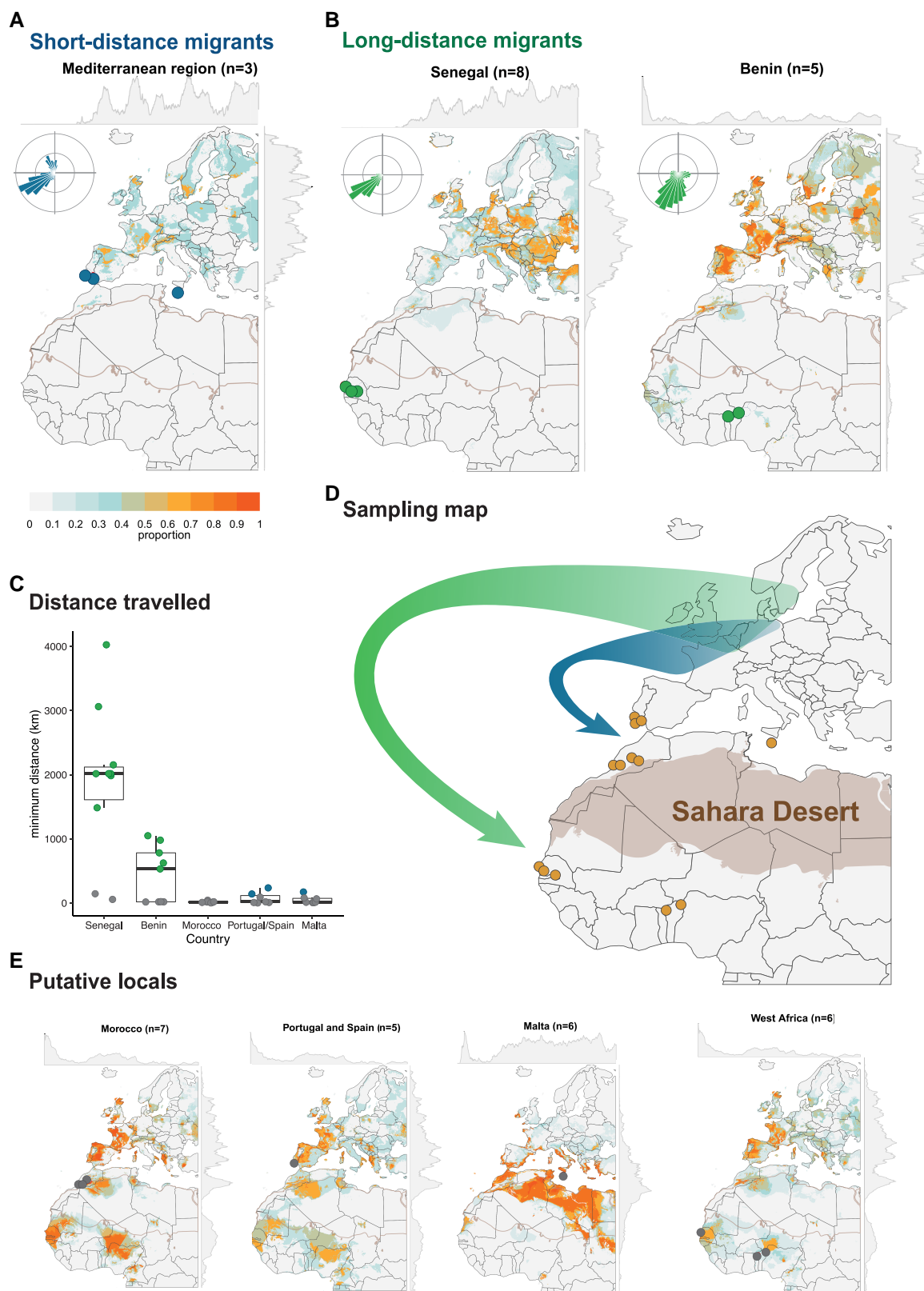
Wing wear scores, a metric scaled from 1 (fresh) to 5 (extremely worn), are often used as a proxy for butterfly age and migratory status (17, 36, 38). In our study, we preferentially sampled butterflies with worn wings (i.e. higher wing wear scores) as it is assumed that high wing wear scores correlate to postmigratory status; thus, wing wear scores ranged from 2 (few scales lost) to 5 (extremely worn). These wing wear scores showed no correlation with migration distance (Spearman rank correlation,  $\rho = -0.06$ ,  $P = 0.70$ ).

Butterfly wing size and shape are sometimes correlated with migratory ability (e.g. 39). Here, wing shape was uncorrelated with migration distance ( $P = 0.2$ ) or the country where the sample was collected ( $P = 0.1$ ). There was a near-significant relationship between wing shape and wing size ( $R^2 = 0.30$ ;  $F_{1,27} = 1.8$ ,  $P = 0.07$ ). Wing shape differed between the sexes ( $F_{1,27} = 3.4$ ,  $P = 0.0002$ ), with the wings of females being more disparate in shape than those of males (Fig. S2; pairwise comparison,  $P = 0.03$ ). Despite the differences in wing shape between males and females, we did not detect any difference in migration distance by sex (Wilcoxon test;  $W = 127$ ,  $P = 0.47$ ). Additionally, wing shape was influenced by the date of collection, with earlier captures having smaller and more elongated wings (Fig. S2;  $F_{1,27} = 3.0$ ,  $P = 0.001$ ).

## Lack of population structure across the Sahara

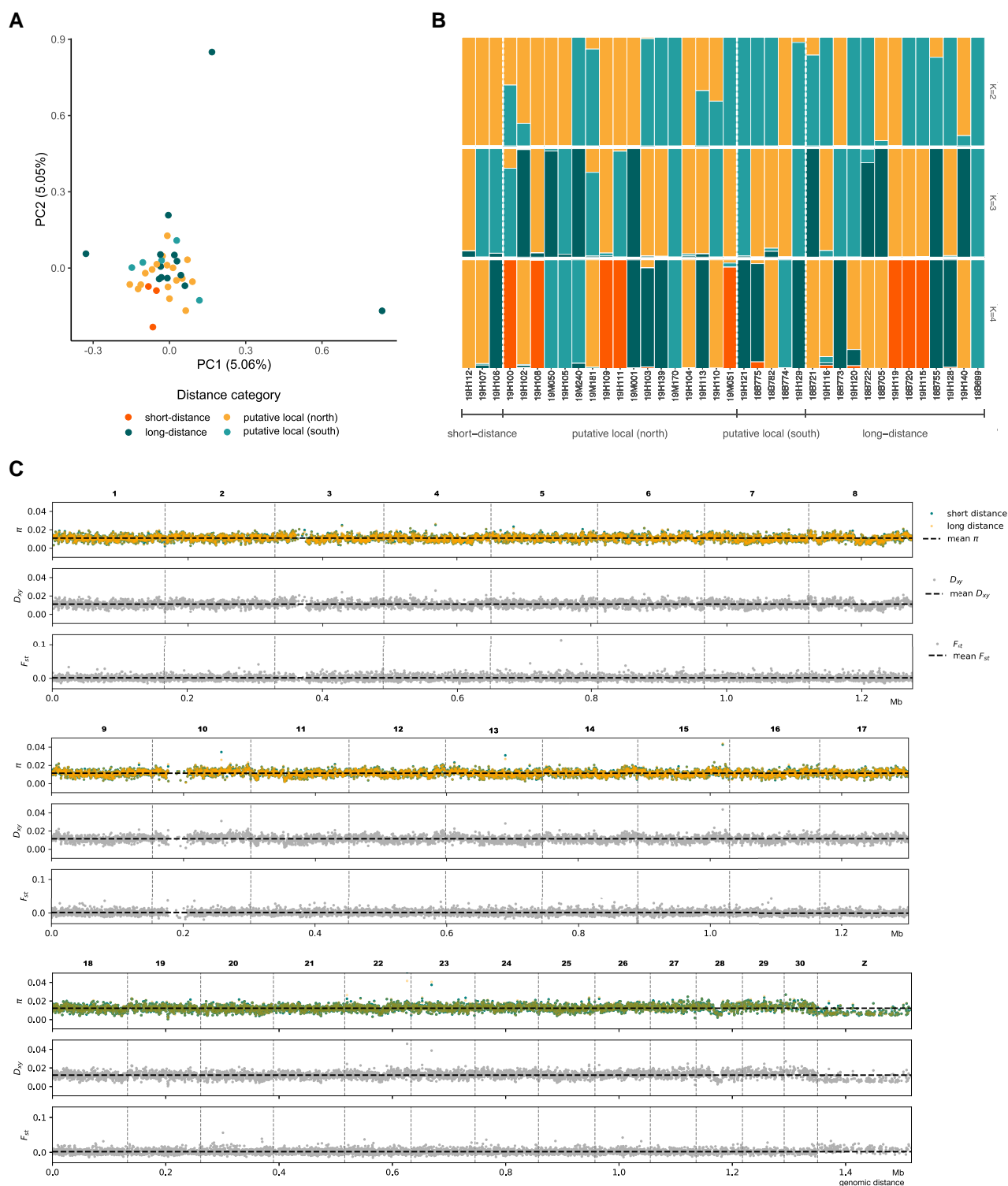
We explored whether the distinct differences in migration distances, coupled with the evident spatiotemporal separation of short- and long-distance migrants, might be reflected in genetic differences between them, which in turn can be an indicator of population structure driven by neutral processes or as early signs of adaptive divergence among migratory phenotypes. We assessed these patterns at the whole-genome level, employing a comprehensive approach that integrated summary statistics, principal component analysis (PCA), and admixture proportion analysis. All approaches yielded consistent results: a lack of detectable population structure within the dataset and negligible differentiation between groups classified as short- and long-distance migrants.

First, using 13,215,545 high-quality callable sites from both coding and noncoding regions (38 individuals, 2 removed due to quality issues), we summarized levels of nucleotide diversity ( $\pi$ ) within groups and genetic divergence ( $d_{\text{XV}}$ ) between groups of *V. cardui* individuals with distinct migratory behavior (short-distance migrants:  $n=21$ ; long-distance migrants:  $n=13$ ), and calculated the fixation index ( $F_{\text{ST}}$ ) as a measure of differentiation (Table S2). We observed relatively high mean values of nucleotide diversity in both short- and long-distance migrants ( $\pi_{\text{autosome}} = 0.012 \pm 0.003$  [1 SD]) compared with other Lepidoptera (40). Levels of the pairwise nucleotide divergence ( $d_{\text{XYautosome}} = 0.012 \pm 0.003$ ) between two phenotypes were elevated, reflecting the effect of high levels of diversity within samples and stochasticity of the estimates. The level of genetic differentiation between autosomes was low, but significant (pairwise  $F_{\text{ST}} = 0.001 \pm 0.006$ ,  $P < 0.001$ ). A slightly higher fixation index was observed for the Z chromosome (pairwise  $F_{\text{ST}} = 0.009 \pm 0.016$ ), which is expected given that the smaller effective population size makes the Z chromosome more susceptible to the effects of genetic drift. We repeated this analysis using groups defined based on capture location (i.e. north vs. south of the Sahara), but found no substantial differences between summary statistics (Table S2). These findings indicate that *V. cardui* that migrate short- vs. long-distances have limited genetic



**Fig. 1.** Summarized posterior probability surfaces and subsequent estimates of migration distance (i.e. minimum distance) and direction based on dual  $\delta^2\text{H}$ - and  $^{87}\text{Sr}/^{86}\text{Sr}$ -based geographic assignment. A, B, E) Summary maps depict the proportion of individuals with a high probability of natal origin in a given area, as defined by the 2:1 odds ratio. The side margins display the mean proportion by latitude and longitude, and only one circle per capture location is shown. The inset rose plot illustrates the average probability-weighted bearing from the estimated natal origin (center of plot) to the capture locations. Maps summarize the results for A) short-distance migrants, B) long-distance migrants, and E) putative locals captured in the circum-Mediterranean region and West Africa. C) Box plot displaying the estimated migration distance (km) categorized by the country of capture. Individuals with a minimum distance of <100 km were considered putative locals. D) Map indicating the capture locations both north (Morocco, Portugal, Spain, and Malta) and south (Senegal and Benin) of the Sahara. Butterflies captured south of the Sahara generally migrated longer distances (large green arrow) compared with those captured north of the Sahara (small blue arrow).





**Fig. 2.** Lack of population structure and genetic differentiation between short- and long-distance migrants. A) PCA biplot illustrating the genetic variation in the samples. B) Genomic admixture patterns based on assumptions of 2 to 4 clusters ( $K = 2$  to 4). No discernible structure was observed, with the optimal model demonstrating  $K = 1$  (not visually depicted). C) Genome scans showing genetic diversity within each group ( $\pi$ ), genetic differentiation ( $D_{xy}$ ), and fixation index ( $F_{st}$ ) between individuals categorized as short- and long-distance migrants. The analysis was performed with sliding, nonoverlapping windows of 20 kb. Mean values are indicated by a black dashed line.

differences and share similar distribution of allele frequencies, and that the effective population size in general is large.

Secondly, we used PCA to investigate potential genetic structuring. After additional filtering steps, we narrowed down our dataset to 813,810 high-quality single-nucleotide polymorphisms

(SNPs). This analysis revealed that the majority of samples clustered together with no separation between short- and long-distance migrants (Fig. 2A), suggesting an overall absence of population structure. However, within this general pattern, two individuals stood out as outliers in the PCA; one showed

differentiation primarily along the first principal component (i.e. 19H128) and the other exhibited differentiation along the second principal component (i.e. 19H115). Both of these outlier samples were collected in September and were among the earliest arrivals to sub-Saharan Africa (19; Table S1). Moreover, these individuals had undertaken particularly long migratory journeys, with one covering over 2,000 km across the Sahara from northwest Africa or Europe (19H115) and the other over 1,000 km from Portugal, France, Sweden, or Finland (19H128; Fig. S3). The clear separation of these individuals in the PCA biplot could be indicative of population structuring. We explored the contribution of specific SNPs to the first two principal components with *pcadapt* and identified 924 outlier SNPs that made significant contributions to the observed segregation of individuals. However, it is essential to note that drawing conclusions about the potential processes underlying the observed differences based on data from just two individuals is inherently limited. Nevertheless, we have provided a list of candidate SNPs that may serve as a valuable starting point for future investigations (see Data Availability).

Finally, we tested for population structure in our dataset using two different software implementations to infer individual ancestry proportions (admixture analysis). Specifically, we employed the sparse nonnegative matrix factorization (SNMF) clustering algorithm and fastSTRUCTURE to test models assuming 1 to 5 subpopulations. In both cases, the best model was  $K = 1$ , indicating that the data are best described by a single panmictic population (Fig. 2B).

### No islands of genetic differentiation between short- and long-distance migrants

The assessment of genetic structure using genome-wide averages may obscure regional differences in genetic differentiation. We therefore used window-based analysis of genetic differentiation between groups (i.e.  $F_{ST}$  scans) based on two sample groupings: (i) migratory phenotype (i.e. short- and long-distance migrants) and (ii) capture location (i.e. north and south of the Sahara). Both comparisons revealed uniform differentiation landscapes characterized by a low average pairwise  $F_{ST}$ . For autosomal regions, the average  $F_{ST}$  was 0.001 and regional peaks had a maximum  $F_{ST}$  of 0.113 (Figs. 2C and S4). For the Z chromosome, the average pairwise  $F_{ST}$  was 0.009 with a maximum peak of 0.072. This level of variation is consistent with stochasticity of the normalization and/or neutral variation, as has previously been demonstrated in simulation studies (41) following simulation regimes applicable to our system. Additionally, none of the local  $F_{ST}$  peaks were accompanied by a simultaneous elevation of the genetic divergence and drop in diversity within each group extended over a substantial region, therefore not meeting the expectations for signatures of divergent selection.

## Discussion

### Contemporary evidence of variation in migratory behavior

Monitoring data and ecological niche modeling of *V. cardui*'s seasonal distributions allowed for hypothesis-driven sampling both north and south of the Sahara in the autumn, targeting butterflies arriving to the disjunct winter breeding grounds (19, 34, 35, 37). Isotope-based estimates of migration distance from these samples were then used to confirm and quantify the natural variation in migration distance exhibited by *V. cardui*. Our findings reveal that butterflies captured in Senegal and Benin during the autumn likely embarked on long migratory journeys, covering up to over 4,000 km as they crossed the

Mediterranean Sea and Sahara from Europe. In contrast, *V. cardui* individuals captured in the circum-Mediterranean region in the same season either had undertaken shorter migrations or had isotopic compositions compatible with their capture area (Fig. 1). We have expanded this isotope-based evidence with the addition of  $^{87}\text{Sr}/^{86}\text{Sr}$ , narrowing down the potential geographical sources of these long-distance migrants and confirming that distinctive migration strategies are displayed by *V. cardui* found north and south of the Sahara during the autumn months.

In our study, we found that estimates of migration distance cannot be explained by wing size, wing shape, or sex. In this regard, *V. cardui* differs from *D. plexippus*, which show strong associations between migratory ability and wing morphology in that larger and elongated wings are associated with the increased flight potential of the "super generation" during autumn migration (39, 42, 43). The variation in migration distance detected in our study appears to have a basis in behavior rather than wing morphology, which may indicate limited selection on migration distance, as behavior is thought to be more plastic than morphology (44).

We additionally found no significant association between wing wear score and migration distance. Although wing wear scores have long been used as a proxy for butterfly age and migratory status, with high wing wear scores corresponding to older butterflies that have completed their migratory journeys (e.g. 17, 24, 45), our results further illustrate that this proxy is highly unreliable (42). For example, the butterfly with the highest estimated migration distance, up to over 4,000 km (i.e. 18B721), had a low wing wear score of 2 (i.e. "few scales lost"; 38), which is often interpreted as belonging to a recently emerged adult. This finding is in agreement with Korkmaz et al. (46), who conjecture that high-altitude, gliding flight during migration results in less wing damage than low-altitude foraging and breeding behavior. Thus, we suggest that while a wing wear score of 1 can distinguish recently emerged, premigratory individuals, wing wear scores over 2 do not have a clear correlation with migratory distance.

### Trans-Saharan panmixia

Our genome-wide analysis revealed no evidence of genetic divergence between *V. cardui* individuals migrating to the north or south of the Sahara during autumn (Fig. 2). Consequently, the *V. cardui* population stretching from Northern Europe to the equatorial regions of sub-Saharan Africa can be characterized as regionally panmictic, confirming the spatial extent of the Afro-Palearctic migratory range. Our findings align well with previous studies on the genetic structure of Palearctic *V. cardui* using mitochondrial DNA markers (47–49). Given that large biogeographic barriers, such as the Mediterranean Sea and Sahara, do not seem to lead to population structure, it is likely that panmixia will be found at even larger scales. However, *V. cardui* migration seems to generally follow latitudinal trajectories, and extending to larger geographical scales might reveal longitudinal structure and population differentiation (e.g. 50). Future studies will need to assess the extent to which global panmixia occurs across the entire distributional range of *V. cardui*, preferably by including genome-wide genetic markers from around the world. These future studies should include remote islands, which sometimes host isolated populations of highly mobile insect species that are phenotypically and genetically differentiated (e.g. 51).

Special attention should be given to understanding how selection and local adaptation act on migratory behavior in insects, as well as the methods for detecting them. In our study, although we

did not detect clear signatures of selection or adaptation, we did detect two outlier samples in the PCA. There could be multiple explanations for the patterns we detected. First, unlike migratory vertebrates that complete round-trip migrations within a single generation, insect migratory cycles are multigenerational. Multigenerational migratory insects exhibit a complex reticular movement pattern driven by the drastic changes in the location and size of the suitable breeding habitat over the annual migratory cycle (19, 34, 37). Therefore, polygenic within-generation selection due to spatially and temporally varying selection pressures could lead to allele frequency changes for a portion of the population (52). However, the same alleles selected for in one generation are unlikely to be favored by selection pressures in the subsequent generation because of spatiotemporal differences in abiotic and biotic conditions during the migratory journey and in the new breeding habitat, thus diluting the effect of within-generation local selection. Similar frameworks have been proposed for some fish species which show signs of within-generation local selection within panmictic populations (52–54). A similar scenario has been outlined for *D. plexippus*, which experiences selection pressures on wing shape and wing size during long-distance autumn migrations, only to have these pressures released in subsequent generations (42).

Secondly, it is important to recognize that the exceptionally large effective population size of *V. cardui* (55) and, consequently, weak genetic drift could impede our ability to detect early, or complete, reproductive isolation between short- and long-distance migrants. In these circumstances, many thousands of generations of reproductive isolation would need to elapse before our analysis would be able to detect significant allele frequency shifts. Therefore, it is possible that reproductive isolation has indeed occurred but remains undetected. Thirdly, it is conceivable that our sampling strategy was unable to detect genetic differentiation between groups of migrants. Notably, the most genetically differentiated migrants (Fig. 2A) were on the leading edge of arrivals to West Africa (September 2019) (19) and likely originated in temperate Europe (Fig. S3). These early arrivals may have been exposed to strong environmental cues in temperate Europe, where seasonal changes occur earlier than in regions closer to the equator, prompting the butterflies to initiate their migrations earlier in the season. Early migrants of *D. plexippus* have been shown to have advantages over their later counterparts, such as enhanced immune defenses (56). Consequently, it is possible that *V. cardui* segregates temporally, rather than by distance. However, it is more plausible that this temporal pattern results from within-generation selection, which is likely to be diluted in the subsequent generation. Despite any ephemeral within-generation local selection, we anticipate that the capacity to tolerate heterogeneous environments relies primarily on phenotypic plasticity (57). Ultimately, interpreting our results as indicative of regional panmixia remains the most parsimonious approach.

Continental-scale panmixia has been demonstrated in other long-distance migratory insect species. For example, whole-genome resequencing data of *D. plexippus* revealed high gene flow and low genetic differentiation between butterflies located east and west of the Rocky Mountains, despite following geographically separated migration routes (11). Regional panmixia has also been suggested through microsatellite-based population genetics for the green darner dragonfly *Anax junius* in eastern North America (58, 59), and global panmixia has been inferred through mitochondrial DNA analysis for the wandering glider dragonfly *Pantala flavescens* (60, 61). Nonetheless, the panmixia described for *V. cardui* across its Afro-Palearctic distributional range, along

with its regular trans-Saharan crossings, are exceptional phenomena for insects and lend additional support to the hypothesis that panmixia across large geographical ranges is a prevalent feature among migratory insect species. The intercontinental panmixia observed in *V. cardui* is particularly striking given that the Palearctic and Afrotropical regions are distinct biogeographic realms with unique faunal compositions, sharing only a relatively small number of species. Based on our conclusions for *V. cardui*, we hypothesize that some of the other insect species shared between the Palearctic and Afrotropical regions, particularly among the Lepidoptera, Odonata, and Orthoptera (62, 63), may also migrate across the Sahara. Consequently, trans-Saharan gene flow may similarly occur in some of these species.

The widespread panmixia observed in many migratory insects raises questions about the life-history strategies promoting panmixia and how this contrasts with migratory vertebrates. We found that neither differences in migration distance nor the potential biogeographic barrier led to isolation, selection, or significant allele frequency differences in any genomic region (Fig. 2C). This finding is in contrast to numerous examples in vertebrate species (e.g. 6, 7, 64, 65) and a few instances in insect species (e.g. 66). However, in the case of many bird species, annual migrations across biogeographic barriers occur within the lifetime of a single individual, and high philopatry can drive population structure and the formation of diversity in migratory behavior (67, 68). In contrast, most insect migrations are multigenerational, featuring complex movement patterns driven by dynamic changes in breeding habitat, promoting intra-population mixing. Additionally, the tendency of long-distance migratory insects to engage in multiple matings, their lack of courtship behavior, and their practice of laying numerous eggs (e.g. 58, 69, 70) hints at a link between panmictic tendencies and mating strategies in migratory insects. This link, while underexplored, is reflected in the very meaning of the term “panmixia.” Large census population sizes in migratory insects, accompanied by intermittent outbreaks, further drive them toward panmixia. Outbreaks increase the number of successful migrants, potentially eroding local adaptation (71). Outbreaks of *V. cardui* from regions with anomalously high vegetation growth (72, 73) may contribute to the observed lack of genetic structure. Similar population dynamics observed in species like *Libytheana carinenta* (74), *A. junius* (75), and some hoverflies (76) hint at outbreaks being pivotal in shaping the population structure of many migratory insects.

The same life-history factors of migratory insects that appear to promote panmixia among migratory insects (i.e. multigenerational migrations, an absence of assortative mating, and large census population sizes with outbreak dynamics) are also associated with large effective population sizes and high genetic diversity. The multigenerational annual migratory cycle results in a greater number of mutations occurring within the population per unit time, compared with species with longer generation times, and large census population sizes are less susceptible to genetic drift, both of which encourage genetic diversity (77). Indeed, our samples showed relatively high genetic diversity ( $\pi_{\text{autosomal}} = 0.012$ ) compared with other butterfly species (40). This observation aligns with the general trend that migratory butterflies tend to possess higher levels of heterozygosity than sedentary species (55). Consequently, this abundant genetic variation provides opportunities for at least some individuals to acclimatize to the heterogeneous environmental conditions that are encountered during their multigenerational migrations, thereby facilitating survival.

Due to the obligate, multigenerational migrations with low philopatry, random mating, and large, outbreaking populations,

local adaptation is implausible within the Afro-Palearctic population of *V. cardui*. We suggest that rather than adapting to short- or long-distance migration, a plastic response to the environment determines the differences we observe in migratory phenotype. It is also possible that the differences in migration distance observed here could also be explained, in whole or in part, by extrinsic mechanisms, including differences in parasite load (78), energy storage and metabolism (79), wind assistance and weather events (80), or even personality (81). However, studies on *D. plexippus* suggest that while migratory state (i.e. resident vs. migratory) is genetically determined (82), variation in migration distance is more likely to be due to many loci of small effect or differential gene expression (11), likely controlled by epigenetic mechanisms (13, 83, 84). Variation in migratory behavior is likely a response to multiple environmental cues, such as resource availability, photoperiod, temperature, humidity, and host plant quality. Recent experiments have demonstrated differences in *V. cardui* gene expression, methylation, and gene regulation due to the density of larval conspecifics, larval food availability, and host plant availability for oviposition (31, 32, 85), suggesting a strong response to environmental cues. The transcriptional basis of differences in migration distance has been explored in the moth *Helicoverpa armigera* and in *D. plexippus*, yielding promising candidate genes (11, 86). Future controlled laboratory studies with *V. cardui* are needed to disentangle the environmental triggers and molecular pathways of phenotypic plasticity in migration distance.

## Conclusion

Here, the combination of isotope-based phenotyping of migration distance and comprehensive whole-genome resequencing provided a unique opportunity to investigate phenotype–genotype associations of migratory behavior in insects, a challenging endeavor until now. We found distinct differences in migration distance, with most individuals captured in Benin and Senegal exhibiting long-distance migrations of up to 4,000 km, traversing the Mediterranean Sea and Sahara from temperate Europe. In contrast, *V. cardui* captured in the circum-Mediterranean region at the same time either undertook short-distance migrations or displayed isotopic compositions indicative of local origins. These differences in migratory behavior did not have a clear association with wing morphology or wing wear. Genome-wide comparisons between these migratory groups found negligible genetic differentiation, leading us to conclude that the Afro-Palearctic population of *V. cardui* is panmictic, with regular bidirectional trans-Saharan crossings during the annual migratory cycle. Given the life-history traits of *V. cardui*, namely multigenerational migration, their mating strategy, and large population sizes, local adaptation appears implausible within this population. Consequently, future research should explore the role of environmental cues and differential gene expression in shaping the diversity in migration distance.

## Materials and methods

### Sample collection

The winter distribution of *V. cardui* in the western Afro-Palearctic is separated by the Sahara into two regions: the circum-Mediterranean region and sub-Saharan Africa (19, 35, 45). These regions are largely unsuitable larval habitat for *V. cardui* in the summer due to high temperatures and low host plant availability, and monitoring data in these regions suggest that *V. cardui* are either absent in the summer or occur in very low numbers (19, 35, 45). Monitoring data in sub-Saharan Africa documents

peaks in the number of *V. cardui* adults in the autumn (August to October), followed, but not preceded, by a peak in larval counts, suggesting that the sudden spike in adults in the autumn is a result of immigration to the region (19). The situation in the circum-Mediterranean region is similar except that this region becomes suitable for *V. cardui* development slightly later in the autumn (September/October), with numbers peaking in November (19). If a large portion of the offspring of immigrants to the circum-Mediterranean region migrated south across the Sahara, monitoring data should show a second peak in adult counts around December, as the offspring of early immigrants to sub-Saharan Africa mix with new immigrants; however, no such peak is observed in monitoring data (19). This suggests that the sampled butterflies represent the founders of the winter generations in each region.

In total, 40 wild, adult *V. cardui* were captured from 6 countries located both north and south of the Sahara during the time when *V. cardui* are arriving to the winter distribution areas (Table S1). South of the Sahara, butterflies were captured in Senegal and Benin from August to November of 2018 and 2019 ( $n = 19$ ). North of the Sahara, butterflies were captured from Morocco, Malta, Portugal, and Spain in November 2019 ( $n = 21$ ).

We preferentially sampled butterflies with worn wings (i.e. wing wear scores  $>1$ ) since a wing wear score of 1, “perfect wing condition,” indicates very recent emergence (36; i.e. that the butterfly was the offspring of early immigrants or summer occupants). We also preferentially sampled butterflies that displayed reproductive behavior (i.e. hilltopping or oviposition) because *V. cardui* delay their reproductive behavior until migration has been completed (33).

## Isotopic analysis

### Hydrogen isotope values

Prior to  $\delta^2\text{H}$  analysis, a forewing from each *V. cardui* was soaked, with agitation, in three successive baths (1 h, 30 min, and 10 min) of 2:1 chloroform:methanol solution to remove superficial dust and lipids, which are known to introduce error into  $\delta^2\text{H}$  measurements (87, 88). Samples were cut from the same membranous region of the forewing to reduce intraindividual variation from pigmentation (87) and wing veins (22). Samples were then weighed ( $0.162 \pm 0.009$  mg) into silver capsules, and surface moisture was dried in a  $40^\circ\text{C}$  oven. Nonexchangeable  $\delta^2\text{H}$  values were determined at the Ján Veizer Stable Isotope Laboratory (University of Ottawa, ON, Canada) using a comparative equilibrium approach (89). All measurements were taken using high-temperature ( $1,400^\circ\text{C}$ ) flash pyrolysis (TCEA, Thermo Finnigan, Germany) with a helium carrier passed through a chromium-filled reactor and, after separation, introduced to a Delta V Plus IRMS (Thermo Finnigan) via a Conflow IV interface (Thermo Finnigan). A three-point calibration was used to calibrate the sample  $\delta^2\text{H}$  values: CBS (caribou hoof;  $-157 \pm 0.9\text{‰}$  (90)), KHS (kudu horn;  $-35.3 \pm 1.1\text{‰}$  (90)), and USGS43 (human hair;  $-44.4 \pm 2.0\text{‰}$  (91)). Internal standards were measured to assess the quality of the measurements, including one keratin reference standard, USGS42 (human hair; measured:  $-75.3 \pm 0.5\text{‰}$ ,  $n = 4$ ; standard:  $-72.9 \pm 2.2\text{‰}$  (91)), and two in-house chitin standards, ground and homogenized *Lymantria dispar* (measured:  $-64.4 \pm 1.8\text{‰}$ ,  $n = 6$ ; long-term average:  $-64 \pm 0.8\text{‰}$ ) and Alfa Aesar chitin (measured:  $-22.8 \pm 0.7\text{‰}$ ,  $n = 4$ ; long-term average:  $-22 \pm 1.2\text{‰}$ ). Based on within-run replicates of the internal standards and repeated sample measurements, precision is estimated to be about  $\pm 2\text{‰}$ . All reported  $\delta^2\text{H}$  values are normalized to the Vienna Standard



Mean Ocean Water—Standard Light Antarctic Precipitation standard scale.

### Strontium isotope ratios

To prepare for  $^{87}\text{Sr}/^{86}\text{Sr}$  analysis, a single forewing from each butterfly was cleaned using pressurized nitrogen gas for 2 min at 69 kPa to remove any surface contaminants (e.g. dust; 24). The wings were then digested in 1 mL of 16 M  $\text{HNO}_3$  (distilled TraceMetal Grade; Fisher Chemical, Canada) using microwave digestion (Anton Paar Multiwave 7000; Austria). The samples were heated from ambient temperature to 250 °C at a steady rate over 20 min and then left at 250 °C for 15 min in a pressurized chamber. An aliquot of 200  $\mu\text{L}$  from each sample was separated and diluted to 2 mL of 2% v/v  $\text{HNO}_3$  and analyzed for Sr content via inductively coupled plasma mass spectrometry (ICP-MS; Agilent 8800 ICP-QQQ; Agilent Technologies Inc., CA, USA) at the Department of Earth and Environmental Science, University of Ottawa (ON, Canada). Calibration standards were prepared using single-element-certified standards purchased from SCP Science (Montreal, Canada). Samples had an average of  $3.1 \pm 1.7$  ng Sr per mg of wing tissue ( $n = 40$ ).

The remaining aliquot of sample digest was dried down and re-dissolved in 0.5 mL 6 M  $\text{HNO}_3$ . The separation of Sr was processed in a 100- $\mu\text{L}$  microcolumn loaded with Sr-spec Resin (100–150  $\mu\text{m}$ ; Eichrom Technologies, LLC). The matrix was rinsed out using 6 M  $\text{HNO}_3$ , and Sr was collected with 0.05 M  $\text{HNO}_3$ . To ensure complete separation, the microcolumn process was repeated. After separation, eluates were dried and re-dissolved in 200  $\mu\text{L}$  2% v/v  $\text{HNO}_3$  for  $^{87}\text{Sr}/^{86}\text{Sr}$  analysis.

The  $^{87}\text{Sr}/^{86}\text{Sr}$  analysis was performed at the Pacific Centre for Isotopic and Geochemical Research (University of British Columbia, BC, Canada) using a Nu-Plasma II high-resolution MC-ICP-MS (Nu Instruments) coupled to a desolvating nebulizer (Aridus II, CETAC Technologies). The isotopes  $^{84}\text{Sr}$ ,  $^{86}\text{Sr}$ , and  $^{87}\text{Sr}$  have isobaric interferences from  $^{84}\text{Kr}$ ,  $^{86}\text{Kr}$ , and  $^{87}\text{Rb}$  respectively, and were corrected for using the  $^{85}\text{Rb}$  and  $^{83}\text{Kr}$  signals. Instrumental mass fractionation was corrected by normalizing  $^{86}\text{Sr}/^{88}\text{Sr}$  to 0.1194 using the exponential law (92). Procedural blanks were negligible. The reproducibility of the  $^{87}\text{Sr}/^{86}\text{Sr}$  measurement for 5 ppb NIST SRM987 is  $0.71025 \pm 0.00009$  (1 SD,  $n = 138$ ) and for 1.4 ppb NIST SRM987,  $0.71019 \pm 0.00011$  ( $n = 48$ ). A matrix-matched chitin internal standard, 5 ppb Alfa Aesar chitin, was also used ( $0.713959 \pm 0.00009$ ;  $n = 3$ ).

### Geographic assignment

Continuous-surface isotope-based geographic assignment was performed via the assignR package (93) in R v4.1.0 (94). The origins of the 40 *V. cardui* individuals were estimated through a probabilistic framework using previously calibrated isoscapes: (i) a hydrogen isoscape calibrated with residential butterfly samples from across the Palearctic and Afrotropical regions (95) and (ii) a regional bio-available strontium isoscape (96). The most likely natal origin of each individual was estimated using  $\delta^2\text{H}$  isotopes alone,  $^{87}\text{Sr}/^{86}\text{Sr}$  isotopes alone, and using both isotopes together (see Data Availability). The 2:1 odds ratio was chosen, in keeping with previous studies (e.g. 17), to create binary surfaces from the posterior probability surfaces of the dual  $\delta^2\text{H}$ - and  $^{87}\text{Sr}/^{86}\text{Sr}$ -based geographic assignment, representing the top third of the probability distribution as high probability of natal origin (i.e. “1”) and everything else as low probability of natal origin (i.e. “0”). Individual binary surfaces were then summed for each capture location, and divided by the number of individuals, to summarize the results of the

geographic assignment into maps of the proportion of individuals with a high probability of natal origin at a given raster cell.

### Distance metric

A consensus has not been reached on the optimal method by which to estimate the distance a migratory individual has travelled from posterior probability surfaces (97). We assessed the ability of seven different metrics to accurately estimate the migration distance of aerial migrants in our study area using synthetic data; we generated isotopic compositions for 250 synthetic migrants, performed dual  $\delta^2\text{H}$  and  $^{87}\text{Sr}/^{86}\text{Sr}$  geographic assignment for each synthetic migrant, and calculated each of seven distance metrics. Most metrics were highly inaccurate, and could under- or over-estimate the actual great circle distance by thousands of kilometers (Fig. S1). However, the “minimum distance” metric consistently underestimated migration distance for our study area and, therefore, acts as a conservative estimate of migration distance (Fig. S1). Therefore, in this study, we used the minimum distance metric to estimate the distance travelled. This metric measures the shortest distance between the capture location and an isopleth threshold, here defined as the top third of the probability distribution (0.3 isopleth; e.g. 98, 99).

Many butterflies had an isotopic composition similar to the local isotopic composition of their capture location. We interpreted individuals with a highly probable natal origin at the capture location (as defined by the 2:1 odds ratio) and/or that had a minimum distance <100 km as “putative locals.” These individuals are likely either the locally sourced offspring of the previous migratory generation or were immigrants from a far-off region with a similar isotopic composition to that of their capture location. Our interpretation of these samples as putative locals is the more parsimonious approach.

### Geometric morphometric analysis

Differences in migration distance or trajectories could potentially be explained by intrinsic factors such as sex, body size, or wing shape. To test for differences in wing shape, photographs of each sample’s forewing were taken and used for morphometrics analysis. Since we preferentially chose samples that had higher wing wear scores, extensive damage to the wings made landmarking of many samples impossible. Therefore, only 28 of the 40 samples were landmarked using tpsDig2 v.2.32 (100) using 12 conventional landmarks for Nymphalid butterflies (e.g. 101, 102). Using the R package geomorph (103), a generalized procrustes analysis removed the translational, rotational, and scaling variation of the shape. A procrustes regression was used to test the effects on wing shape of sex, month of collection, estimated migration distance, and butterfly wing size.

### Genomic analysis and genotyping

#### Whole-genome resequencing

Library preparation for all 40 samples was done using TruSeq PCR-free kits (Illumina, San Diego, CA, USA) at the Institut Botànic de Barcelona (IBB—CSIC, Barcelona, Spain). Whole-genome resequencing was performed with paired-end  $2 \times 151$  bp reads on one Illumina NovaSeq 6000 S4 lane at the National Genomics Infrastructure (NGI; Stockholm, Sweden) to generate a mean depth of coverage per individual of about 48 reads.

#### SNP calling

A modified Sarek Nextflow pipeline (104) was used for preprocessing of the whole-genome resequencing data and for detecting

SNPs. At the first step of the Sarek pipeline, we evaluated the quality of the raw reads with FastQC (version 0.11.9; 105). Reads were stringently trimmed using fastp (106, 107) to remove (i) sequences from the 5' end (12 bp) and 3' end (3 bp) of each read, (ii) adaptor sequences, (iii) low-quality reads (overall Phred score <30), and (iv) short reads (<30 bp). Read quality after trimming was re-evaluated using FastQC. At the next step, a mean of 96.6% of the reads were mapped to the reference genome (108) using BWA-MEM (version 0.7.17; 109) and then sorted and indexed using SAMtools (version 1.9; 110). Optical and PCR duplicates were marked using GATK MarkDuplicates (version 4.1.4.1; 111). These duplicates, as well as sequences with low mapping quality (MAPQ < 30), were removed using SAMtools. Since prior SNP data for *V. cardui* were unavailable, the base calibration step module in the Sarek pipeline was skipped.

The first steps of variant calling were performed using GATK Best Practices (112), as implemented through the Sarek pipeline. In order to decrease computational time, the Sarek pipeline calls each chromosome separately. MultiQC (version 1.8) was used to assess the quality of the output from all steps of the pipeline. The resulting per chromosome vcf files included from 1.3 to 3 million unfiltered SNPs (78.9 million SNPs total). The last step of variant calling was performed outside of the pipeline to accommodate for invariant site calling (-all-sites option in GenotypeGVCFs), which is necessary for reliable downstream estimation of population genetic summary statistics, such as  $\pi$  and  $d_{XY}$ . This variant filtering included multiple steps: (i) hard filtering, (ii) removing SNPs overlapping with TE annotation, (iii) separate quality and coverage filtering for variant and invariant sites, and (iv) merging of variant and invariant site vcfs for the  $F_{ST}$  estimation. Hard filtering, which is meant to substitute for the variant quality score recalibration step in GATK Best Practices (112), was performed on raw vcf files, which included both variant and invariant sites. Using bcftools v1.15 (113), we filtered for variant quality by depth (QD < 2.0), Fisher strand bias (FS > 60.0), root mean square mapping quality (MQ < 40.0),  $u$ -based  $z$ -approximation from the Rank Sum Test for mapping qualities (MQRankSum < -12.5), and  $u$ -based  $z$ -approximation from the Rank Sum Test for site position (ReadPosRankSum < -8.0). We proceeded to remove indels and multivariate sites (-exclude-type indels, other --max-alleles 2) and reindexed the resulting vcf file. Next, the coordinates of transposable elements and simple repeats were obtained with a custom bash script from a previously published annotation of transposable elements (114), and used to remove SNPs appearing in these regions.

The quality of the resulting files (one vcf file per chromosome) was evaluated using summary statistics implemented with vcftools (i.e. allele frequency [-freq2], mean depth [-depth], mean per site depth [-site-mean-depth], per site quality [-site-quality], per sample missingness [-missing-indv], per site missingness [-missing-indv], and heterozygosity estimation [-het]; 115). Two individuals were removed due to quality issues (sample IDs: 19H137 and 19H114). Variant and invariant sites were split into two files for downstream analysis. Informed by the summary statistics, we applied filters for quality and depth to the variant sites vcf file (i.e. minor allele frequency [-max-maf 0], minimum mean-depth value [-min-meanDP 30], and maximum mean-depth value [-max-meanDP 80]). We also applied additional filters to the invariant sites vcf file (i.e. quality value [-minQ 30] and minimum mean-depth value [-min-meanDP 30]). Separate files with both variant and invariant calls were kept as separate chromosomes to decrease computational time. The W chromosome was also removed from the analysis. A region with divergent values of populational statistics was detected in the first

third of chromosome 10 corresponding to a large area of transposable element proliferation (Fig. 2C; 114). Since this region likely does not carry information about genetic structure it was also filtered from the population structure analyses. On a technical note, it is important to mention that this region was retained after stringent SNP filtration, and future research on *V. cardui* should be aware that this region may necessitate special treatment (e.g. manual filtering).

### Population genetic analysis

For PCA and admixture analysis, variant files per chromosome were concatenated. To investigate genetic variation between individuals, a PCA was performed using PLINK v1.90 b 4.9 (Fig. 2A). For this analysis, we additionally filtered out singleton SNP sites (i.e. MAF = 0.03), and SNPs that were in linkage (i.e. SNPs with  $r^2$  values > 0.1 using sliding windows of 50 SNPs with a step size of 10 SNPs); a total of 813,810 high-quality variants remained after filtering. Outliers in the first two principal components were detected using the R package pcadapt (116) using Bonferroni-corrected P-values ( $\alpha = 0.05$ ). The PCA was repeated excluding the two outliers, but no additional structure was detected (Fig. S5).

Reverse migrations from sub-Saharan Africa to the Maghreb are thought to occur in the autumn, even while southward trans-Saharan migrations are still underway (36). All of the samples captured in Morocco were designated as putative locals, but the posterior probability surfaces of some showed a possibility of originating south of the Sahara (Fig. 1E). To test whether these samples could have biased our results, we repeated the PCA without samples collected in Morocco (Fig. S5). The resulting PCA was based on 647,906 SNPs and showed a single cluster with the same outlier individuals as the original PCA.

After excluding SNPs on the Z chromosome and in coding sequences, genetic structure was also assessed with a likelihood method, the SNMF clustering algorithm in the R package LEA v3.6.0 (117). Each number of ancestral populations or clusters (K) from 1 to 5 was run independently with five repetitions with the regularization parameter set to 100, after which cross-validation errors calculated using the cross-entropy criterion were compared. This process was repeated using a Bayesian framework via fastSTRUCTURE (118) using the simple allele frequency prior with five repetitions. The number of population clusters from 1 to 5 that best explained the observed structure was chosen with the chooseK.py tool.

### Genomic scan for selection

To detect any fine-scale regions of differentiation along the genome, a genome scan was performed on groupings based on capture location (i.e. north vs. south of the Sahara) and migratory phenotype (i.e. long vs. short migration distance). For the grouping based on migratory phenotype, individuals from south of the Sahara that were putative locals ( $n = 6$ ) were removed, while putative locals from north of the Sahara ( $n = 19$ ) were included with the short-distance migrants ( $n = 3$ ) to balance the sample size. For genome scans, we retained SNPs falling into coding regions and having minor allele frequencies. We used pixy v1.2.7 (119, 120) to calculate nucleotide diversity ( $\pi$ ), absolute divergence ( $d_{XY}$ ), and genetic differentiation ( $F_{ST}$ ; Hudson method [-hudson]) in 20 kb nonoverlapping windows using all sites. Windows with <10% of sites represented (i.e. <2,000 bp) were removed from the analysis. Resulting per-window statistics were summarized using custom python scripts. For low levels of  $F_{ST}$ , we tested whether the

index significantly differed from 0 using a t test in the python scipy package ( $P$ -value:  $1.3e-216$ ).

## Acknowledgments

The authors acknowledge support from the NGI in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and Swedish National Infrastructure for Computing (SNIC)/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden and the SNIC in Uppsala, partially funded by the Swedish Research Council through grant agreement nos. 2022-06725 and 2018-05973. The authors are grateful to members of the Backström lab and Sean Stankowski for insightful discussions about the genome scan analysis. They thank Smita Mohanty for her help with the ICP-MS, Kerry Klassen and Paul Middlestead at the Ján Veizer Stable Isotope Laboratory for their help with the hydrogen isotope analysis, and Kathy Gordon at the Pacific Centre for Isotopic and Geochemical Research for her help with the strontium isotope analysis. The authors thank Leonardo Dapporto, Leonardo Platania, and Aurora García-Berro for sharing code and knowledge related to the wing geometric morphometrics analysis, and Eude Goudégnon, Roger Vila, Mattia Menchetti, and Tomasz Suchan for help in sample collection. The authors also thank Andrea Contina, Gabriel Bowen, and Michael Wunder for sharing code and insight relating to the migration distance estimates.

## Supplementary Material

Supplementary material is available at PNAS Nexus online.

## Funding

This study was funded by grant 2018-00738 of the Government of Canada (New Frontiers in Research Fund) to G.T. and C.P.B.; by grants PID2020-117739GA-I00, PID2023-152239NB-I00 (MCIN/AEI/10.13039/501100011033), and 2021-SGR-01334 Generalitat de Catalunya (Departament de Recerca i Universitats) to G.T.; by grant COOPB23061 to F.B., K.K., and G.T.; and by grant LINKA20399 from the Consejo Superior de Investigaciones Científicas (Spanish National Research Council, CSIC) iLink program to N.B., C.P.B., and G.T. M.S.R. was supported by the Government of Ontario (Queen Elizabeth II Graduate Scholarship in Science and Technology) (QEII-GSST), the Government of Ontario (Ontario Graduate Scholarship), and the National Science Foundation (2021 ORIGIN Project; DBI-1565128). N.B. is supported by a research grant from the Swedish Research Council FORMAS (grant 2019-00670). D.S. is supported by the Knut and Alice Wallenberg Foundation and the Erling Persson Family Foundation (Postdoctoral Thunberg Fellowship Programme of Swedish Collegium for Advanced Study).

## Author Contributions

G.T. and C.P.B. conceived the idea for the study and organized the sample collection. G.T., M.S.R., K.K., and F.B. collected the samples. M.S.R. and C.P.B. performed the isotope analysis component. D.S., M.S.R., G.T., V.T., and N.B. performed the genomic analysis

component. J.L.B. did the landmarking for the geometric morphometric analysis. M.S.R. and D.S. led the analysis and writing of the manuscript. All authors contributed to the drafts and gave final approval for publication.

## Preprints

Preprint servers: bioRxiv (<https://doi.org/10.1101/2023.12.10.569105>).

## Data Availability

The raw sequence data are deposited in the European Nucleotide Archive with accession number PRJEB70298. All data and code pertaining to the isotopic, geometric morphometric, and genomic analyses, including bash, R, and python scripts, can be found on the Open Science Foundation (<https://osf.io/dzh2g/>).

## References

- Dingle H, Drake VA. 2007. What is migration? *BioScience*. 57: 113–121.
- Chapman JW, Reynolds DR, Wilson K. 2015. Long-range seasonal migration in insects: mechanisms, evolutionary drivers and ecological consequences. *Ecol Lett*. 18:287–302.
- Turbek SP, Scordato ESC, Safran RJ. 2018. The role of seasonal migration in population divergence and reproductive isolation. *Trends Ecol Evol*. 33:164–175.
- Berthold P. 1999. A comprehensive theory for the evolution, control and adaptability of avian migration. *Ostrich*. 70:1–11.
- Helbig AJ. 1996. Genetic basis, mode of inheritance and evolutionary changes of migratory directions in Palearctic warblers (Aves: Sylviidae). *J Exp Biol*. 199:49–55.
- Sokolovskis K, et al. 2023. Migration direction in a songbird explained by two loci. *Nat Commun*. 14:165.
- Delmore K, et al. 2020. The evolutionary history and genomics of European blackcap migration. *eLife*. 9:e54462.
- Delmore K, Fox JW, Irwin DE. 2012. Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. *Proc Biol Sci*. 279: 4582–4589.
- Mueller JC, Pulido F, Kempaers B. 2011. Identification of a gene associated with avian migratory behaviour. *Proc R Soc Lond B Biol Sci*. 278:2848–2856.
- Shaw AK. 2020. Causes and consequences of individual variation in animal movement. *Mov Ecol*. 8:1–12.
- Talla V, et al. 2020. Genomic evidence for gene flow between monarchs with divergent migratory phenotypes and flight performance. *Mol Ecol*. 29:2567–2582.
- Chowdhury S, Fuller RA, Dingle H, Chapman JW, Zalucki MP. 2021. Migration in butterflies: a global overview. *Biol Rev Camb Philos Soc*. 96(4):1462–1483.
- Liedvogel M, Åkesson S, Bensch S. 2011. The genetics of migration on the move. *Trends Ecol Evol*. 26:561–569.
- Menz MHM, et al. 2022. Individual tracking reveals long-distance flight-path control in a nocturnally migrating moth. *Science*. 377:764–768.
- Knight SM, Pitman GM, Flockhart DTT, Norris DR. 2019. Radio-tracking reveals how wind and temperature influence the pace of daytime insect migration. *Biol Lett*. 15:20190327.
- Wikelski M, et al. 2006. Simple rules guide dragonfly migration. *Biol Lett*. 2:325–329.

- 17 Flockhart DTT, et al. 2013. Tracking multi-generational colonization of the breeding grounds by monarch butterflies in eastern North America. *Proc Biol Sci.* 280:20131087.
- 18 Sun W, et al. 2022. Population source of third-generation oriental armyworm in Jilin, China, determined by entomology radar, trajectory analysis, and mitochondrial COI sequences. *Environ Entomol.* 51:621–632.
- 19 Talavera G, et al. 2023. The Afrotropical breeding grounds of the Palearctic-African migratory painted lady butterflies (*Vanessa cardui*). *Proc Natl Acad Sci U S A.* 120:e2218280120.
- 20 Flockhart DTT, Kyser TK, Chipley D, Miller NG, Norris DR. 2015. Experimental evidence shows no fractionation of strontium isotopes ( $^{87}\text{Sr}/^{86}\text{Sr}$ ) among soil, plants, and herbivores: implications for tracking wildlife and forensic science. *Isotopes Environ Health Stud.* 51:372–381.
- 21 Hobson KA, Wassenaar LI, Taylor OR. 1999. Stable isotopes ( $\delta\text{D}$  and  $\delta^{13}\text{C}$ ) are geographic indicators of natal origins of monarch butterflies in eastern North America. *Oecologia.* 120:397–404.
- 22 Lindroos EE, Bataille CP, Holder PW, Talavera G, Reich MS. 2023. Temporal stability of  $\delta^2\text{H}$  in insect tissues: implications for isotope-based geographic assignments. *Front Ecol Evol.* 11:1060836.
- 23 Reich MS, et al. 2023. Metals and metal isotopes incorporation in insect wings: implications for geolocation and pollution exposure. *Front Ecol Evol.* 11:1085903.
- 24 Reich MS, Flockhart DTT, Norris DR, Hu L, Bataille CP. 2021. Continuous-surface geographic assignment of migratory animals using strontium isotopes: a case study with monarch butterflies. *Methods Ecol Evol.* 12(12):2445–2457.
- 25 Wassenaar LI, Hobson KA. 1998. Natal origins of migratory monarch butterflies at wintering colonies in Mexico: new isotopic evidence. *Proc Natl Acad Sci U S A.* 95:15436–15439.
- 26 Bataille CP, Crowley BE, Wooller MJ, Bowen GJ. 2020. Advances in global bioavailable strontium isoscapes. *Palaeogeogr Palaeoclimatol Palaeoecol.* 555:109849.
- 27 Crowley BE, Bataille CP, Haak BA, Sommer KM. 2021. Identifying nesting grounds for juvenile migratory birds with dual isotope: an initial test using North American raptors. *Ecosphere.* 12:e03765.
- 28 Kruszynski C, et al. 2021. Identifying migratory pathways of Nathusius' pipistrelles (*Pipistrellus nathusii*) using stable hydrogen and strontium isotopes. *Rapid Commun Mass Spectrom.* 35:1–11.
- 29 Shields O. 1992. World distribution of the *Vanessa cardui* group (Nymphalidae). *J Lepid Soc.* 46:235–238.
- 30 Williams CB. 1970. The migrations of the painted lady butterfly, *Vanessa cardui* (Nymphalidae), with special reference to North America. *Lepid Soc.* 24(3):157–175.
- 31 Shipilina D, et al. 2024. Gene expression responses to environmental cues shed light on components of the migratory syndrome in butterflies. *bioRxiv* 602486. <https://doi.org/10.1101/2024.07.17.602486>, preprint: not peer reviewed.
- 32 Nasvall K, Shipilina D, Vila R, Talavera G, Backstrom N. 2023. Resource availability affects activity profiles of regulatory elements in a long-distance butterfly migrant. <https://doi.org/10.22541/au.167827909.99815237/v1>, preprint: not peer reviewed.
- 33 Stefanescu C, Ubach A, Wiklund C. 2021. Timing of mating, reproductive status and resource availability in relation to migration in the painted lady butterfly. *Anim Behav.* 172:145–153.
- 34 Menchetti M, Guéguen M, Talavera G. 2019. Spatio-temporal ecological niche modelling of multigenerational insect migrations. *Proc R Soc Lond B Biol Sci.* 286:20191583.
- 35 Stefanescu C, et al. 2013. Multi-generational long-distance migration of insects: studying the painted lady butterfly in the Western Palearctic. *Ecography.* 36:474–486.
- 36 Stefanescu C, Soto DX, Talavera G, Vila R, Hobson KA. 2016. Long-distance autumn migration across the Sahara by painted lady butterflies: exploiting resource pulses in the tropical savannah. *Biol Lett.* 12(10):20160561.
- 37 Talavera G, Vila R. 2016. Discovery of mass migration and breeding of the painted lady butterfly *Vanessa cardui* in the Sub-Saharan: the Europe-Africa migration revisited. *Biol J Linn Soc Lond.* 120:274–285.
- 38 Miller NG, Wassenaar LI, Hobson KA, Norris R. 2012. Migratory connectivity of the monarch butterfly (*Danaus plexippus*): patterns of spring re-colonization in eastern North America. *PLoS One.* 7:e31891.
- 39 Altizer S, Davis A. 2010. Populations of monarch butterflies with different migratory behaviors show divergence in wing morphology. *Evolution.* 64:1018–1028.
- 40 Mackintosh A, et al. 2019. The determinants of genetic diversity in butterflies. *Nat Commun.* 10:3466.
- 41 Lohse K. 2017. Come on feel the noise—from metaphors to null models. *J Evol Biol.* 30:1506–1508.
- 42 Flockhart DTT, et al. 2017. Migration distance as a selective episode for wing morphology in a migratory insect. *Mov Ecol.* 5:7.
- 43 Li Y, Pierce AA, de Roode JC. 2016. Variation in forewing size linked to migratory status in monarch butterflies. *Animal Migration.* 3:27–34.
- 44 West-Eberhard MJ. 1989. Phenotypic plasticity and the origins of diversity. *Annu Rev Ecol Syst.* 20:249–278.
- 45 Stefanescu C, et al. 2017. Back to Africa: autumn migration of the painted lady butterfly *Vanessa cardui* is timed to coincide with an increase in resource availability. *Ecol Entomol.* 42:737–747.
- 46 Korkmaz R, Rajabi H, Eshghi S, Gorb SN, Büscher TH. 2022. The frequency of wing damage in a migrating butterfly. *Insect Sci.* 30(5):1507–1517.
- 47 Dapporto L, et al. 2019. Integrating three comprehensive data sets shows that mitochondrial DNA variation is linked to species traits and paleogeographic events in European butterflies. *Mol Ecol Resour.* 19:1623–1636.
- 48 Pfeiler E, Markow TA. 2017. Population connectivity and genetic diversity in long-distance migrating insects: divergent patterns in representative butterflies and dragonflies. *Biol J Linn Soc Lond.* 122:479–486.
- 49 Vodá R, et al. 2016. Historical and contemporary factors generate unique butterfly communities on islands. *Sci Rep.* 6:28828.
- 50 Singh VK, Kumar S, Awasthi S, Gupta SK. 2021. Phylogeography and population expansion of Western Himalayan highly-erratically migratory painted lady butterfly (Lepidoptera: Nymphalidae: *Vanessa cardui*). *J Asia Pac Entomol.* 24:1116–1121.
- 51 Alvial IE, et al. 2019. Isolation on a remote island: genetic and morphological differentiation of a cosmopolitan odonate. *Heredity (Edinb).* 122:893–905.
- 52 Ehrlich MA, Wagner DN, Oleksiak MF, Crawford DL. 2021. Polygenic selection within a single generation leads to subtle divergence among ecological niches. *Genome Biol Evol.* 13:evaa257.
- 53 Als TD, et al. 2011. All roads lead to home: panmixia of European eel in the Sargasso Sea. *Mol Ecol.* 20:1333–1346.
- 54 Pujolar JM, et al. 2014. Genome-wide single-generation signatures of local selection in the panmictic European eel. *Mol Ecol.* 23:2514–2528.



- 55 García-Berro A, et al. 2023. Migratory behaviour is positively associated with genetic diversity in butterflies. *Mol Ecol.* 32: 560–574.
- 56 Satterfield DA, Wright AE, Altizer S. 2013. Lipid reserves and immune defense in healthy and diseased migrating monarchs *Danaus plexippus*. *Curr Zool.* 59:393–402.
- 57 Enbody ED, et al. 2021. Ecological adaptation in European eels is based on phenotypic plasticity. *Proc Natl Acad Sci U S A.* 118: e2022620118.
- 58 Matthews JH. *Research in motion: patterns of large-scale migration in dragonflies and birds*. The University of Texas at Austin, 2007.
- 59 May ML, Matthews JH. 2008. Migration in Odonata: a case study of *Anax junius*. In: Cordoba-Aguilar A, Beatty C, Bried J, editors. *Dragonflies and damselflies: model organisms for ecological and evolutionary research*. Oxford University Press. p. 63–78.
- 60 Ware J, et al. 2022. Evidence for widespread gene flow and migration in the Globe Skimmer dragonfly *Pantala flavescens*. *Int J Odonatol.* 25:43–55.
- 61 Troast D, Suhling F, Jingui H, Sahlén G, Ware J. 2016. A global population genetic study of *Pantala flavescens*. *PLoS One.* 11: e0148949.
- 62 Chelli A, Moulaï R. 2019. Ecological characterization of the odonotofauna in lotic and lentic waters of northeast Algeria. *Ann Soc entomol Fr.* 55:430–445.
- 63 Pedgley DE, Reynolds DR, Tatchell GM. 1995. Long-range insect migration in relation to climate and weather: Africa and Europe. In: Gatehouse G, Drake A, editors. *Insect migration: tracking resources through space and time*. Cambridge University Press. p. 3–29.
- 64 Gu Z, et al. 2021. Climate-driven flyway changes and memory-based long-distance migration. *Nature.* 591:259–564.
- 65 Thompson NF, et al. 2020. A complex phenotype in salmon controlled by a simple change in migratory timing. *Science.* 370: 609–613.
- 66 Niitepöld K, et al. 2009. Flight metabolic rate and Pgi genotype influence butterfly dispersal rate in the field. *Ecology.* 90: 2223–2232.
- 67 Pegan TM, et al. 2024. Population genetic consequences of the seasonal migrations of birds. *bioRxiv* 601242. <https://doi.org/10.1101/2024.06.28.601242>, preprint: not peer reviewed.
- 68 Conklin JR, et al. 2024. High dispersal ability versus migratory traditions: fine-scale population structure and post-glacial colonisation in bar-tailed godwits. *Mol Ecol.* 33:e17452.
- 69 Pliske TE. 1975. Courtship behavior of the monarch butterfly, *Danaus plexippus* L. *Ann Entomol Soc Am.* 68:143–151.
- 70 Wiklund C, Friberg M. 2022. Testing the migration syndrome: comparative fecundity of migratory and non-migratory nymphaline butterflies. *Ecol Entomol.* 47:1061–1067.
- 71 Chapuis M-P, et al. 2009. Outbreaks, gene flow and effective population size in the migratory locust, *Locusta migratoria*: a regional-scale comparative survey. *Mol Ecol.* 18:792–800.
- 72 Gorki JL, et al. 2024. Pollen metabarcoding reveals the origin and multigenerational migratory pathway of an intercontinental-scale butterfly outbreak. *Curr Biol.* 34(12):2684–2692.e6.
- 73 López-Mañas R, et al. 2022. Erratic spatiotemporal vegetation growth anomalies drive population outbreaks in a trans-Saharan insect migrant. *Proc Natl Acad Sci U S A.* 119:3–5.
- 74 Neck RW. 1983. Causal analysis of a migration of the snout butterfly, *Libytheana bachmanii larvata* (Strecker) (Libytheidae). *J Lepid Soc.* 37:121–128.
- 75 Russell RW, May ML, Soltesz KL, Fitzpatrick JW. 1998. Massive swarm migrations of dragonflies (Odonata) in eastern North America. *Am Midl Nat.* 140:325–342.
- 76 Menz MHM, Brown BV, Wotton KR. 2019. Quantification of migrant hoverfly movements (Diptera: Syrphidae) on the west coast of North America. *R Soc Open Sci.* 6:190153.
- 77 Buffalo V. 2021. Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's Paradox. *eLife.* 10:e67509.
- 78 Altizer S, Hobson KA, Davis AK, De Roode JC, Wassenaar LI. 2015. Do healthy monarchs migrate farther? Tracking natal origins of parasitized vs. uninfected monarch butterflies overwintering in Mexico. *PLoS One.* 10:e0141371.
- 79 Du B, Ding D, Ma C, Guo W, Kang L. 2022. Locust density shapes energy metabolism and oxidative stress resulting in divergence of flight traits. *Proc Natl Acad Sci U S A.* 119:e2115753118.
- 80 Gatehouse AG. 1994. Insect migration: variability and success in a capricious environment. *Popul Ecol.* 36:165–171.
- 81 Chapman BB, et al. 2011. To boldly go: individual differences in boldness influence migratory tendency. *Ecol Lett.* 14:871–876.
- 82 Zhan S, et al. 2014. The genetics of monarch butterfly migration and warning coloration. *Nature.* 514:317–321.
- 83 Merlin C, Liedvogel M. 2019. The genetics and epigenetics of animal migration and orientation: birds, butterflies and beyond. *J Exp Biol.* 222:1–12.
- 84 Van Petegem KHP, et al. 2015. Empirically simulated spatial sorting points at fast epigenetic changes in dispersal behaviour. *Evol Ecol.* 29:299–310.
- 85 Boman J, et al. 2023. Environmental stress during larval development induces DNA methylation shifts in the migratory painted lady butterfly (*Vanessa cardui*). *Mol Ecol.* 32(13):3513–3523.
- 86 Jones CM, et al. 2015. Genomewide transcriptional signatures of migratory flight activity in a globally invasive insect pest. *Mol Ecol.* 24:4901–4911.
- 87 Hobson KA, et al. 2017. Within-wing isotopic ( $\delta^2\text{H}$ ,  $\delta^{13}\text{C}$ ,  $\delta^{15}\text{N}$ ) variation of monarch butterflies: implications for studies of migratory origins and diet. *Animal Migration.* 4:8–14.
- 88 Paritte JM, Kelly JF. 2009. Effect of cleaning regime on stable-isotope ratios of feathers in Japanese Quail (*Coturnix japonica*). *Auk.* 126:165–174.
- 89 Wassenaar LI, Hobson KA. 2003. Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. *Isotopes Environ Health Stud.* 39:211–217.
- 90 Soto DX, Koehler G, Wassenaar LI, Hobson KA. 2017. Re-evaluation of the hydrogen stable isotopic composition of keratin calibration standards for wildlife and forensic science applications. *Rapid Commun Mass Spectrom.* 31:1193–1203.
- 91 Coplen TB, Qi H. 2012. USGS42 and USGS43: human-hair stable hydrogen and oxygen isotopic reference materials and analytical methods for forensic science and implications for published measurement results. *Forensic Sci Int.* 214:135–141.
- 92 Moore LJ, Murphy TJ, Barnes IL, Paulsen PJ. 1982. Absolute isotopic abundance ratios and atomic weight of a reference sample of strontium. *J Res Natl Bur Stand (1977).* 87:1–8.
- 93 Ma C, Vander Zanden HB, Wunder MB, Bowen GJ. 2020. Assignr: an R package for isotope-based geographic assignment. *Methods Ecol Evol.* 11(8):996–1001.
- 94 R Core Team. 2021. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Deposited 2021.
- 95 Ghouri S, et al. 2024. A hydrogen isoscape for tracing the migration of herbivorous lepidopterans across the Afro-Palearctic range. *Rapid Comm Mass Spectrometry.* 38:e9675.

- 96 Reich MS, et al. 2024. Trans-Saharan migratory patterns in *Vanessa cardui* and evidence for a southward leapfrog migration. *iScience*. 27(12):111342.
- 97 Vander Zanden HB, Nelson DM, Wunder MB, Conkling TJ, Katzner T. 2018. Application of isoscapes to determine geographic origin of terrestrial wildlife for conservation and management. *Biol Conserv*. 228:268–280.
- 98 Hallworth MT, Marra PP, McFarland KP, Zahendra S, Studds CE. 2018. Tracking dragons: stable isotopes reveal the annual cycle of a long-distance migratory insect. *Biol Lett*. 14(12):20180741.
- 99 Katzner TE, et al. 2017. Golden Eagle fatalities and the continental-scale consequences of local wind-energy generation. *Conserv Biol*. 31:406–415.
- 100 Rohlf FJ. 2016. *tpsDig2* (Version 2.32) [Software]. Department of Ecology and Evolution, State University of New York at Stony Brook; [accessed 2016]. <https://www.sbmorphometrics.org/soft-dataacq.html>.
- 101 Escobar-Suárez S, et al. 2023. A geometric morphometrics and genetics characterization of *Vanessa carye* in an extreme elevational gradient in the Chilean Altiplano. *Zool Anz*. 304:105–112.
- 102 Soule AJ, Decker LE, Hunter MD. 2020. Effects of diet and temperature on monarch butterfly wing morphology and flight ability. *J Insect Conserv*. 24:961–975.
- 103 Adams DC, Otárola-Castillo E. 2013. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods Ecol Evol*. 4:393–399.
- 104 Garcia M, et al. 2020. Sarek: a portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Res*. 9:63.
- 105 Andrews S. 2010. *FastQC: a quality control tool for high throughput sequence data* (Version 0.10.0) [Software]. Babraham Bioinformatics; [accessed 2010]. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- 106 Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*. 2:e107.
- 107 Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34:i884–i890.
- 108 Lohse K, Wright C, Talavera G, García-Berro A. 2021. The genome sequence of the painted lady, *Vanessa cardui* Linnaeus 1758. *Wellcome Open Res*. 6:324.
- 109 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, *arXiv:1303.3997*. <https://doi.org/10.48550/arXiv.1303.3997>, preprint: not peer reviewed.
- 110 Li H, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25:2078–2079.
- 111 DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 43:491–498.
- 112 Van Der Auwera GA, et al. 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 11:1–43.
- 113 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 27:2987–2993.
- 114 Shipilina D, et al. 2022. Linkage mapping and genome annotation give novel insights into gene family expansions and regional recombination rate variation in the painted lady (*Vanessa cardui*) butterfly. *Genomics*. 114:110481.
- 115 Danecek P, et al. 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156–2158.
- 116 Privé F, Luu K, Vilhjálmsson BJ, Blum MGB. 2020. Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Mol Biol Evol*. 37:2153–2154.
- 117 Frichot E, Mathieu F, Trouillon T, Bouchard G, François O. 2014. Fast and efficient estimation of individual ancestry coefficients. *Genetics*. 196:973–983.
- 118 Raj A, Stephens M, Pritchard JK. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 197:573–589.
- 119 Samuk K, Korunes K, Trevisani MD, Moshiri N. 2022. ksamuk/pixy: pixy 1.2.7.beta1. Deposited 2022. <https://doi.org/10.5281/ZENODO.6551490>.
- 120 Korunes KL, Samuk K. 2021. Pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol Ecol Resour*. 21:1359–1368.