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

Megan S. Reich[a,1]*, Daria Shipilina[b,1]*, Venkat Talla[b], Farid Bahleman[c], Khadim Kébé[d], Johanna L. Berger[e,f], Niclas Backström[b], Gerard Talavera[f,2], Clément P. Bataille[g,2]

[a]Department of Biology, University of Ottawa, Ottawa, K1N 7N9 Canada

[b]Evolutionary Biology Program, Department of Ecology and Genetics, Uppsala University, Uppsala, 752 36 Sweden

[c]SOS Savane - NGO, Tanguiéta, 145 Benin

[d]GRBA-BE, Department of Chemical Engineering and Applied Biology, Polytechnic Higher School of Dakar, Dakar, 5085 Senegal

[e]Ecological Networks, Technische Universität Darmstadt, Darmstadt, 64287 Germany

[f]Institut Botànic de Barcelona (IBB), CSIC-CMCNB, Barcelona, 08038 Catalonia, Spain

[g]Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, K1N 7N9 Canada

[4]These authors contributed equally to this work

[5]Joint Senior Authors

*Corresponding authors:

Daria Shipilina; Norbyvägen 18D, Uppsala, Sweden; 752 36 ; +43-664-883-26278; daria.shipilina@ebc.uu.se

Megan S. Reich; 25 Templeton St, Ottawa, ON; K1N 7N9; 1-604-793-3808; megan.reich@uottawa.ca

Author Contributions: GT and CB conceived the idea for the study and organised the sample collection. GT, MR, KK, and FB collected the samples. MR and CB performed the isotope analysis component. DS, MR, GT, VT, and NB performed the genomic analysis component. JB did the landmarking for the geometrics morphometrics analysis. MR and DS

led the analysis and writing of the manuscript. All authors contributed to the drafts and gave final approval for publication.

Competing Interest Statement: The authors declare no competing interest.

Classification: Biological Sciences, Evolution

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

This file includes:

Main Text

Figures 1 to 2

Abstract

The painted lady butterfly *Vanessa cardui* is renowned for its virtually cosmopolitan distribution and the remarkable long-distance migrations that are part of its annual, multi-generational migratory cycle. Recently, *V. cardui* individuals were found north and south of the Sahara in the autumn, suggesting distinct migratory behaviours within the species. However, the evolutionary and ecological factors shaping these differences in migratory behaviour remain largely unexplored. Here, we performed whole-genome resequencing and analysed 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 both 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 (pairwise F_{ST} = 0.001 ± 0.006). Nonetheless, two individuals, which were early arrivals to West Africa and covered 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.

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 *V.cardui*, to explore the genetic basis of variation in insect migration behavior. We leverage new generation isotope geolocation techniques to uncover striking differences in butterfly behaviour, with some individuals migrating short distances within 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. This study provides a ground-breaking framework to study migration in insects.

Main Text

Introduction

Migration has evolved several times in insects (1) and is often viewed as an advantageous adaptation, as insects move to favourable environmental conditions, exploit seasonal resources, and escape parasites and predators (2). Many migratory animals show variation in migratory behaviour that results in distinctive spatiotemporal movement patterns. As an innate behaviour in insects, the proximate mechanisms for migration are likely held in the genome (3,4), but the environmental and genetic factors driving intraspecific variation in insect migratory behaviour are only beginning to be understood (5).

The painted lady butterfly *Vanessa cardui* is renowned for its obligate, annual, multi-generational migratory cycles throughout Africa, Asia, Europe, and North America (6,7). As part of a migratory cycle covering between 8 and 10 generations and involving both the Palearctic and Afrotropical regions, late summer and autumn generations migrate southward from temperate Europe towards more suitable breeding grounds. While a substantial portion of these southward migrating butterflies journey across the Mediterranean Sea and Sahara to subsequently breed in sub-Saharan Africa (8–10), others migrate to the circum-

111 Mediterranean region (11,12). The factors driving these differences in migratory behaviour,
112 resulting in the disjunct winter distribution of *V. cardui* across the Sahara, remains unknown.
113 Migratory phenotypes include observable characteristics such as migratory
114 propensity, timing, orientation, and wing morphology (13). However, migratory behaviours
115 are complex traits and are therefore difficult to quantify, especially in insects. The distance
116 an individual butterfly covers during migration is a consequence of this complex phenotype,
117 which creates a distribution of migration distances across individuals, reflecting the extent of
118 the species' ecological niche. In cases where the suitable niche is geographically
119 continuous, it is expected that individuals at the extremes of the migration distance
120 distribution (i.e., those who migrate the shortest or longest possible distances) are selected
121 against due to the inherent risks associated with extremes. In such scenarios, stabilising
122 selection may favour migrants that cover intermediate distances. Conversely, when the
123 suitable niche is fragmented by large biogeographic barriers, such as the Sahara, extreme
124 migration distances may be favoured by disruptive selection. Selection for specific migration
125 distances can have important repercussions for survival and growth. For example,
126 pronounced climatic differences between regions north and south of the Sahara influence
127 the composition and quantity of host plant species, predators, parasitoids, and weather
128 events, as well as influencing development time (9). Altogether, this scenario leads us to
129 ask: can we detect the action of adaptive evolutionary forces acting on the extremes of the
130 migration distance distribution and effectively disentangle their action from neutral
131 evolutionary processes?
132 Differences in migratory routes, distances, and orientation has led to genetic
133 differentiation and subsequent divergent selection in multiple bird species, such as the
134 European blackcap *Sylvia atricapilla,* Swainson's thrush *Catharus ustulatus,* and willow
135 warbler *Phylloscopus trochilus* (14–17). Likewise, differences in migration distance could
136 lead to similar spatiotemporal separation of *V. cardui* lineages, resulting in reduced genetic
137 exchange between *V. cardui* that migrate across the Sahara and those that remain in the
138 circum-Mediterranean region, thus leading to genetic differentiation. However, exploring the

genetic underpinnings of natural variation in migratory behaviour requires not only whole-genome sequencing data, but also accurate migratory phenotyping (4). Studies on insect migration have lagged behind those of birds and mammals because identifying migratory phenotypes in insects is challenging (18). Extrinsic markers such as biologgers commonly record migratory trajectories for birds and mammals, but few studies have used these techniques on insects (19-21) 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), behaviour (e.g., insect-monitoring radar), or regional field monitoring (e.g., spatiotemporal location).

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 local environmental isotopic signatures through their diet and incorporate these isotopic signatures into their developing tissues (22,23). Insect wings are relatively inactive tissues and largely preserve the isotopic signature of the natal environment (24,25). Thus, the isotopic signature of the wing of a wild-caught insect can be measured and then compared to 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 (26). Dual isotope-based geographic assignment combining hydrogen isotope values (δ^2H), driven by climatic variables (e.g., precipitation; 23,27), and strontium isotope ratios ($^{87}Sr/^{86}Sr$), primarily driven by geological processes (e.g., bedrock type and age; 28), have previously been used to classify migratory behaviours in birds and bats (e.g., 29,30). However, isotopes have not often been leveraged to phenotype migratory behaviour and investigate the potential genomic drivers of variation in insect migratory behaviour.

Here, we combined isotope-based phenotyping of migration distance with whole-genome resequencing data to investigate the mechanisms driving variation in migratory behaviour for *V. cardui* individuals. By coupling δ^2H and $^{87}Sr/^{86}Sr$ into a dual geographic assignment framework, we estimated the natal origin and migration distance of 40 adult *V. cardui* synchronically collected from north and south of the Sahara in the autumn. A genome-wide population resequencing approach was used to investigate potential genetic structuring associated with contrasting migratory phenotypes. This study endeavours to shed light on the underlying factors shaping the remarkable trans-Saharan migratory journeys of *V. cardui*.

Results

Isotope-based phenotyping reveals differences in migration distance

We estimated the natal origins of 40 *V. cardui* of unknown origin captured during the autumn southward migration from both north (i.e., Morocco, Spain, Portugal, Malta) and south (i.e., Benin, Senegal) of the Sahara using δ^2H and $^{87}Sr/^{86}Sr$ following the geographic assignment framework of Reich *et al.* (26) (S1 Table; Fig 1D). From the posterior probability surfaces, we then calculated a conservative estimate of migration distance (i.e., minimum distance; S1 Fig; Fig 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 over 500 to over 4,000 km), (ii) short-distance migrants (n = 3, 7%; with minimum migration distances ranging from over 140 to over 240 km), and (iii) individuals of putatively local origin (n = 24, 60%; having potentially travelled less than 100 km). Isotopic signatures corresponding to long-distance migrations were found exclusively in individuals captured south of the Sahara, specifically in Senegal and Benin. Based on the summarised 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

194 of the individuals (n = 6, representing 32% of the samples from south of the Sahara; Fig 1E)
195 had a high probability of originating from areas near their capture points (< 100 km) and
196 were designated as “putative locals.” Notably, no individuals from south of the Sahara were
197 classified in the short-distance migration category.

198 Among the samples collected north of the Sahara, some belonged to the short-
199 distance migration category (n = 3, 14%; Fig 1A). However, the majority of these samples
200 exhibited isotopic signatures consistent with local origin (n = 18, 86%; Fig 1E). The isotope-
201 based analysis doesn't clarify whether this local signature indicates (i) pre-migratory, locally-
202 sourced offspring of early immigrants to the region or (ii) immigrants from a distant region
203 with an isotopic signature similar to their capture location. We also considered natural history
204 observations to interpret these patterns. The majority of individuals collected in the circum-
205 Mediterranean region during October and November are likely the descendants of earlier
206 arrivals to the collection area. This inference is substantiated by the favourable breeding
207 conditions in the region during late September and October (9,11), as well as the observed
208 presence of eggs, larvae, and recently-emerged adults (i.e., expelling meconium) at the time
209 of collection in Morocco, Malta, and Portugal. Thus, we categorise any samples with a high
210 probability of originating near their capture location as putative locals.

211 To assess the genomic differentiation between migratory phenotypes, it is crucial to
212 integrate the isotopic signature data with spatiotemporal information from individual samples.
213 Consequently, for subsequent analysis, we grouped individuals that travelled long distances
214 from Senegal and Benin, explicitly excluding individuals with putatively local origins, into the
215 “long-distance” group. To balance sample sizes, we grouped individuals from north of the
216 Sahara, encompassing both the short-distance (n = 3) and putatively local categories (n =
217 18), into the “short-distance” group, operating under the assumption that these putative
218 locals are the next-generation offspring of short-distance migrants.
219

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 (8,31,32). 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 post-migratory 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., 33). 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 (S2 Fig; 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 (S2 Fig; $F_{1,27} = 3.0$, $p = 0.001$).

Lack of population structure across the Sahara

We explored if 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. Remarkably, 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.

248 First, using 13,215,545 high-quality callable sites from both coding and non-coding
249 regions (38 individuals, two removed due to quality issues), we summarised levels of
250 nucleotide diversity (π) within groups and genetic divergence (d_{XY}) between groups of *V.*
251 *cardui* individuals with distinct migratory behaviour (short-distance migrants: n = 21; long-
252 distance migrants: n = 13), and calculated the fixation index (F_{st}) as a measure of
253 differentiation (S2 Table). We observed relatively high mean values of nucleotide diversity in
254 both short- and long-distance migrants ($\pi_{autosome}$ = 0.012 ± 0.003(1 SD)) compared to other
255 Lepidoptera (34). Levels of the pairwise nucleotide divergence ($d_{XYautosome}$ = 0.012 ± 0.003)
256 between two phenotypes were elevated, reflecting the effect of high levels of diversity within
257 samples and stochasticity of the estimates. The level of genetic differentiation between
258 autosomes was low, but significant (pairwise F_{ST} = 0.001 ± 0.006, $p < 0.001$). A slightly
259 higher fixation index was observed for the Z-chromosome (pairwise F_{ST} = 0.009 ± 0.016),
260 which is expected given that the smaller effective population size makes the Z-chromosome
261 more susceptible to the effects of genetic drift. We repeated this analysis using groups
262 based on capture location (i.e., north vs. south of the Sahara), but found no substantial
263 differences between summary statistics (S2 Table). These findings indicate that *V. cardui*
264 that migrate short- versus long-distances have limited genetic differences and share similar
265 distribution of allele frequencies, and that the effective population size in general is large.
266 Secondly, we used PCA to investigate potential genetic structuring. After additional
267 filtering steps, we narrowed down our dataset to 813,810 high-quality single nucleotide
268 polymorphisms (SNPs). This analysis revealed that the majority of samples clustered
269 together with no separation between short- and long-distance migrants (Fig 2A), suggesting
270 an overall absence of population structure. However, within this general pattern, two
271 individuals stood out as outliers in the PCA; one showed differentiation primarily along the
272 first principal component (i.e., 19H128) and the other exhibited differentiation along the
273 second principal component (i.e., 19H115). Both of these outlier samples were collected in
274 September and were among the earliest arrivals to sub-Saharan Africa (9; S1 Table).
275 Moreover, these individuals had undertaken particularly long migratory journeys, with one

276 covering over 2,000 km across the Sahara from northwest Africa or Europe (19H115) and
277 the other over 1,000 km from Portugal, France, Sweden, or Finland (19H128) (S3 Fig). The
278 clear separation of these individuals in the PCA biplot could be indicative of population
279 structuring. We explored the contribution of specific SNPs to the first two principal
280 components with *pcadapt* and identified 924 outlier SNPs that made significant contributions
281 to the observed segregation of individuals. However, it is essential to note that drawing
282 conclusions about the potential processes underlying the observed differences based on
283 data from just two individuals is inherently limited. Nevertheless, we have provided a list of
284 candidate SNPs that may serve as a valuable starting point for future investigations (see
285 Data Availability Statement).

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

292

293 **No islands of genetic differentiation between short- and long-distance**
294 **migrants**

295 The assessment of genetic structure using genome-wide averages may obscure regional
296 differences in genetic differentiation. We therefore used window-based analysis of genetic
297 differentiation between populations (i.e., F_{ST} scans) to quantify genetic differentiation across
298 the genome and search for potential signatures of selection. Our analyses were based on
299 two sample groupings: (i) migratory phenotype (i.e., short- and long-distance migrants) and
300 (ii) capture location (i.e., north and south of the Sahara). Both comparisons revealed uniform
301 differentiation landscapes characterized by a low average pairwise F_{ST}. For autosomal
302 regions, the average F_{ST} was 0.001 and regional peaks had a maximum F_{ST} of 0.113 (Fig
303 2C; S4 Fig). For the Z-chromosome, the average pairwise F_{ST} was 0.009 with a maximum

304 peak of 0.072. This level of variation is consistent with stochasticity of the normalisation
305 and/or neutral variation, as has previously been demonstrated in simulation studies (35)
306 following simulation regimes applicable to our system. Additionally, none of the local F_{ST}
307 peaks were accompanied by a simultaneous elevation of the genetic divergence and drop in
308 diversity within each population extended over a substantial region, therefore not meeting
309 the expectations for signatures of divergent selection.
310

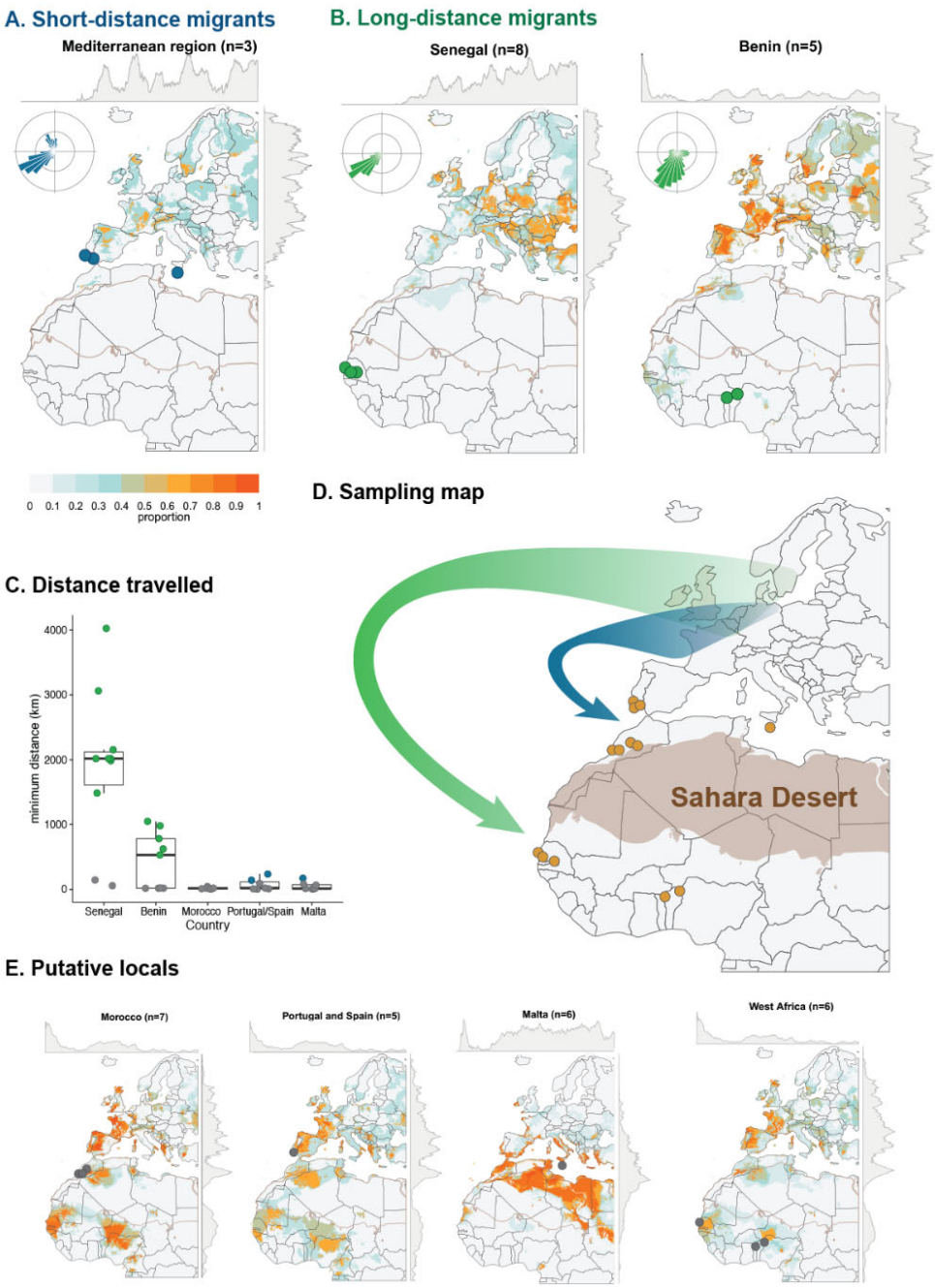

Fig 1. Summarised posterior probability surfaces and subsequent estimates of migration distance (i.e., minimum distance) and direction based on dual δ^2H and $^{87}Sr/^{86}Sr$-based geographic assignment. (A,B,E) Summary maps depict the proportion of individuals with a high probability of natal origin 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 (centre of plot) to the capture locations. Maps summarise 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)** Boxplot displaying the estimated migration distance (km) categorised by the country of capture. Individuals with a minimum distance of less than 100 km were considered putative locals. **(D)** Map indicating the capture locations both north (Morocco, Portugal and Spain, and Malta) and south (Senegal and Benin) of the Sahara. Butterflies captured south of the Sahara generally migrated longer distances (green arrow) compared to those captured north of the Sahara (blue arrow).

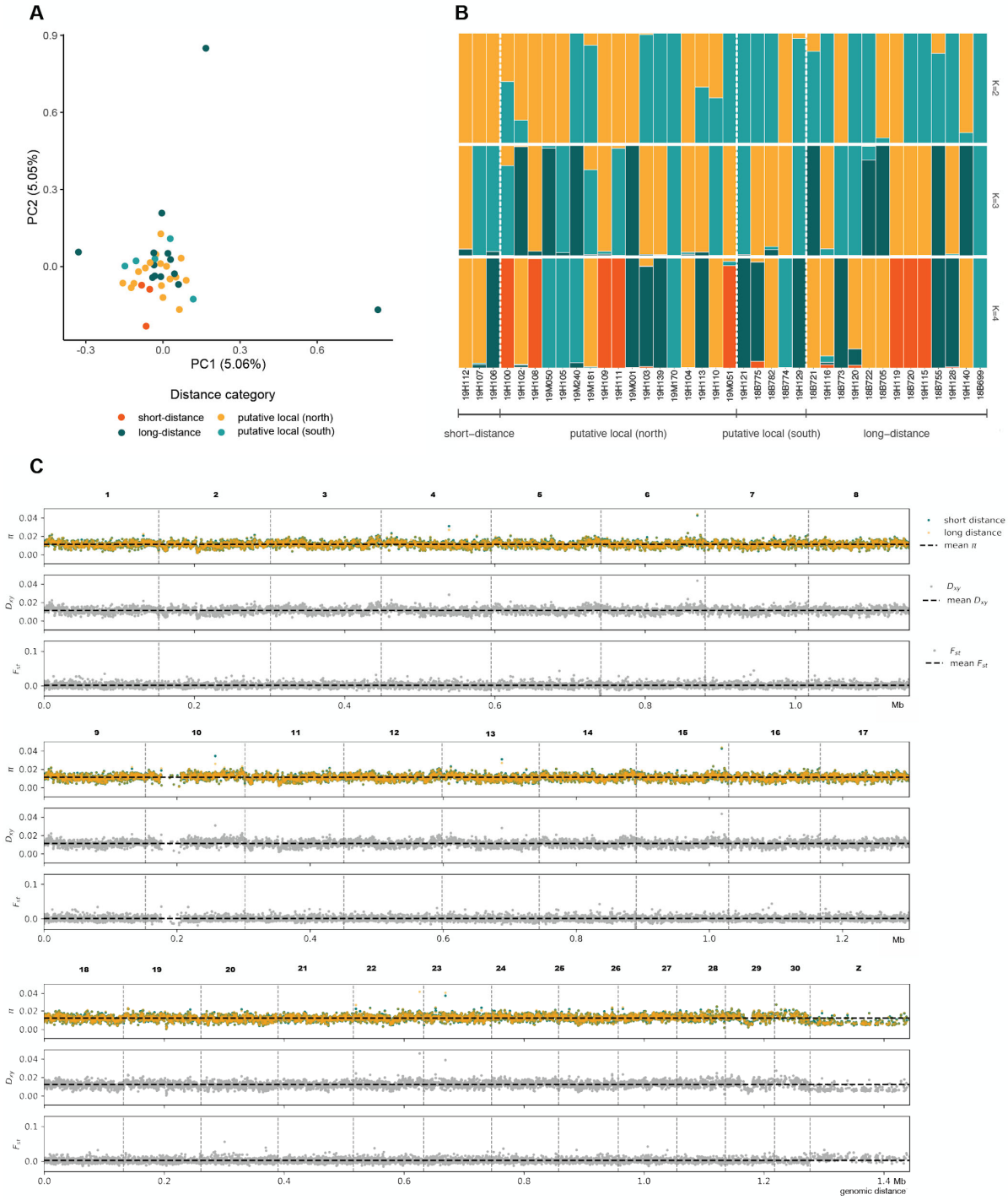

Fig 2. Lack of population structure and genetic differentiation between short- and long-distance migrants. (A) Principal component analysis (PCA) biplot illustrating the genetic variation in the samples; **(B)** Genomic admixture patterns based on assumptions of 2 to 4 ancestral populations (K = 2 to 4). No discernible structure was observed, with the optimal model demonstrating K = 1 (not visually depicted); **(C)** Genome scans showing within-population diversity (π), genetic differentiation (d_{XY}) and fixation index (F_{ST}) between individuals categorised as short- and long-distance migrants. The analysis was performed with sliding, non-overlapping windows of 20 kb. Mean values are indicated by a black dashed line.

Discussion

Contemporary evidence of variation in migratory behaviour

We used isotope-based estimates of migration distance to 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 had either undertaken shorter migrations or had isotopic signatures compatible with their capture area (Fig 1). Despite the increasing evidence that *V. cardui* was involved in regular trans-Saharan crossings as part of its annual migratory cycle, this evidence was limited to the monitoring of migratory arrivals in the sub-Sahara (9,10) and δ^2H-based geolocation (8,36). We have expanded this isotope-based evidence with the addition of ^{87}Sr/^{86}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 North American monarch butterflies *Danaus 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 (33,37,38). The variation in migration distance detected in our study appears to have a basis in behaviour rather than wing morphology, which may indicate limited selection on migration distance, as behaviour is thought to be more plastic than morphology (39).

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., 26,31,40), our results further illustrate that this

345 proxy is highly unreliable (37). For example, the butterfly with the highest estimated
346 migration distance, up to over 4,000 km (i.e., 18B721), had a low wing wear score of 2 (i.e.,
347 “few scales lost”; 32), which is often interpreted as belonging to a recently-emerged adult.
348 This finding is in agreement with Korkmaz et al. (41), who conjecture that high altitude,
349 gliding flight during migration results in less wing damage than low-altitude foraging and
350 breeding behaviour.
351

352 **Trans-Saharan panmixia**

353 Our genome-wide analysis revealed no evidence of genetic divergence between *V. cardui*
354 individuals migrating to north or south of the Sahara during autumn (Fig 2). Consequently,
355 the *V. cardui* population stretching from Northern Europe to the equatorial regions of sub-
356 Saharan Africa can be characterised as regionally panmictic, confirming the spatial extent of
357 the Afro-Palearctic migratory range. Our findings align well with previous studies using
358 mitochondrial DNA markers (42–44). Given that large biogeographic barriers, such as the
359 Mediterranean Sea and Sahara, do not seem to lead to population structure, it is likely that
360 panmixia will be found at even larger scales. However, *V. cardui* migration seems to
361 generally follow latitudinal trajectories, and extending to larger geographical scales might
362 reveal longitudinal structure and population differentiation (e.g., 45). Future studies will need
363 to assess the extent to which global panmixia occurs across the entire distributional range of
364 *V. cardui*, preferably by including genome-wide genetic markers from around the world,
365 including remote islands which sometimes host isolated populations that are phenotypically
366 and genetically differentiated (e.g., 46).

367 Special attention should be given to understanding how selection and local
368 adaptation act on migratory behaviour in insects, as well as the methods for detecting them.
369 In our study, although we did not detect clear signatures of selection or adaptation, we did
370 detect two outlier samples in the PCA. There could be multiple explanations for the patterns
371 we detected. First, unlike migratory vertebrates that complete round-trip migrations within a
372 single generation, insect migratory cycles are multi-generational. Multi-generational

373 migratory insects exhibit a complex reticular movement pattern driven by the drastic changes
374 in the location and size of the suitable breeding habitat over the annual migratory cycle
375 (9,11). Therefore, polygenic within-generation selection due to spatially- and temporally-
376 varying selection pressures could lead to allele frequency changes for a portion of the
377 population (47). However, the same alleles selected for in one generation are unlikely to be
378 favoured by selection pressures in the subsequent generation because of spatiotemporal
379 differences in abiotic and biotic conditions during the migratory journey and in the novel
380 breeding habitat, thus diluting the effect of within-generation local selection. Similar
381 frameworks have been proposed for some fish species which show signs of within-
382 generation local selection within panmictic populations (47–49). A similar scenario has been
383 outlined for *D. plexippus*, which experiences selection pressures on wing shape and wing
384 size during long-distance autumn migrations, only to have these pressures released in
385 subsequent generations (37).

386 Second, it is important to recognize that the exceptionally large effective population
387 size of *V. cardui* (50) and, consequently, weak genetic drift could impede our ability to detect
388 early, or complete, reproductive isolation between short- and long-distance migrants. In
389 these circumstances, many thousands of generations of reproductive isolation would need to
390 elapse before our analysis would be able to detect significant allele frequency shifts.
391 Therefore, it is possible that reproductive isolation has indeed occurred, but remains
392 undetected. Third, it is conceivable that our sampling strategy was unable to detect genetic
393 differentiation between groups of migrants. Notably, the most genetically differentiated
394 migrants (Fig 2A) were on the leading edge of arrivals to West Africa (September 2019) (9)
395 and likely originated in temperate Europe (S3 Fig). These early arrivals may have been
396 exposed to strong environmental cues in temperate Europe, where seasonal changes occur
397 earlier than in regions closer to the equator, prompting the butterflies to initiate their
398 migrations earlier in the season. Early migrants of *D. plexippus* have been shown to have
399 advantages over their later counterparts, such as enhanced immune defences (51).
400 Consequently, it is possible that *V. cardui* segregates temporally, rather than by distance.

401 However, it is more plausible that this temporal pattern results from within-generation
402 selection, which is likely to be diluted in the subsequent generation. Despite any ephemeral
403 within-generation local selection, we anticipate that the capacity to tolerate heterogeneous
404 environments relies primarily on phenotypic plasticity (52). Ultimately, interpreting our results
405 as indicative of regional panmixia remains the most parsimonious approach.

406 Panmixia across an Afro-Palearctic distributional range and regular trans-Saharan
407 crossings are exceptional phenomena for insects, as evidenced by the fact that the
408 Palearctic and Afrotropical regions are distinct biogeographic realms with unique faunal
409 compositions, sharing only a relatively small number of species. We found that neither
410 differences in migration distance nor the potential biogeographic barrier led to isolation,
411 selection, or significant allele frequency shifts in any genomic region (Fig 2C). This finding is
412 in contrast to numerous examples in vertebrate species (e.g., 15,17,53,54) and a few
413 instances in insect species (e.g., 55). Some of the species of insects shared between the
414 Palearctic and the Afrotropical regions are suspected to be migratory, particularly among the
415 Lepidoptera, Odonata, and Orthoptera (56,57), and thus it is likely that trans-Saharan gene
416 flow occurs in some of these other species as well. Continental-scale panmixia has been
417 demonstrated in other long-distance migratory insect species, suggesting that panmixia is a
418 prevalent feature among migratory species (but see 58). For example, whole-genome
419 resequencing of *D. plexippus* revealed high gene flow and low genetic differentiation
420 between butterflies located east and west of the Rocky Mountains, despite following
421 geographically separated migration routes (59). Regional panmixia has also been suggested
422 through microsatellite-based population genetics for the green darner dragonfly *Anax junius*
423 in eastern North America (60,61) and global panmixia has been inferred through
424 mitochondrial DNA analysis for the wandering glider dragonfly *Pantala flavescens* (62,63).

425 The widespread panmixia observed in many migratory insects raises questions about
426 the life history strategies promoting panmixia. First, most insect migrations are multi-
427 generational, featuring complex movement patterns driven by dynamic changes in breeding
428 habitat, promoting intra-population mixing. Additionally, the tendency of long-distance

429 migratory insects to engage in multiple matings, their lack of courtship behaviour, and their
430 practice of laying numerous eggs (e.g., 60,64,65) hints at an eponymous, yet underexplored,
431 link between panmictic tendencies and mating strategies in migratory insects. Large census
432 population sizes in migratory insects, accompanied by intermittent outbreaks, further drive
433 them towards panmixia. Outbreaks increase the number of successful migrants, potentially
434 eroding local adaptation (66). Outbreaks of *V. cardui* from regions with anomalously high
435 vegetation growth (67) may contribute to the observed lack of genetic structure. Similar
436 population dynamics observed in species like *Libytheana carinenta* (68), *A. junius* (69), and
437 some hoverflies (70) hint at outbreaks being pivotal in shaping the population structure of
438 many migratory insects.

439 The same life-history factors of migratory insects that appear to promote panmixia
440 among migratory insects (i.e., multi-generational migrations, an absence of assortative
441 mating, and large census population sizes with outbreak dynamics) are also associated with
442 large effective population sizes and high genetic diversity. The multi-generational annual
443 migratory cycle results in a greater number of mutations occurring within the population per
444 unit time, compared to species with longer generation times, and large census population
445 sizes are less susceptible to genetic drift, both of which encourage genetic diversity (71).
446 Indeed, our samples showed relatively high genetic diversity ($\pi_{autosome}$ = 0.012) compared to
447 other butterfly species (34). This observation aligns with the general trend that migratory
448 butterflies tend to possess higher levels of heterozygosity than sedentary species (50).
449 Consequently, this abundant genetic variation provides opportunities for at least some
450 individuals to acclimatise to the heterogeneous environmental conditions that are
451 encountered during their multi-generational migrations, thereby facilitating survival.

452 Due to the obligate, multi-generational migrations with low philopatry, random mating,
453 and large, outbreaking populations, local adaptation is implausible within the Afro-Palearctic
454 population of *V. cardui*. We suggest that rather than adapting to short or long-distance
455 migration, a plastic response to the environment determines the differences we observe in
456 migratory phenotype. It is also possible that the differences in migration distance observed

457 here could also be explained, in whole or in part, by extrinsic mechanisms, including
458 differences in parasite load (72), energy storage and metabolism (73), wind-assistance and
459 weather events (74), or even personality (75). However, studies on *D. plexippus* suggest that
460 while migratory state (i.e., resident vs migratory) is genetically determined (76), variation in
461 migration distance is more likely to be due to many loci of small effect or differential gene
462 expression (59), likely controlled by epigenetic mechanisms (4,18,77). Variation in migratory
463 behaviour is likely a response to multiple environmental cues such as resource availability,
464 photoperiod, temperature, humidity, and host plant quality. Recent experiments have
465 demonstrated differences in *V. cardui* gene expression, methylation, and gene regulation
466 due to the density of larval conspecifics, larval food availability, and host plant availability for
467 oviposition (78,79), suggesting a strong response to environmental cues. The transcriptional
468 basis of differences in migration distance has been explored in the moth *Helicoverpa*
469 *armigera* and in *D. plexippus*, yielding promising candidate genes (59,80). Future controlled
470 laboratory studies with *V. cardui* are needed to disentangle the environmental triggers and
471 molecular pathways of phenotypic plasticity in migration distance.
472

473 **Conclusion**

474 Here, the combination of isotope-based phenotyping of migration distance and
475 comprehensive whole-genome resequencing provided a unique opportunity to investigate
476 phenotype-genotype associations of migratory behaviour in insects, a challenging endeavour
477 until now. We found distinct differences in migration distance, with most individuals captured
478 in Benin and Senegal exhibiting long-distance migrations of up to 4,000 km, traversing the
479 Mediterranean Sea and Sahara from temperate Europe. In contrast, *V. cardui* captured in
480 the circum-Mediterranean region at the same time either undertook short-distance
481 migrations or displayed isotopic signatures indicative of local origins. These differences in
482 migratory behaviour did not have a clear association with wing morphology or wing wear.
483 Genome-wide comparisons between these migratory groups found negligible genetic
484 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 multi-generational 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

1. Sample collection

Wild, adult *V. cardui* were captured from six countries located both north and south of the Sahara (S1 Table). 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 used behavioural, demographic and timing evidence to maximise the probability of targeting migratory individuals at these locations. First, we preferentially sampled butterflies with worn wings (i.e., higher wing wear scores) and showing reproductive behaviour (i.e., hilltopping or oviposition) because *V. cardui* typically delay their reproductive behaviour until migration has been completed (81). Second, demographic data in some of the sampling locations supported the arrival of migrants. A monitoring station in Senegal saw a rising number of *V. cardui* adults in the autumn (August through October) during 2018 and 2019, 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 (9). Similarly, a monitoring station in Benin saw high numbers of adults in September and October, followed, but not preceded, by high larval counts in October and November (9). Besides this demographic evidence, the time of the year when these butterflies were collected matches the first estimated arrival of migrants at the sampling locations. South of the Sahara, monitoring data does not indicate the presence of *V. cardui* year-round (9). North of the Sahara, much of the circum-Mediterranean region is unsuitable for *V. cardui* development in July, August, and

September (Morocco) due to high temperatures and low host plant availability (40), suggesting that these individuals represent the first arrival of migrants in the region (9, 40).

2. Isotopic analysis

2.1 Hydrogen isotope values

Prior to δ^2H analysis, a forewing from each *V. cardui* was soaked, with agitation, in three successive baths (1 h, 30 min., 10 min.) of 2:1 chloroform:methanol solution to remove surficial dust and lipids, which are known to introduce error into δ^2H measurements (82,83). Samples were cut from the same membranous region of the forewing to reduce intra-individual variation from pigmentation (82) and wing veins (24). Samples were then weighed (0.162 ± 0.009 mg) into silver capsules and surface moisture was dried in a 40 °C oven. Non-exchangeable δ^2H values were determined at the Ján Veizer Stable Isotope Laboratory (University of Ottawa, ON, Canada) using a comparative equilibrium approach (84). All measurements were taken using high temperature (1400 °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, Germany) via a Conflow IV interface (Thermo Finnigan, Germany). A three-point calibration was used to calibrate the sample δ^2H values: CBS (caribou hoof; -157 ± 0.9‰; 85), KHS (kudu horn; -35.3 ± 1.1‰; 85), and USGS43 (human hair; -44.4 ± 2.0‰; 86). Internal standards were measured to assess the quality of the measurements, including one keratin reference standard, USGS42 (human hair; measured: -75.3 ± 0.5‰, n = 4; standard: -72.9 ± 2.2‰ (86)), and two in-house chitin standards, ground and homogenised *Lymantria dispar* (measured: -64.4 ± 1.8‰, n = 6; long-term average: -64 ± 0.8‰) and Alfa Aesar chitin (measured: -22.8 ± 0.7‰, n = 4; long-term average: -22 ± 1.2‰). Based on within-run replicates of the internal standards and repeated sample measurements, precision is estimated to be about ± 2‰. All reported δ^2H values are normalised to the Vienna Standard Mean Ocean Water - Standard Light Antarctic Precipitation (VSMOW-SLAP) standard scale.

539 **2.2 Strontium isotope ratios**

540 To prepare for $^{87}Sr/^{86}Sr$ analysis, a single forewing from each butterfly was cleaned using
541 pressurised nitrogen gas for 2 min. at 69 kPa to remove any surface contaminants (e.g.,
542 dust; 26). The wings were then digested in 1 mL of 16 M HNO_3 (distilled TraceMetal™
543 Grade; Fisher Chemical, Canada) using microwave digestion (Anton Paar Multiwave 7000;
544 Austria). The samples were heated from ambient temperature to 250 °C at a steady rate
545 over 20 min. and then left at 250 °C for 15 min. in a pressurised chamber. An aliquot of 200
546 µL from each sample was separated and diluted to 2 mL of 2% v/v HNO_3 and analysed for
547 Sr content via inductively coupled plasma mass spectrometry (ICP-MS; Agilent 8800 ICP-
548 QQQ, Agilent Technologies Inc., CA, USA) at the Department of Earth and Environmental
549 Science, University of Ottawa (ON, Canada). Calibration standards were prepared using
550 single element certified standards purchased from SCP Science (Montreal, Canada).
551 Samples had an average of 3.1 ± 1.7 ng Sr per mg of wing tissue (n = 40).

552 The remaining aliquot of sample digest was dried down and re-dissolved in 0.5 mL 6
553 M HNO_3. The separation of Sr was processed in a 100 µL microcolumn loaded with Sr-spec
554 Resin (100 - 150 µm; Eichrom Technologies, LLC). The matrix was rinsed out using 6 M
555 HNO_3, and Sr was collected with 0.05 M HNO_3. To ensure complete separation, the
556 microcolumn process was repeated. After separation, eluates were dried and re-dissolved in
557 200 µL 2% v/v HNO_3 for $^{87}Sr/^{86}Sr$ analysis.

558 The $^{87}Sr/^{86}Sr$ analysis was performed at the Pacific Centre for Isotopic and
559 Geochemical Research (University of British Columbia, BC, Canada) using a Nu-Plasma II
560 high-resolution MC-ICP-MS (Nu Instruments) coupled to a desolvating nebulizer (Aridus II,
561 CETAC Technologies). The isotopes ^{84}Sr, ^{86}Sr, and ^{87}Sr have isobaric interferences from
562 ^{84}Kr, ^{86}Kr, and ^{87}Rb respectively, and were corrected for using the ^{85}Rb and ^{83}Kr signals.
563 Instrumental mass fractionation was corrected by normalising $^{86}Sr/^{88}Sr$ to 0.1194 using the
564 exponential law (87). Procedural blanks were negligible. The reproducibility of the $^{87}Sr/^{86}Sr$
565 measurement for 5 ppb NIST SRM987 is 0.71025 ± 0.00009 (1 SD, n = 138) and for 1.4 ppb

NIST SRM987, 0.71019 ± 0.00011 (n = 48). A matrix-matched chitin internal standard, 5 ppb Alfa Aesar chitin, was also used (0.713959 ± 0.00009; n = 3).

2.3 Geographic assignment

Continuous-surface isotope-based geographic assignment was performed via the assignR package (88) in R v4.1.0 (89). The origins of the unknown origin *V. cardui* 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 (90), and (ii) a regional bioavailable strontium isoscape (91). The most likely natal origin of each individual was estimated using δ^2H isotopes alone, $^{87}Sr/^{86}Sr$ isotopes alone, and using both isotopes together (see Data Availability Statement). The 2:1 odds ratio was chosen, in keeping with previous studies (e.g. 31), to create binary surfaces from the posterior probability surfaces of the dual δ^2H and $^{87}Sr/^{86}Sr$-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 summarise 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.

2.4 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 (92). 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 signatures for 250 synthetic migrants, performed dual δ^2H and $^{87}Sr/^{86}Sr$ 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 kilometres (S1 Fig). However, the “minimum distance” metric consistently underestimated

594 migration distance for our study area and, therefore, acts as a conservative estimate of
595 migration distance (S1 Fig). Therefore, in this study, we used the minimum distance metric
596 to estimate the distance travelled. This metric measures the shortest distance between the
597 capture location and an isopleth threshold, here defined as the top third of the probability
598 distribution (0.3 isopleth; e.g., 93,94).

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

606

607 **3. Geometric morphometric analysis**

608 Differences in migration distance or trajectories could potentially be explained by intrinsic
609 factors such as sex, body size, or wing shape. To test for differences in wing shape,
610 photographs of each sample's forewing were taken and used for morphometrics analysis.
611 Since we preferentially chose samples that had higher wing wear scores, extensive damage
612 to the wings made landmarking of many samples impossible. Therefore, only 28 of the 40
613 samples were landmarked using TPSDig2 v.2.32 (95) using 12 conventional landmarks for
614 Nymphalid butterflies (e.g., 96,97). Using the R package geomorph (98), a generalised
615 procrustes analysis removed the translational, rotational, and scaling variation of the shape.
616 A procrustes regression was used to test the effects on wing shape of sex, month of
617 collection, estimated migration distance, and butterfly wing size.

618

619 **4. Genomic analysis and genotyping**

620 **4.1 Whole genome resequencing**

621 Library preparation for all 40 samples was done using TruSeq PCR-free kits (Illumina, San
622 Diego, USA) at the Institut Botànic de Barcelona (IBB – CSIC, Barcelona, Spain). Whole-
623 genome resequencing was performed with paired-end 2x151 bp reads on one Illumina
624 NovaSeq 6000 S4 lane at the National Genomics Infrastructure (NGI; Stockholm, Sweden)
625 to generate a mean depth of coverage per individual of about 48 reads.
626
627 **4.2 SNP calling**
628 A modified Sarek Nextflow pipeline (99) was used for preprocessing of the whole genome
629 resequencing data and for detecting SNPs. At the first step of the Sarek pipeline, we
630 evaluated the quality of the raw reads with FastQC (version 0.11.9; 100). Reads were
631 stringently trimmed using fastp (101,102) to remove (i) sequences from the 5` end (12 bp)
632 and 3` end (3 bp) of each read, (ii) adaptor sequences, (iii) low quality reads (overall Phred
633 score < 30), and (iv) short reads (< 30 bp). Read quality after trimming was re-evaluated
634 using FastQC. At the next step, a mean of 96.6% of the reads were mapped to the reference
635 genome (103) using BWA-MEM (version 0.7.17; 104) and then sorted and indexed using
636 SAMtools (version 1.9; 105). Optical and PCR duplicates were marked using GATK
637 MarkDuplicates (version 4.1.4.1; 106). These duplicates, as well as sequences with low
638 mapping quality (MAPQ < 30), were removed using SAMtools. Since prior SNP data for *V.*
639 *cardui* were unavailable, the base calibration (BQSR) step module in the Sarek pipeline was
640 skipped.
641 The first steps of variant calling were performed using GATK Best Practices (107), as
642 implemented through the Sarek pipeline. In order to decrease computational time, the Sarek
643 pipeline calls each chromosome separately. MultiQC (version 1.8) was used to assess the
644 quality of the output from all steps of the pipeline. The resulting per chromosome vcf files
645 included from 1.3 to 3 million unfiltered SNPs (78.9 million SNPs total). The last step of
646 variant calling was performed outside of the pipeline to accommodate for invariant site
647 calling (-all-sites option in GenotypeGVCFs), which is necessary for reliable downstream
648 estimation of population genetic summary statistics, such as π and d_{XY}. This variant filtering

649 included multiple steps: (i) hard filtering, (ii) removing SNPs overlapping with TE annotation,
650 (iii) separate quality and coverage filtering for variant and invariant sites, and (iv) merging of
651 variant and invariant site vcfs for the F_{ST} estimation. Hard filtering, which is meant to
652 substitute for the variant quality score recalibration (VQSR) step in GATK Best Practices
653 (107), was performed on raw vcf files, which included both variant and invariant sites. Using
654 bcftools v1.15 (108), we filtered for variant quality by depth (QD < 2.0), Fisher strand bias
655 (FS > 60.0), root mean square mapping quality (MQ < 40.0), u-based z-approximation from
656 the Rank Sum Test for mapping qualities (MQRankSum < -12.5), and u-based z-
657 approximation from the Rank Sum Test for site position (ReadPosRankSum < -8.0). We
658 proceeded to remove indels and multivariate sites (--exclude-types indels, other --max-
659 alleles 2) and reindexed the resulting vcf file. Next, the coordinates of transposable elements
660 and simple repeats were obtained with a custom bash script from a previously published
661 annotation of transposable elements (109), and used to remove SNPs appearing in these
662 regions.

663 The quality of the resulting files (one vcf file per chromosome) was evaluated using
664 summary statistics implemented with vcftools (i.e., allele frequency (--freq2), mean depth (--
665 depth), mean per site depth (--site-mean-depth), per site quality (--site-quality), per sample
666 missingness (--missing-indv), per site missingness (--missing-indv), heterozygosity
667 estimation (--het); 110). Two individuals were removed due to quality issues (Sample IDs:
668 19H137 and 19H114). Variant and invariant sites were split into two files for downstream
669 analysis. Informed by the summary statistics, we applied filters for quality and depth to the
670 variant sites vcf file (i.e., minor allele frequency (--max-maf 0), mean depth value (--min-
671 meanDP 30), mean depth value (--max-meanDP 80)). We also applied additional filters to
672 the invariant sites vcf file (i.e., quality value (--minQ 30), mean depth value (--min-meanDP
673 30)). Separate files with both variant and invariant calls were kept as separate chromosomes
674 to decrease computational time. The W chromosome was also removed from the analysis. A
675 region with divergent values of populational statistics was detected in the first third of
676 chromosome 10 corresponding to a large area of transposable element proliferation (Fig 2C;

677 109). Since this region likely does not carry information about genetic structure it was also
678 filtered from the population structure analyses. On a technical note, it is important to mention
679 that this region was retained after stringent SNP filtration, and future research on *V. cardui*
680 should be aware that this region may necessitate special treatment (e.g., manual filtering).
681
682 **4.3 Population genetic analysis**
683 For PCA and admixture analysis, variant files per chromosome were concatenated. To
684 investigate genetic variation between individuals, a PCA was performed using PLINK v1.90
685 b 4.9 (Fig 2A). For this analysis, we additionally filtered out singleton SNP sites (i.e., MAF =
686 0.03), and SNPs that were in linkage (i.e., SNPs with r^2 values > 0.1 using sliding windows of
687 50 SNPs with a step size of 10 SNPs); a total of 813,810 high quality variants remained after
688 filtering. Outliers in the first two principal components were detected using the R package
689 pcadapt (111) using Bonferroni-corrected p-values ($\alpha = 0.05$). The PCA was repeated
690 excluding the two outliers, but no additional structure was detected (S5 Fig).
691 Reverse migrations from sub-Saharan Africa to the Maghreb are thought to occur in
692 the autumn, even while southwards trans-Saharan migrations are still underway (8). All of
693 the samples captured in Morocco were designated as putative locals, but the posterior
694 probability surfaces of some showed a possibility of originating south of the Sahara (Fig 1E).
695 To test whether these samples could have biased our results, we repeated the PCA without
696 samples collected in Morocco (S5 Fig). The resulting PCA was based on 647,906 SNPs and
697 showed a single cluster with the same outlier individuals as the original PCA.
698 After excluding SNPs on the Z chromosome and in coding sequences, genetic
699 structure was also assessed with a likelihood method, the sparse non-negative matrix
700 factorization clustering algorithm, SNMF, in the R package LEA v3.6.0 (112). Each number
701 of population clusters (K) from 1 to 5 was run independently with 5 repetitions with the
702 regularisation parameter set to 100, after which cross-validation errors calculated using the
703 cross-entropy criterion were compared. This process was repeated using a Bayesian
704 framework via fastSTRUCTURE (113) using the simple allele frequency prior with 5

705 repetitions. The number of population clusters from 1 to 5 that best explained the observed
706 structure was chosen with the chooseK.py tool.
707
708 **4.4 Genomic scan for selection**
709 To detect any fine-scale regions of differentiation along the genome, a genome scan was
710 performed on groupings based on capture location (i.e., north vs. south of the Sahara) and
711 migratory phenotype (i.e., long vs. short migration distance). For the grouping based on
712 migratory phenotype, individuals from south of the Sahara that were putative locals (n = 6)
713 were removed, while putative locals from north of the Sahara (n = 19) were included with the
714 short-distance migrants (n = 3) to balance the sample size. For genome scans we retained
715 SNPs falling into coding regions and having minor allele frequencies. We used pixy v1.2.7
716 (114,115) to calculate nucleotide diversity (π), absolute divergence (d_{XY}), and genetic
717 differentiation (F_{ST}; Hudson method (-hudson)) in 20 kb non-overlapping windows using all
718 sites. Windows with less than 10% of sites represented (i.e., < 2000 bp) were removed from
719 the analysis. Resulting per window statistics were summarized using custom python scripts.
720 For low levels of F_{ST}, we tested if the index significantly differed from 0 using a t-test in the
721 python scipy package (p-value: 1.3e-216).
722
723 ## Acknowledgements
724 This study was funded by grant 2018-00738 of the New Frontiers in Research Fund
725 (Government of Canada) to GT and CB; by grants PID2020-117739GA-I00 MCIN / AEI /
726 10.13039/501100011033 and WW1-300R-18 (National Geographic Society) to GT; and by
727 grant LINKA20399 from the CSIC iLink program to NB, CB and GT. MR was supported by
728 the Queen Elizabeth II Graduate Scholarship in Science and Technology (QEII-GSST), the
729 Ontario Graduate Scholarship, and the 2021 ORIGIN project (NSF award DBI-1565128). NB
730 is supported by a research grant from the Swedish Research Council FORMAS (grant 2019-
731 00670). DS is supported by the Thurberg postdoctorate fellowship program of Swedish
732 Collegium for Advanced Science (Knut and Alice Wallenberg Foundation). The authors

acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and 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 (NAISS) and the Swedish National Infrastructure for Computing (SNIC) in Uppsala, partially funded by the Swedish Research Council through grant agreement no. 2022-06725 and no. 2018-05973. We are grateful to members of the Backström lab and Sean Stankowski for insightful discussions about the genome scan analysis. We would like to thank Smita Mohanty for help with the ICP-MS, Kerry Klassen and Paul Middlestead at the Ján Veizer Stable Isotope Laboratory for help with the hydrogen isotope analysis, and Kathy Gordon at the Pacific Centre for Isotopic and Geochemical Research for help with the strontium isotope analysis. We would like to 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. We would also like to thank Andrea Contina, Gabriel Bowen, and Michael Wunder for sharing code and insight relating to the migration distance estimates.

Data Availability Statement

The raw sequence data is deposited in the European Nucleotide Archive (ENA) 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/?view_only=6dc9f9bd496f4b538b0708d5f13cfce8).

References

1. Dingle H, Drake VA. What is migration? BioScience. 2007;57(2):113–21. doi:

10.1641/B570206

2. Chapman JW, Reynolds DR, Wilson K. Long-range seasonal migration in insects: Mechanisms, evolutionary drivers and ecological consequences. Ecology Letters. 2015;18(3):287–302. doi: 10.1111/ele.12407

3. Dingle H. Evolutionary genetics of migration. In: Migration: The biology of life on the move. Oxford University Press; 2014. p. 248–68.

4. Merlin C, Liedvogel M. The genetics and epigenetics of animal migration and orientation: Birds, butterflies and beyond. J Exp Biol. 2019;222:1–12. doi: 10.1242/jeb.191890

5. Shaw AK. Causes and consequences of individual variation in animal movement. Mov Ecol. 2020;8(1):1–12. doi: 10.1186/s40462-020-0197-x

6. Shields O. World distribution of the *Vanessa cardui* group (Nymphalidae). J Lepidopterists Soc. 1992;46(3):235–8.

7. Williams CB. The migrations of the painted lady butterfly, *Vanessa cardui* (Nymphalidae), with special reference to North America. Lepidopterists Soc. 1970;24(3).

8. Stefanescu C, Soto DX, Talavera G, Vila R, Hobson KA. Long-distance autumn migration across the Sahara by painted lady butterflies: Exploiting resource pulses in the tropical savannah. Biol Lett. 2016;12:1–4. doi: 10.1098/rsbl.2016.0561

9. Talavera G, García-Berro A, Talla VNK, Ng'iru I, Bahleman F, Kébé K, et al. The Afrotropical breeding grounds of the Palearctic-African migratory painted lady butterflies (*Vanessa cardui*). Proc Natl Acad Sci. 2023;120(16):e2218280120. doi: 10.1073/pnas.2218280120

10. Talavera G, Vila R. Discovery of mass migration and breeding of the painted lady butterfly *Vanessa cardui* in the Sub-Sahara: The Europe-Africa migration revisited. Biol J Linn Soc. 2016;120(2):274–85. doi: 10.1111/bij.12873

11. Menchetti M, Guéguen M, Talavera G. Spatio-temporal ecological niche modelling of multigenerational insect migrations. Proc R Soc B Biol Sci. 2019;286(1910):20191583.

doi: 10.1098/rspb.2019.1583

12. Stefanescu C, Páramo F, Åkesson S, Alarcón M, Ávila A, Brereton T, et al. Multi-generational long-distance migration of insects: Studying the painted lady butterfly in the Western Palaearctic. Ecography. 2013;36(4):474–86. doi: 10.1111/j.1600-0587.2012.07738.x

13. Turbek SP, Scordato ESC, Safran RJ. The role of seasonal migration in population divergence and reproductive isolation. Trends Ecol Evol. 2018;33(3):164–75. doi: 10.1016/j.tree.2017.11.008

14. Delmore K, Fox JW, Irwin DE. Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. Proc R Soc B Biol Sci. 2012;279(1747):4582–9. doi: 10.1098/rspb.2012.1229

15. Delmore K, Illera JC, Pérez-Tris J, Segelbacher G, Lugo Ramos JS, Durieux G, et al. The evolutionary history and genomics of European blackcap migration. eLife. 2020;9:e54462. doi: 10.7554/eLife.54462

16. Mueller JC, Pulido F, Kempenaers B. Identification of a gene associated with avian migratory behaviour. Proc R Soc B Biol Sci. 2011;278(1719):2848–56. doi: 10.1098/rspb.201

17. Sokolovskis K, Lundberg M, Åkesson S, Willemoes M, Zhao T, Caballero-Lopez V, et al. Migration direction in a songbird explained by two loci. Nat Commun. 2023;14(1):165. doi: 10.1038/s41467-023-35788-7

18. Liedvogel M, Åkesson S, Bensch S. The genetics of migration on the move. Trends Ecol Evol. 2011;26(11):561–9. doi: 10.1016/j.tree.2011.07.009

19. Knight SM, Pitman GM, Flockhart DTT, Norris DR. Radio-tracking reveals how wind and temperature influence the pace of daytime insect migration. Biol Lett. 2019;15(7): 20190327. doi: 10.1098/rsbl.2019.0327

20. Menz MHM, Scacco M, Bürki-Spycher HM, Williams HJ, Reynolds DR, Chapman JW, et al. Individual tracking reveals long-distance flight-path control in a nocturnally migrating moth. Science. 2022;377(6607):764–8. doi: 10.1126/science.abn1663

21. Wikelski M, Moskowitz D, Adelman JS, Cochran J, Wilcove DS, May ML. Simple rules guide dragonfly migration. Biol Lett. 2006;2:325–9. doi: 10.1098/rsbl.2006.0487

22. Flockhart DTT, Kyser TK, Chipley D, Miller NG, Norris DR. Experimental evidence shows no fractionation of strontium isotopes ($^{87}Sr/^{86}Sr$) among soil, plants, and herbivores: Implications for tracking wildlife and forensic science. Isotopes Environ Health Stud. 2015;51(3):372–81. doi: 10.1080/10256016.2015.1021345

23. Hobson KA, Wassenaar LI, Taylor OR. Stable isotopes (δD and $\delta^{13}C$) are geographic indicators of natal origins of monarch butterflies in eastern North America. Oecologia. 1999;120(3):397–404.

24. Lindroos EE, Bataille CP, Holder PW, Talavera G, Reich MS. Temporal stability of $\delta^{2}H$ in insect tissues: Implications for isotope-based geographic assignments. Front Ecol Evol. 2023;11:1060836. doi: 10.3389/fevo.2023.1060836

25. Reich MS, Kindra M, Dargent F, Hu L, Flockhart DTT, Norris DR, et al. Metals and metal isotopes incorporation in insect wings: Implications for geolocation and pollution exposure. Front Ecol Evol. 2023;11:1085903. doi: 10.3389/fevo.2023.1085903

26. Reich MS, Flockhart DTT, Norris DR, Hu L, Bataille CP. Continuous-surface geographic assignment of migratory animals using strontium isotopes: A case study with monarch butterflies. Methods Ecol Evol. 2021;00:1–13. doi: 10.1111/2041-210X.13707

27. Wassenaar LI, Hobson KA. Natal origins of migratory monarch butterflies at wintering colonies in Mexico: New isotopic evidence. Proc Natl Acad Sci U S Am. 1998;95:15436–9.

28. Bataille CP, Crowley BE, Wooller MJ, Bowen GJ. Advances in global bioavailable strontium isoscapes. Palaeogeogr Palaeoclimatol Palaeoecol. 2020;555:109849. doi: 10.1016/j.palaeo.2020.109849

29. Crowley BE, Bataille CP, Haak BA, Sommer KM. Identifying nesting grounds for juvenile migratory birds with dual isotope: An initial test using North American raptors. Ecosphere. 2021;12(10):0–21. doi: https://doi.org/10.1002/ecs2.3765

844 30. Kruszynski C, Bailey LD, Courtiol A, Bach L, Bach P, Göttsche M, et al. Identifying
845 migratory pathways of Nathusius' pipistrelles (*Pipistrellus nathusii*) using stable
846 hydrogen and strontium isotopes. Rapid Commun Mass Spectrom. 2021;35(6):1–11.
847 doi: 10.1002/rcm.9031
848 31. Flockhart DTT, Wassenaar LI, Martin TG, Hobson KA, Wunder MB, Norris DR.
849 Tracking multi-generational colonization of the breeding grounds by monarch butterflies
850 in eastern North America. Proc R Soc B. 2013;280(20131087). doi:
851 10.1098/rspb.2013.1087
852 32. Miller NG, Wassenaar LI, Hobson KA, Norris R. Migratory connectivity of the monarch
853 butterfly (*Danaus plexippus*): Patterns of spring re-colonization in eastern North
854 America. PLoS ONE. 2012;7(3):e31891. doi: 10.1371/journal.pone.0031891
855 33. Altizer S, Davis A. Populations of monarch butterflies with different migratory behaviors
856 show divergence in wing morphology. Evolution. 2010;64(4):1018–28. doi:
857 10.1111/j.1558-5646.20
858 34. Mackintosh A, Laetsch DR, Hayward A, Charlesworth B, Waterfall M, Vila R, et al. The
859 determinants of genetic diversity in butterflies. Nat Commun. 2019;10(1):3466. doi:
860 10.1038/s41467-019-11308-4
861 35. Lohse K. Come on feel the noise – from metaphors to null models. J Evol Biol.
862 2017;30(8):1506–8. doi: 10.1111/jeb.13109
863 36. Talavera G, Bataille C, Benyamini D, Gascoigne-Pees M, Vila R. Round-trip across the
864 Sahara: Afrotropical painted lady butterflies recolonize the Mediterranean in early
865 spring. Biol Lett. 2018;14:20180274. doi: 10.1098/rsbl.2018.0274
866 37. Flockhart DTT, Fitz-gerald B, Brower LP, Derbyshire R, Altizer S, Hobson KA, et al.
867 Migration distance as a selective episode for wing morphology in a migratory insect.
868 Mov Ecol. 2017;5(7):1–9. doi: 10.1186/s40462-017-0098-9
869 38. Li Y, Pierce AA, de Roode JC. Variation in forewing size linked to migratory status in
870 monarch butterflies. Anim Migr. 2016;3:27–34. doi: 10.1515/ami-2016-0003
871 39. West-Eberhard MJ. Phenotypic plasticity and the origins of diversity. Annu Rev Ecol

872 Syst. 1989;20(1):249–78.
873 40. Stefanescu C, Puig-Montserrat X, Samraoui B, Izquierdo R, Ubach A, Arrizabalaga A.
874 Back to Africa: Autumn migration of the painted lady butterfly *Vanessa cardui* is timed
875 to coincide with an increase in resource availability. Ecol Entomol. 2017;42(6):737–47.
876 doi: 10.1111/een.12442
877 41. Korkmaz R, Rajabi H, Eshghi S, Gorb SN, Büscher TH. The frequency of wing damage
878 in a migrating butterfly. Insect Sci. 2022;0:1-11. doi: 10.1111/1744-7917.13153
879 42. Dapporto L, Cini A, Vodă R, Dincă V, Wiemers M, Menchetti M, et al. Integrating three
880 comprehensive data sets shows that mitochondrial DNA variation is linked to species
881 traits and paleogeographic events in European butterflies. Mol Ecol Resour.
882 2019;19(6):1623–36. doi: 10.1111/1755-0998.13059
883 43. Pfeiler E, Markow TA. Population connectivity and genetic diversity in long-distance
884 migrating insects: Divergent patterns in representative butterflies and dragonflies. Biol J
885 Linn Soc. 2017;122:479–86.
886 44. Vodă R, Dapporto L, Dincă V, Shreeve TG, Khaldi M, Barech G, et al. Historical and
887 contemporary factors generate unique butterfly communities on islands. Sci Rep.
888 2016;6(28828):1-11. doi: 10.1038/srep28828
889 45. Singh VK, Kumar S, Awasthi S, Gupta SK. Phylogeography and population expansion
890 of Western Himalayan highly-erratically migratory painted lady butterfly (Lepidoptera:
891 Nymphalidae: *Vanessa cardui*). J Asia-Pac Entomol. 2021;24(4):1116–21. doi:
892 10.1016/j.aspen.2021.10.005
893 46. Alvial IE, Vargas HA, Marinov M, Esquivel C, Araya J, Araya-Donoso R, et al. Isolation
894 on a remote island: Genetic and morphological differentiation of a cosmopolitan
895 odonate. Heredity. 2019;122(6):893–905. doi: 10.1038/s41437-018-0165-z
896 47. Ehrlich MA, Wagner DN, Oleksiak MF, Crawford DL. Polygenic selection within a single
897 generation leads to subtle divergence among ecological niches. Genome Biol Evol.
898 2021;13(2). doi: 10.1093/gbe/evaa257
899 48. Als TD, Hansen MM, Maes GE, Castonguay M, Riemann L, Aarestrup K, et al. All

900 roads lead to home: Panmixia of European eel in the Sargasso Sea. Mol Ecol.
901 2011;20(7):1333–46. doi: 10.1111/j.1365-294X.2011.05011.x
902 49. Pujolar JM, Jacobsen MW, Als TD, Frydenberg J, Munch K, Jónsson B, et al. Genome-
903 wide single-generation signatures of local selection in the panmictic European eel. Mol
904 Ecol. 2014;23(10):2514–28. doi: 10.1111/mec.12753
905 50. García-Berro A, Talla V, Vila R, Wai HK, Shipilina D, Chan KG, et al. Migratory
906 behaviour is positively associated with genetic diversity in butterflies. Mol Ecol.
907 2023;32(3):560–74. doi: 10.1111/mec.16770
908 51. Satterfield DA, Wright AE, Altizer S. Lipid reserves and immune defense in healthy and
909 diseased migrating monarchs *Danaus plexippus*. Curr Zool. 2013;59(3):393–402. doi:
910 10.1093/czoolo/59.3.393
911 52. Enbody ED, Pettersson ME, Sprehn CG, Palm S, Wickström H, Andersson L.
912 Ecological adaptation in European eels is based on phenotypic plasticity. Proc Natl
913 Acad Sci. 2021;118(4):e2022620118. doi: 10.1073/pnas.2022620118
914 53. Gu Z, Pan S, Lin Z, Hu L, Dai X, Chang J, et al. Climate-driven flyway changes and
915 memory-based long-distance migration. Nature. 2021;591:259–564. doi:
916 10.1038/s41586-021-03265-0
917 54. Thompson NF, Anderson EC, Clemento AJ, Campbell MA, Pearse DE, Hearsey JW, et
918 al. A complex phenotype in salmon controlled by a simple change in migratory timing.
919 Science. 2020;370(6516):609–13. doi: 10.1126/science.aba9059
920 55. Niitepõld K, Smith AD, Osborne JL, Reynolds DR, Carreck L, Martin AP, et al. Flight
921 metabolic rate and *Pgi* genotype influence butterfly dispersal rate in the field. Ecology.
922 2009;90(8):2223–32.
923 56. Chelli A, Moulaï R. Ecological characterization of the odonatofauna in lotic and lentic
924 waters of northeast Algeria. Ann Société Entomol Fr NS. 2019;55(5):430–45. doi:
925 10.1080/00379271.2019.1660215
926 57. Pedgley DE, Reynolds DR, Tatchell GM. Long-range insect migration in relation to
927 climate and weather: Africa and Europe. In: Insect migration: tracking resources

928 through space and time. Cambridge University Press; 1995. p. 3–29.
929 58. Zhang L, Li Z, Peng Y, Liang X, Wilson K, Chipabika G, et al. Global genomic signature
930 reveals the evolution of fall armyworm in the Eastern hemisphere. Mol Ecol. 2023;00:1-
931 16. doi: 10.1111/mec.17117
932 59. Talla V, Pierce AA, Adams KL, de Man TJB, Nallu S, Villablanca FX, et al. Genomic
933 evidence for gene flow between monarchs with divergent migratory phenotypes and
934 flight performance. Mol Ecol. 2020;29(14):2567–82. doi: 10.1111/mec.15508
935 60. Matthews JH. Research in motion: Patterns of large-scale migration in dragonflies and
936 birds [dissertation]. The University of Texas at Austin; 2007.
937 61. May ML, Matthews JH. Migration in Odonata: A case study of *Anax junius*. In:
938 Dragonflies and damselflies: Model organisms for ecological and evolutionary research.
939 Oxford University Press; 2008. p. 63–78.
940 62. Troast D, Suhling F, Jinguji H, Sahlén G, Ware J. A global population genetic study of
941 *Pantala flavescens*. PLOS ONE. 2016;11(3):e0148949. doi:
942 10.1371/journal.pone.0148949
943 63. Ware J, Kohli MK, Mendoza CM, Troast D, Jinguji H, Hobson KA, et al. Evidence for
944 widespread gene flow and migration in the Globe Skimmer dragonfly *Pantala*
945 *flavescens*. Int J Odonatol. 2022;25:43–55. doi: 10.48156/1388.2022.1917166
946 64. Pliske TE. Courtship behavior of the monarch butterfly, *Danaus plexippus* L. Ann
947 Entomol Soc Am. 1975;68(1):143–51. doi: 10.1093/aesa/68.1.143
948 65. Wiklund C, Friberg M. Testing the migration syndrome: Comparative fecundity of
949 migratory and non-migratory nymphaline butterflies. Ecol Entomol. 2022;47:1061–7.
950 doi: 10.1111/een.13195
951 66. Chapuis MP, Loiseau A, Michalakis Y, Lecoq M, Franc A, Estoup A. Outbreaks, gene
952 flow and effective population size in the migratory locust, *Locusta migratoria*: A
953 regional-scale comparative survey. Mol Ecol. 2009;18(5):792–800. doi: 10.1111/j.1365-
954 294X.2008.04072.x
955 67. López-Mañas R, Pascual-Díaz JP, García-Berro A, Bahleman F, Reich MS, Pokorny L,

et al. Erratic spatiotemporal vegetation growth anomalies drive population outbreaks in a trans-Saharan insect migrant. Proc Natl Acad Sci. 2022;119(19):3–5. doi: 10.1073/pnas.2121249119

68. Neck RW. Causal analysis of a migration of the snout butterfly*, Libytheana bachmanii larvata* (Strecker) (Liibytheidae). J Lepidopterists Soc. 1983;37(2):121–8.

69. Russell RW, May ML, Soltesz KL, Fitzpatrick JW. Massive swarm migrations of dragonflies (Odonata) in eastern North America. Am Midl Nat. 1998;140(2):325–42. doi: 10.1674/0003-0031(1998)140[0325:MSMODO]2.0.CO;2

70. Menz MHM, Brown BV, Wotton KR. Quantification of migrant hoverfly movements (Diptera: Syrphidae) on the west coast of North America. R Soc Open Sci. 2019;6: 190153. doi: 10.1098/rsos.190153

71. Buffalo V. Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's Paradox. eLife. 2021;10:e67509. doi: 10.7554/eLife.67509

72. Altizer S, Hobson KA, Davis AK, De Roode JC, Wassenaar LI. Do healthy monarchs migrate farther? Tracking natal origins of parasitized vs. uninfected monarch butterflies overwintering in Mexico. Breuker CJ, editor. PLOS ONE. 2015;10(11):e0141371. doi: 10.1371/journal.pone.0141371

73. Du B, Ding D, Ma C, Guo W, Kang L. Locust density shapes energy metabolism and oxidative stress resulting in divergence of flight traits. Proc Natl Acad Sci. 2022;119(1):e2115753118. doi: 10.1073/pnas.2115753118

74. Gatehouse AG. Insect migration: Variability and success in a capricious environment. Popul Ecol. 1994;36(2):165–71. doi: 10.1007/BF02514932

75. Chapman BB, Hulthén K, Blomqvist DR, Hansson LA, Nilsson JÅ, Brodersen J, et al. To boldly go: Individual differences in boldness influence migratory tendency. Ecol Lett. 2011;14(9):871–6. doi: 10.1111/j.1461-0248.2011.01648.x

76. Zhan S, Zhang W, Niitepõld K, Hsu J, Haeger JF, Zalucki MP, et al. The genetics of monarch butterfly migration and warning colouration. Nature. 2014;514:317–21. doi:

10.1038/nature13812

77. Van Petegem KHP, Pétillon J, Renault D, Wybouw N, Van Leeuwen T, Stoks R, et al. Empirically simulated spatial sorting points at fast epigenetic changes in dispersal behaviour. Evol Ecol. 2015;29:299–310. doi: 10.1007/s10682-015-9756-9

78. Boman J, Zhu Y, Höök L, Vila R, Talavera G, Backström N. Environmental stress during larval development induces DNA methylation shifts in the migratory painted lady butterfly (*Vanessa cardui*). Mol Ecol. 2023;00:1-11. doi: 10.1111/mec.16957

79. Näsvall K, Shipilina D, Vila R, Talavera G, Backstrom N. Resource availability affects activity profiles of regulatory elements in a long-distance butterfly migrant [preprint]. 2023. doi: 10.22541/au.167827909.99815237/v1

80. Jones CM, Papanicolaou A, Mironidis GK, Vontas J, Yang Y, Lim KS, et al. Genomewide transcriptional signatures of migratory flight activity in a globally invasive insect pest. Mol Ecol. 2015;24(19):4901–11. doi: 10.1111/mec.13362

81. Stefanescu C, Ubach A, Wiklund C. Timing of mating, reproductive status and resource availability in relation to migration in the painted lady butterfly. Animal Behaviour. 2021;172:145–53. doi: 10.1016/j.anbehav.2020.12.013

82. Hobson KA, Plint T, Serrano EG, Alvarez XM, Ramirez I, Longstaffe FJ. Within-wing isotopic (δ^2H, δ^{13}C, δ^{15}N) variation of monarch butterflies: Implications for studies of migratory origins and diet. Anim Migr. 2017;4:8–14. doi: 10.1515/ami-2017-0002

83. Paritte JM, Kelly JF. Effect of cleaning regime on stable-isotope ratios of feathers in Japanese Quail (*Coturnix japonica*). The Auk. 2009;126(1):165–74. doi: 10.1525/auk.2009.07187

84. Wassenaar LI, Hobson KA. Comparative equilibration and online technique for determination of non-exchangeable hydrogen of keratins for use in animal migration studies. Isotopes Environ Health Stud. 2003;39(3):211–7. doi: 10.1080/1025601031000096781

85. Soto DX, Koehler G, Wassenaar LI, Hobson KA. Re-evaluation of the hydrogen stable isotopic composition of keratin calibration standards for wildlife and forensic science

applications. Rapid Commun Mass Spectrom. 2017;31(14):1193–203. doi: 10.1002/rcm.7893

86. Coplen TB, Qi H. 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. 2012;214:135–41. doi: 10.1016/j.forsciint.2011.07.035

87. Moore LJ, Murphy TJ, Barnes IL, Paulsen P. J. Absolute isotopic abundance ratios and atomic weight of a reference sample of strontium. J Res Natl Bur Stand. 1982;87(1):1–8.

88. Ma C, Vander Zanden HB, Wunder MB, Bowen GJ. assignR: An R package for isotope-based geographic assignment. Methods Ecol Evol. 2020;00:1–6. doi: 10.1111/2041-210x.13426

89. R Core Team. R: A language and environment for statistical computing [xtky| fwj]. Vienna, Austria: R Foundation for Statistical Computing. 2021. Available from: https://www.R-project.org/

90. Ghouri S, Reich MS, Lopez-Mañas R, Talavera G, Bowen G, Vila R, et al. A hydrogen isoscape for tracing the migration of terrestrial herbivorous insects across the Afro-Palearctic range [Preprint]. Authorea. 2023. doi: 10.22541/au.168897547.72773944/v1

91. Reich, MS, Ghouri S, Zabudsky S, Hu L, Le Corre M, Ng'iru I, Benyamini D, Shipilina S, Collins SC, Martins DJ, Vila R, Talavera G, Bataille CP. Trans-Saharan migratory patterns in *Vanessa cardui* and evidence for a southward leapfrog migration. [In review].

92. Vander Zanden HB, Nelson DM, Wunder MB, Conkling TJ, Katzner T. Application of isoscapes to determine geographic origin of terrestrial wildlife for conservation and management. Biol Conserv. 2018;228:268–80. doi: 10.1016/j.biocon.2018.10.019

93. Hallworth MT, Marra PP, McFarland KP, Zahendra S, Studds CE. Tracking dragons: Stable isotopes reveal the annual cycle of a long-distance migratory insect. Biol Lett. 2018;14(12). doi: 10.1098/rsbl.2018.0741

1040 94. Katzner TE, Nelson DM, Braham MA, Doyle JM, Fernandez NB, Duerr AE, et al.
1041 Golden Eagle fatalities and the continental-scale consequences of local wind-energy
1042 generation. Conserv Biol. 2017;31(2):406–15. doi: 10.1111/cobi.12836
1043 95. Rohlf FJ. tpsDig2 ver. 2.32. [software]. Ecology & Evolution, SUNY, Stony Brook; 2016.
1044 96. Escobar-Suárez S, Villalobos-Leiva A, Fabres A, Órdenes-Clavería R, Cruz-Jofré F,
1045 Laroze D, et al. A geometric morphometrics and genetics characterization of *Vanessa*
1046 *carye* in an extreme elevational gradient in the Chilean Altiplano. Zool Anz.
1047 2023;304:105–12. doi: 10.1016/j.jcz.2023.04.001
1048 97. Soule AJ, Decker LE, Hunter MD. Effects of diet and temperature on monarch butterfly
1049 wing morphology and flight ability. J Insect Conserv. 2020;24(6):961–75. doi:
1050 10.1007/s10841-020-00267-7
1051 98. Adams DC, Otárola-Castillo E. geomorph: An R package for the collection and analysis
1052 of geometric morphometric shape data. Methods Ecol Evol. 2013;4:393–9. doi:
1053 10.1111/2041-210X.12035
1054 99. Garcia M, Juhos S, Larsson M, Olason PI, Martin M, Eisfeldt J, et al. Sarek: A portable
1055 workflow for whole-genome sequencing analysis of germline and somatic variants.
1056 F1000Research. 2020;9(63):1-20. doi: 10.12688/f1000research.16665.2
1057 100. Andrews S. FastQC: A quality control tool for high throughput sequence data
1058 [software]. 2010.
1059 101. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and
1060 deduplication using fastp. iMeta. 2023;2:e107. doi: 10.1002/imt2.107
1061 102. Chen S, Zhou Y, Chen Y, Gu J. fastp: An ultra-fast all-in-one FASTQ preprocessor.
1062 Bioinformatics. 2018;34:i884–90. doi: 10.1093/bioinformatics/bty560
1063 103. Lohse K, Wright C, Talavera G, García-Berro A. The genome sequence of the painted
1064 lady, *Vanessa cardui* Linnaeus 1758. Wellcome Open Res. 2021;6:324. doi:
1065 10.12688/wellcomeopenres.17358.1
1066 104. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM
1067 [preprint]. arXiv. 2013. Available from: http://arxiv.org/abs/1303.3997

1068 105. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
1069 Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9. doi:
1070 10.1093/bioinformatics/btp352
1071 106. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework
1072 for variation discovery and genotyping using next-generation DNA sequencing data.
1073 Nat Genet. 2011;43(5):491–8. doi: 10.1038/ng.806
1074 107. Van Der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A,
1075 et al. From FastQ data to high-confidence variant calls: The genome analysis toolkit
1076 best practices pipeline. Curr Protoc Bioinforma. 2013;11(1110):1-43. doi:
1077 10.1002/0471250953.bi1110s43
1078 108. Li H. A statistical framework for SNP calling, mutation discovery, association mapping
1079 and population genetical parameter estimation from sequencing data. Bioinformatics.
1080 2011;27(21):2987–93. doi: 10.1093/bioinformatics/btr509
1081 109. Shipilina D, Nasvall K, Höök L, Vila R, Talavera G, Backström N. Linkage mapping and
1082 genome annotation give novel insights into gene family expansions and regional
1083 recombination rate variation in the painted lady (*Vanessa cardui*) butterfly. Genomics.
1084 2022;114(6):110481. doi: 10.1101/2022.04.14.488360
1085 110. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant
1086 call format and VCFtools. Bioinformatics. 2011;27(15):2156–8. doi:
1087 10.1093/bioinformatics/btr330
1088 111. Privé F, Luu K, Vilhjálmsson BJ, Blum MGB. Performing highly efficient genome scans
1089 for local adaptation with R package pcadapt Version 4. Rosenberg M, editor. Mol Biol
1090 Evol. 2020;37(7):2153–4. doi: 10.1093/molbev/msaa053
1091 112. Frichot E, Mathieu F, Trouillon T, Bouchard G, François O. Fast and efficient estimation
1092 of individual ancestry coefficients. Genetics. 2014;196(4):973–83.
1093 113. Raj A, Stephens M, Pritchard JK. fastSTRUCTURE: Variational inference of population
1094 structure in large SNP data sets. Genetics. 2014;197(2):573–89. doi:
1095 10.1534/genetics.113.160572

114. Korunes KL, Samuk K. pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. Mol Ecol Resour. 2021;21(4):1359–68. doi: 10.1111/1755-0998.13326

115. Samuk K, Korunes K, Trevisani MD, Moshiri N. ksamuk/pixy: pixy 1.2.7.beta1 [software]. 2022. Available from: https://zenodo.org/record/6551490