



The Afrotropical breeding grounds of the Palearctic-African migratory painted lady butterflies (*Vanessa cardui*)

Gerard Talavera^{a,b,1} , Aurora García-Berro^a , Valery N. K. Talla^c, Ivy Ng'iru^d , Farid Bahleman^e, Khadim Kébé^f , Kelvin M. Nzala^g, Dulce Plasencia^h, Mohammad A. J. Marafiiⁱ , Abeje Kassie^j , Eude O. A. Goudégnon^k, Martial Kiki^k, Dubi Benyamini^l, Megan S. Reich^m , Roger López-Mañás^a , Fulvia Benetelloⁿ, Steve C. Collins^o, Clément P. Bataille^p , Naomi E. Pierce^b , Dino J. Martins^{a,d,q} , Tomasz Suchan^r , Mattia Menchetti^s , and Roger Vila^s

Edited by David Denlinger, The Ohio State University, Columbus, OH; received November 3, 2022; accepted February 25, 2023

Migratory insects are key players in ecosystem functioning and services, but their spatiotemporal distributions are typically poorly known. Ecological niche modeling (ENM) may be used to predict species seasonal distributions, but the resulting hypotheses should eventually be validated by field data. The painted lady butterfly (*Vanessa cardui*) performs multigenerational migrations between Europe and Africa and has become a model species for insect movement ecology. While the annual migration cycle of this species is well understood for Europe and northernmost Africa, it is still unknown where most individuals spend the winter. Through ENM, we previously predicted suitable breeding grounds in the subhumid regions near the tropics between November and February. In this work, we assess the suitability of these predictions through i) extensive field surveys and ii) two-year monitoring in six countries: a large-scale monitoring scheme to study butterfly migration in Africa. We document new breeding locations, year-round phenological information, and host-plant use. Field observations were nearly always predicted with high probability by the previous ENM, and monitoring demonstrated the influence of the precipitation seasonality regime on migratory phenology. Using the updated dataset, we built a refined ENM for the Palearctic-African range of *V. cardui*. We confirm the relevance of the Afrotropical region and document the missing natural history pieces of the longest migratory cycle described in butterflies.

insect migration | species distribution modelling | population dynamics | painted lady butterfly | afrotropics

“Butterfly migration” has long fascinated scientists and naturalists across the world (1–3), but our understanding of most of their seasonal movements remains limited. Ecological niche modeling can fill gaps in partially known species distributions, including those of migratory species (4, 5). However, model predictions outside the geographic scope of the researchers’ initial dataset ultimately require validation by in situ fieldwork (6–9).

Mapping species distributions of migratory animals undergoing continuous or seasonal movements is often challenging. Considerable efforts have been made in gathering biodiversity data, and global open-access online repositories are now producing a wealth of information. Nevertheless, primary data for biogeography research are often incomplete and biased (10–13). Biases can be i) spatial: toward areas where access is easier or toward densely populated areas, ii) temporal: toward specific time periods when surveying activity is more intense, and iii) taxonomic or by life stage: toward particular species or stages that are more easily found and/or identified in nature. An additional major issue is how spatial and temporal absences can be reliably inferred from field surveys, biodiversity data repositories (e.g., museum collections, citizen science online platforms) and long-term monitoring schemes (11).

Insects account for the majority of terrestrial migrants and their movements are of prime ecological importance (14, 15). Understanding how distribution ranges shift at each generational step is important for conservation, pest management, and community ecology. But, seasonal distributions and, particularly, their connectivity are poorly known for most migratory insects (4, 16). Moreover, insects entail a wide variety of migratory strategies and scales of movement (17–20) and, thus, their shifting distributions are hardly comparable across taxa and require specific assessments. While some insects, such as migratory locusts, might perform opportunistic and erratic movements, often involving large outbreaks (21), other taxa such as a number of migratory Lepidoptera carry out annual cycles through recurrent distributional shifts that can be predicted to some extent (15, 22). Two butterfly species performing spectacular migrations have become models for the study and dissemination of insect migration: the monarch (*Danaus plexippus*) and

Significance

Insect migration is a poorly understood global phenomenon, with important consequences for ecosystems. A trans-Saharan migrant, the painted lady butterfly, has become a model system and a flagship species, together with the monarch. Despite state-of-the-art modeling, the localization of the overwintering grounds is highly debated, due to a stark lack of data from tropical Africa. We present an African long-term monitoring scheme for “butterfly migration” encompassing six countries, describe the yearly population dynamics, and show extensive winter breeding in the Afrotropics. A spatiotemporal niche model approach proves key to identifying these breeding regions. These results expand the Palearctic-African migratory annual cycle up to 15,000 kms and highlight the importance of field evidence in validating and refining model-based predictions.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Although PNAS asks authors to adhere to United Nations naming conventions for maps (<https://www.un.org/geospatial/mapsgeo/>), our policy is to publish maps as provided by the authors.

¹To whom correspondence may be addressed. Email: gerard.talavera@csic.es.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2218280120/-DCSupplemental>.

Published April 10, 2023.

the painted lady (*Vanessa cardui*). Although similar methodologies can be applied to their study, they represent remarkably different ecological strategies. The monarch butterfly, and particularly the eastern North American population, travels from breeding grounds in the United States and Canada in the summer, where four or even five generations transpire, to small overwintering areas in the highlands of central Mexico, where they overwinter in diapause until the following spring, when they migrate again toward breeding grounds in the North (5).

The painted lady butterfly undergoes year-round continuous movements without entering a diapause phase, involving up to ten generations per year (4, 23). Virtually distributed worldwide, the species performs multigenerational round-trip migratory circuits encompassing large geographical ranges (20). Painted Ladies in the Palearctic and the Afrotropics are strongly connected by regular trans-Saharan migrations, meaning that *V. cardui* rely on both temperate and tropical regions to complete their annual migratory cycle (23–25). However, delimiting the spatiotemporal distribution of *V. cardui* is complex because available occurrence data are strongly biased. First, records are biased toward some seasons and regions of the annual migratory range, while others are either undersampled or neglected. Second, occurrence data are also biased toward sampling adult stages, and the limited sampling of immature stages hampers the investigation on reproductive habitats (4).

Due to these biases, the distribution of *V. cardui* during the European winter months remains largely unknown. Since the discovery that key overwintering breeding sites of the European *V. cardui* butterflies could occur south of the Sahara (23, 24), there has been growing interest in understanding the spatiotemporal dynamics of the species within this region (4, 23–29). Based on a comprehensive dataset of observations of immatures (eggs, larvae, and pupae) along the entire migratory range (646 breeding sites from 30 countries), we previously used ecological niche modeling to predict the regions with the highest probabilities of presence of *V. cardui* immatures at monthly resolution, bringing solid hypotheses regarding the distribution of the species' breeding regions in Africa during European winters (4). The model predicted that the semiarid sub-Saharan Sahelian region and the Sudanian Savannah become highly suitable in September with suitability increasing in October. The arrival of migratory flocks and the massive presence of immatures in the Sahelian region and the Sudanian Savannah at that time were validated through dedicated field expeditions (23). After October, the predicted suitable habitat progressively disappears from the Sudanian Savannah and shifts south toward subhumid and humid regions closer to the tropical latitudes, reaching as far south as 5°N in Western Africa and the equatorial line, or even further south, in Eastern Africa in January. It is thus expected that the offspring of the European immigrants follow this habitat retreat toward the South, but this hypothetical distributional shift has not yet been documented or tested.

Another major unresolved mystery required to understand the transcontinental migration of the species is the lack of information on the larval hostplants used in the Afrotropical grounds and their phenology. Finding acceptable hostplants at the appropriate stage of their development is necessary for the butterflies to survive and continue their multigenerational migration circuit (22, 30, 31). Because *V. cardui* is highly polyphagous, it is often assumed that females will find suitable hostplants if there is any vegetation in an area (28, 32), but this may not be always the case. Hostplants are therefore infrequently recorded. As seasonality is particularly pronounced in the Afrotropics, mainly due to a shifting precipitation regime driven by the Intertropical Convergence Zone

(ITCZ), vegetation growth is often ephemeral, and some of these stepping stones may represent a limiting factor for migration.

In this work, we characterize the last unknown leg of the annual Palearctic-African migratory circuit of *V. cardui*. We evaluate the accuracy of a previously inferred ecological niche modeling (ENM) in ref. 4 through: i) in situ validations of presence/absence probabilities using field surveys in sub-Saharan Africa from December to January, including immature occurrence validation in six different countries (Cote d'Ivoire, Ghana, Benin, Cameroon, Uganda, and Kenya) and contributions documenting the diversity of larval hostplants; and ii) long-term, year-round monitoring of immature and adult presence/absence in the Canary Islands, Senegal, Benin, Cameroon, Ethiopia, and Kenya. With an updated dataset, for which immature occurrences in the Afrotropics are particularly improved, we build a refined, monthly spatiotemporal distribution model for the Afrotropical breeding grounds of *V. cardui* linked to the Palearctic. We aim to answer the following questions: i) do the Afrotropics represent vast breeding grounds from December to February (the northern hemisphere winter)? ii) Is there a distributional shift toward the tropics after the colonization of the semiarid Sahelian belt in the European fall? iii) What larval hostplants sustain the new populations generated in the Afrotropics during this part of the annual cycle? This study represents a concerted effort to fill a gap in our knowledge of the annual phenology and breeding activity of *V. cardui*, so we could obtain a more complete picture of the Palearctic-African migratory range.

Results

Field Surveys, Larval Hostplants, and Field-Based Model Validation. *V. cardui* immature occurrences were found in most of the surveyed countries in December/January, with the exception of the southeast region in Ghana. The recorded breeding sites were not uniformly distributed but followed spatial variation in latitude, longitude, and altitude (Fig. 1 and Dataset S1). Documented larval hostplants are shown in Fig. 2. Details on the field surveys' findings in each country and the larval hostplant diversity, phenology, and distribution are described in SI Appendix.

December/January surveys were conducted in countries that included both suitable and unsuitable breeding habitats for *V. cardui*, as predicted by the species distribution model (4), allowing us to explore a diversity of sites from low to high probabilities. Likewise, year-round monitoring surveys allowed us to test for temporal fluctuations of probabilities at permanent sites. The average probability of the pixels where immatures were empirically confirmed was $99.18 \pm 0.29\%$, while the average probability of the pixels where no immatures were observed was $91.03 \pm 1.90\%$. The probability of finding breeding at highly probable pixels was significantly higher than the probability of finding absences (t test, $t = 2.9874$, P value < 0.0032). Thus, the model shows a reliable level of prediction for the suitable breeding conditions of the species, although these seem to be restricted to the pixels with the highest probabilities. The surveys resulted in absences (or likely absences) at certain sites with probabilities $>90\%$. It is possible that these sites could be climatically suitable for *V. cardui*, but hostplants may be missing because of their different climatic needs, and/or stochasticity affecting any number of additional biotic and abiotic factors, including anthropogenic change. In summary, the climatic conditions captured by the model seem to be necessary but not sufficient to explain distributions.

Monitoring Results and Population Dynamics across Regions. We documented 2,755 caterpillars in the monitored breeding sites and 1,198 adults at hilltops. In line with the expectations

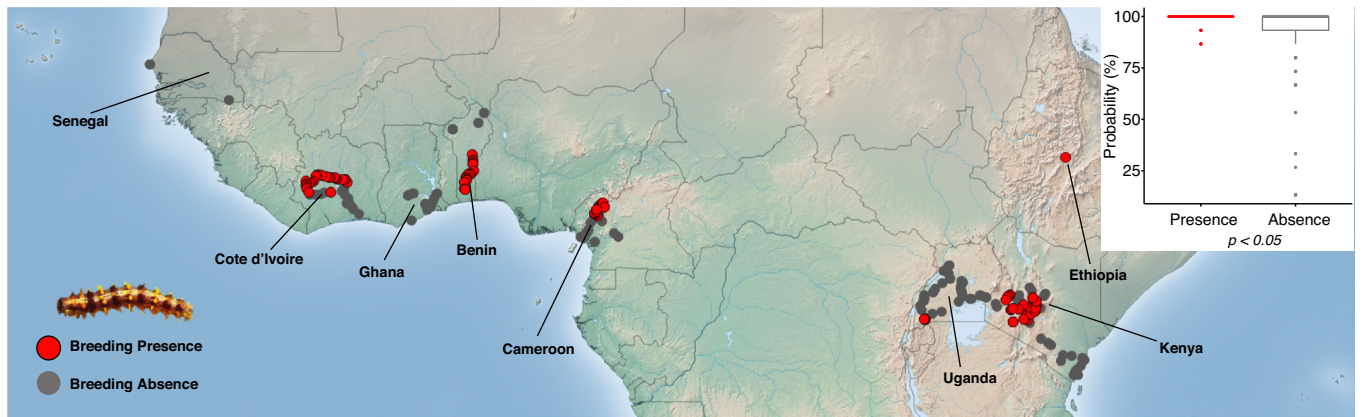


Fig. 1. (A) Localities where presence (red dots) or absence (gray dots) of *V. cardui* breeding was documented in the sub-Saharan during December (2017) and January (2018, 2019, and 2020), including Kenya, Uganda, Cameroon, Benin, Ghana, Cote d'Ivoire, and Senegal. The inset plot shows estimated *V. cardui* habitat suitability in ref. 4 spatiotemporal niche model for pixels where either breeding presence or absence was recorded. A *t* test indicated that significantly higher suitability values were found a priori at sites with presence confirmed a posteriori (mean $\pm$ SE: Presence: 991.8367 $\pm$ 29.3; absence: 910.3028 $\pm$ 190.0; *t* test, *T* = 2.9874, *P* value < 0.003186).

of the seasonality regimes in Africa driven by the ITCZ annual shift, the monitoring data showed two annual peaks of *V. cardui* abundance in highland sites in countries closer to the equator (Cameroon and Kenya), corresponding to the two rainy seasons. However, only one annual peak was observed in other sites (Canary Islands, Senegal, Benin, and Ethiopia), corresponding to the expected vegetation growth occurring right after the only annual rainy season.

The migratory phenology of *V. cardui* in the different sites was: i) Canary Islands (Tenerife): from September to May; ii) Senegal: from July to December; iii) Northern Benin: from August to December; iv) Cameroon highlands: year-round background presence, but with peaks between October and February and between March and July; v) Ethiopia (Addis Ababa): possibly year-round but with an abundance peak between October and February; vi) Kenya (Laikipia): a peak between May and August and a peak between November and January; vii) Kenya (Narok): a peak between September and March and a peak between May and August (Fig. 3). Thus, monitoring data showed a clear influence of seasonality on the migratory phenology of the different sites, especially in the savannah/grassland environments of Senegal, Benin, Canary Islands, and in Laikipia (Kenya), but less so at highland sites in Kenya, Ethiopia, or Cameroon.

As a general trend in all sites, an increase in the number of adults is observed preceding an increase in the number of larvae, and this is followed by a second increase in adult numbers, suggesting a pattern of colonization by adults, followed by a second generation (first immatures, then newly emerged adults) (Fig. 3).

A Refined Model for the Afrotropical Region. Predictive accuracy was consistently high across individual models, being the ranges of mean score values (10 runs) between 0.853 and 0.996 for True Skill Statistic (TSS) and 0.938 and 1.0 for area under the curve (AUC) (SI Appendix, Table S2). The predicted monthly distributions in the Afrotropics from September to February were generally consistent with those found by ref. 4 (Fig. 4 and SI Appendix, Fig. S2). Nevertheless, the new models decrease the areas with high probabilities (>70%) and increase the areas with low probabilities (<50%) for all months, indicating that the improved dataset leads to more specific predictions (SI Appendix, Fig. S3). Comparative subtractions between rasters show that the new models mainly discard dense rainforest areas with respect to ref. 4 and that increasing observational data from the Afrotropics greatly help the

model to restrict the predicted breeding habitat in the sub-Saharan (SI Appendix, Fig. S4). In an extreme test where no observational data from the Afrotropics was used, the predicted suitable habitat in the sub-Saharan becomes marginal (SI Appendix, Fig. S5). No important differences were observed between the models using observational data for the Palearctic + Afrotropic and those for the Afrotropical range only, although the latter produced even more restricted spatial predictions (Fig. 4 and SI Appendix, Figs. S2 and S4). Considering the whole Afrotropical region, the greatest occupancy seems to occur in November, and the lowest in September and February. The general pattern describes early occupancy of the Sahelian belt and a further progressive shift toward the South. In the Ethiopian highlands, occupancy is at its maximum in November, while in the equatorial countries of Uganda, Kenya, Rwanda, Tanzania, Burundi, and Somalia, patterns of occupancy are patchy and we observe a minimum in November. The pattern of occupancy generally matches the semiarid zone from September to November and the subhumid zone from November to February, according to the agro-ecological zones of Africa (33).

Discussion

The Afrotropical Breeding Grounds of *V. cardui*: Multiple Generations Shift South Toward the Tropics. Previous field research demonstrated that *V. cardui* colonizes the Afrotropics south of the Sahara during the European autumn (23). The monitoring data presented here for the sub-Saharan show that the colonization is recurrent across years and represents therefore a predictable pattern. While systematic monitoring was conducted for two years in a limited number of sites, the results agree with opportunistic field observations contributed over the years in supporting the view that trans-Saharan migrations are not anecdotal. Still, an important question remained: what is the fate of the offspring of these immigrants? This was still an enigma and arguably the missing spatiotemporal gap necessary to fully understand the multigenerational cycle of *V. cardui* between Africa and the Palearctic. Due to the high temperatures in the Afrotropics, we expect that the developmental time of the caterpillars is short and a single generation might require one month (23). This would mean that four or five generations could transpire in the Afrotropics from October to February.

No immature data were available until now from December to February, and debate on the potential breeding locations had developed around modeled scenarios (4, 24, 26, 28, 29, 32). Here,

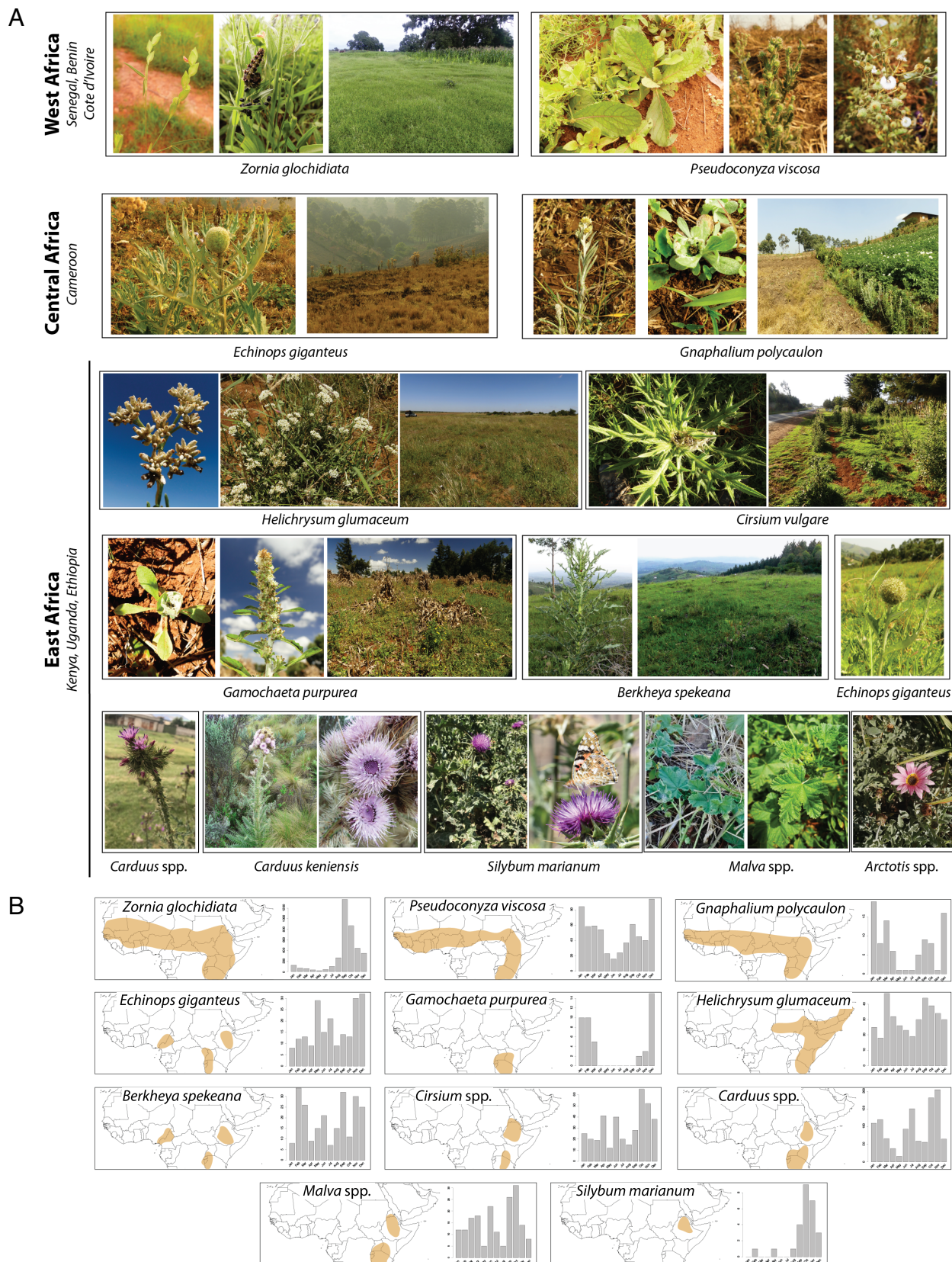


Fig. 2. (A) Hostplants used by *V. cardui* in the sub-Sahara from September to February. Other plants not documented in this work could also be used, but are likely less abundant. (B) Approximate distribution and monthly occurrences of reported hostplants according to the observations available in GBIF. Entries for January 1st, dominant within this month, were excluded as they potentially represent observations without precise dates, but likely assigned to January 1st as a default date during the databasing process.

we show that the semiarid zone in West and Central Africa (Sahelian belt and northern Savannah) is colonized in September and is then abandoned around November, until the next episode

of rains in July of the following year. There is evidence that movements of *V. cardui* occur northwards to the Maghreb and the Mediterranean basin around November, coinciding with decreased

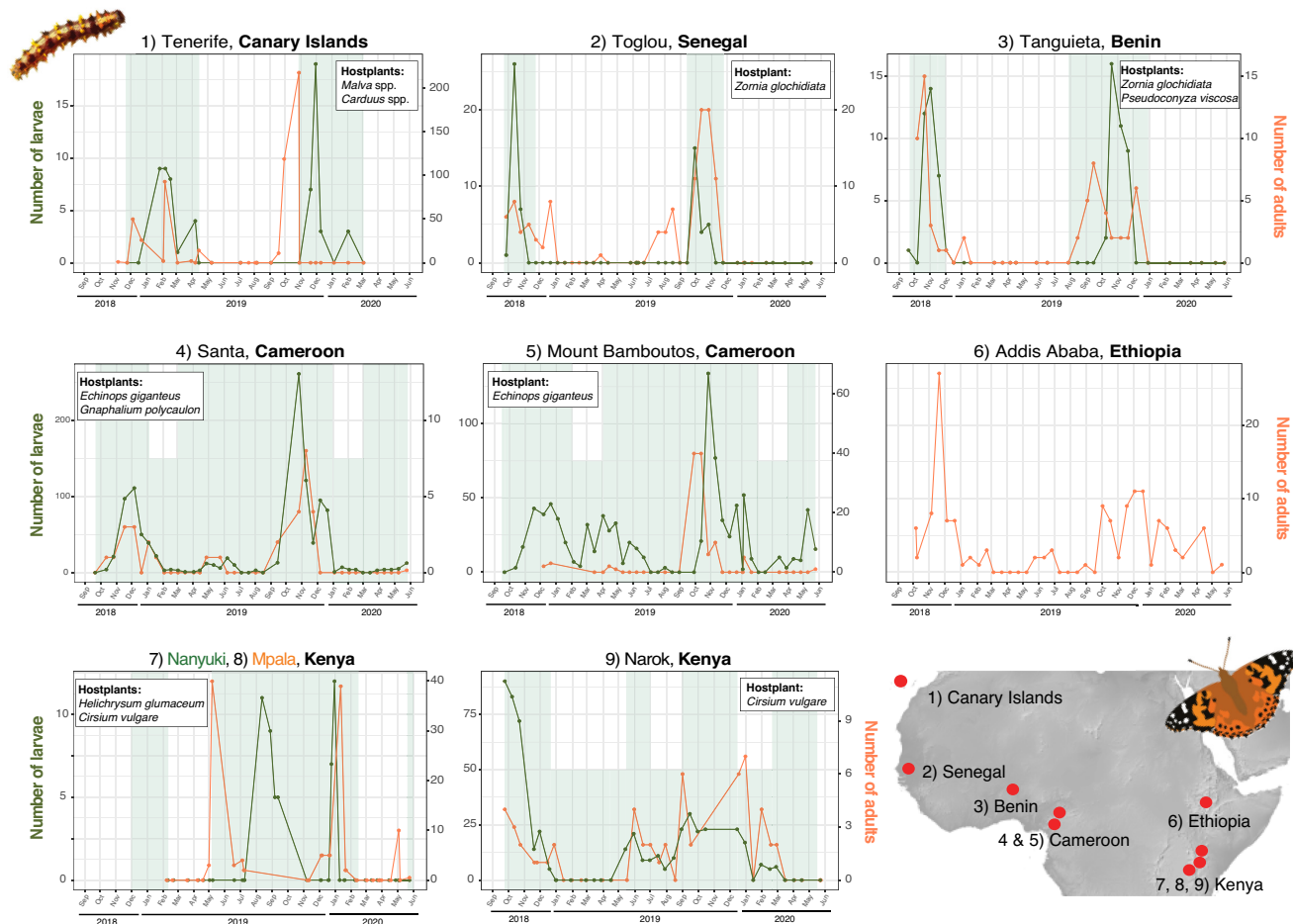


Fig. 3. Monitoring data of *V. cardui* abundances across a 2-y period (2018 to 2020). Immature and/or adult counts were obtained from nine established monitoring sites in Senegal, Benin, Cameroon, Ethiopia, Kenya, and the Canary Islands. Monitoring visits were done every 2 wk, approximately. Green bars indicate periods of active phenology of larval hostplants in the monitored sites. Full bars mean high counts of plants per plot and phenology from low growth to flowering (stages of active growth, mostly used by *V. cardui* for oviposition); half bars represent relatively low counts of plants per plot and phenology ranging from fully grown to flowered to dry; no bars indicate hostplant absence in the plot.

resources in the northernmost sub-Saharan grasslands and a synchronous increase of resources north of the Sahara (26, 34). However, the presence of *V. cardui* immatures surviving over the winter in the South-Western Palearctic or Eastern Mediterranean seems to be marginal (28, 35, 36). According to previous hypotheses (4, 23), a progressive shift south and/or to higher altitude by the Afrotropical generations is expected, and may be influenced by the progressive droughts in the grasslands from the North toward the South, a collateral effect of the retreat of the ITCZ. This seasonal southward shift has been found in other insects (37, 38). Indeed, it may resemble the movements happening in the Palearctic part of the *V. cardui* cycle in the spring and summer, where successive generations shift, by migrating relatively short distances, from the Mediterranean basin to northern Europe and up to mountain ranges (28, 36).

More heterogeneous locations in landscape, particularly the highlands in Eastern Africa and Cameroon, show higher variability in the transition patterns of presence and absence of *V. cardui*, and generally have longer seasons of high occupancy too. The monitoring data also suggest that direct migration from the Palearctic around October may not only reach the northernmost part of the Afrotropical range (i.e., the Sahelian belt or the Ethiopian highlands), but also as far south as Cameroon or Kenya. This hypothesis is supported by suitable conditions predicted by the niche model at these latitudes. The idea that some flocks migrate longer than

others, or that there are early- and late-migrants throughout the season, is not surprising considering the large number of breeding groups in the Palearctic, its large extent, and factors including the location, conditions, and timing at the natal grounds, or the particularities of each migratory journey. For example, massive colonization of the Maghreb occurs around November (34), the same time as those observed in the sub-Saharan (23, 26). The timeframe to expect European immigrants in the Afrotropics also lasts up to four months, from August to November, which, combined with a large extent of suitable grounds at that time, represents a wide range of possible spatiotemporal destinations for multiple waves of immigrants to settle down and reproduce. This pattern agrees with the idea that generations are not discrete units but rather that they may overlap in space and time to some degree, and that migratory movements in *V. cardui* should be understood as a complex connectivity network (20). The monitoring data show two annual peaks of presence in countries with a bimodal rainfall distribution pattern, with two distinct rainy seasons (i.e., Kenya, Cameroon). This is consistent with the idea that there are populations in the Afrotropics that are not linked to the Palearctic. These strict Afrotropical migrants may correspond to those peaks observed outside the September-February timeframe (4).

The set of observations of immatures gathered in the field and the modeled distributions indicate that the most likely breeding regions from December to February are concentrated in the

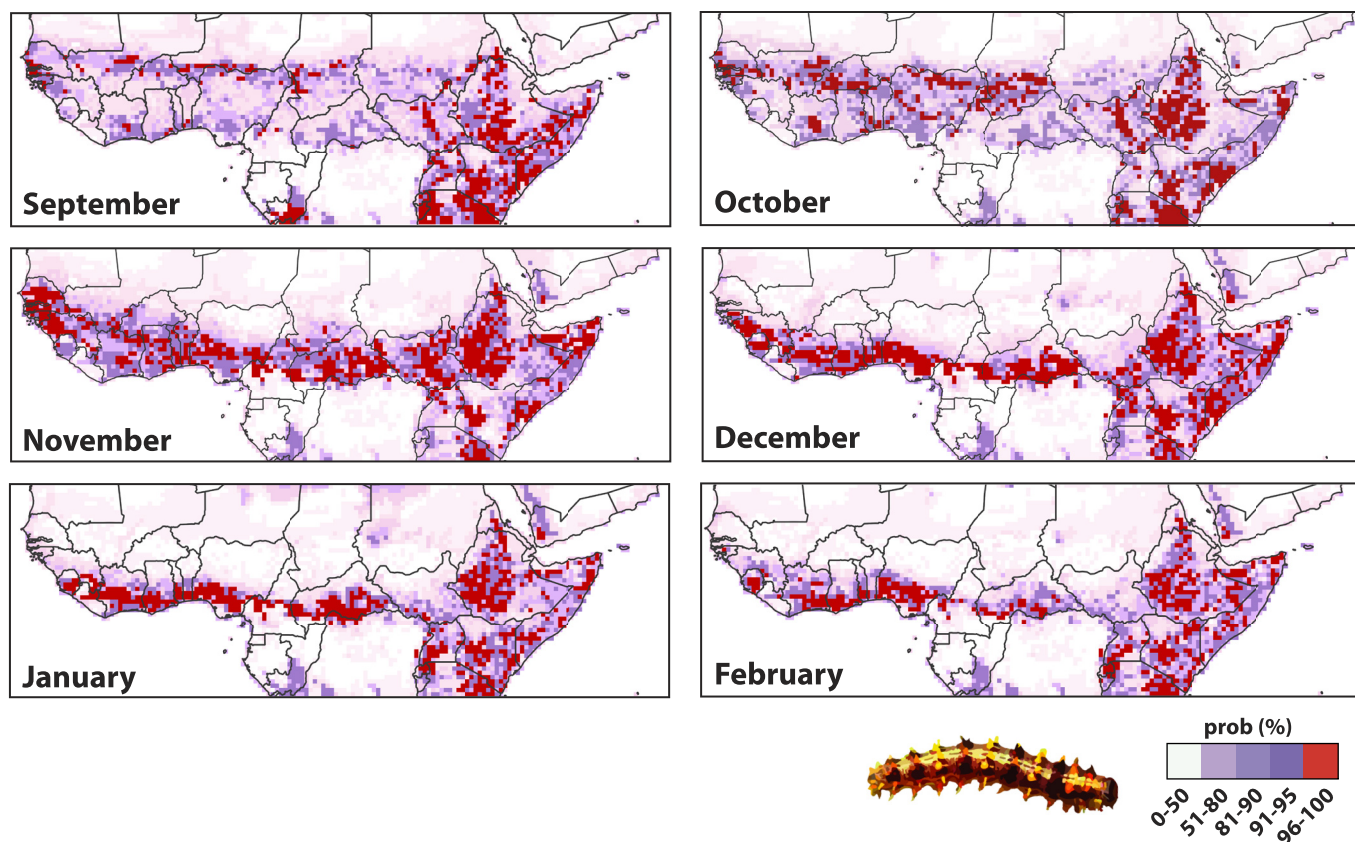


Fig. 4. Monthly projections of the refined ensemble ENM for the Afrotropical part of *V. cardui* Palearctic-Africa migratory cycle. Color denotes probabilities from 0 (white) to 100 (red).

subhumid and even humid parts of the tropics (Figs. 1, 4, and 5). These include the Guinea forest-savannah ecoregion, southern lowland Nigeria, the highlands in Cameroon/Nigeria, and the highlands in East Africa, including Ethiopia, Kenya, Tanzania, Uganda, and Rwanda (Fig. 5). Patchy areas within the semiarid zone, such as bushlands and thickets, and montane woodlands in Somalia are also retrieved by the model as being highly suitable. Areas in the subtropics, like the Canary Islands, are also suitable locations for *V. cardui* to overwinter in addition to the previously identified patches in North Africa and Arabia (4). Although no immatures were monitored in Ethiopia, a drastic decrease of observed hilltopping adults in December suggests that the region might also be partially abandoned, as the niche model reflects with a reduction of its highly suitable area (Fig. 4). While adults of *V. cardui* have departed the savannah grasslands in Western Africa by November, the grasslands monitored in Kenya are colonized around December, suggesting that *V. cardui* phenology is influenced by a latitudinal factor in addition to the type of habitat.

The models predict that the suitable range for breeding progressively decreases in the Afrotropics from December to February (Fig. 5). Although a decrease in the extent of suitable area should not necessarily reflect a decrease in population abundance, a general perception from field surveys is that abundances in the Afrotropics are lower from December to February than those from September to November. This may involve a population decrease in the Afrotropics, as it also seems to occur for breeding groups in the subtropical areas in North Africa (28, 36). Such fluctuations in census population sizes, however, do not seem to affect the species' global effective population size in the long term (20). A progressive decrease of suitable resources may trigger the last generation of Afrotropical butterflies to move northward in February/March and colonize the Maghreb, Arabia, and the Mediterranean basin.

In summary, we provide ground-truthed evidence on the distribution of the breeding grounds of *V. cardui* from December to February (Fig. 5). Suitable areas are not necessarily contiguous, as it is in the previous months, and that they are restricted to the wettest sites, mainly involving highlands and coastal areas still influenced by precipitation.

Diversity and Phenology of Larval Hostplants in the Afrotropics.

V. cardui larvae have been recorded feeding on plants in numerous families, and is thus frequently considered to be highly polyphagous. Ackery (39), for example, compiled information for 63 genera in 11 plant families. On the contrary, given the effective cosmopolitan distribution range of *V. cardui*, this hostplant diversity is not available to every butterfly generation. Generally, plant diversity is lower in arid and semiarid regions in comparison to tropical or temperate zones.

During field surveys in the Afrotropics, we reported larval hostplants belonging to 12 genera in three families (Asteraceae, Fabaceae, and Malvaceae) (Fig. 2). However, we documented only two larval hostplant species in Western Africa: *Zornia glochidiata* (Fabaceae) and *Pseudoconyza viscosa* (Asteraceae). These two plants were found to be dominant in typical habitats for *V. cardui* and were extensively exploited, often with many caterpillars per plant, suggesting that these are the preferred hostplants in the region. It is important to note that some of the most important food sources observed in other parts of the world such as plants in the genus *Malva* (Malvaceae) and thistles in the genera *Cirsium*, *Carduus*, *Silybum*, or *Onopordum* (Asteraceae) are virtually absent in West Africa. *Gnaphalium polycaulon* (Asteraceae) is used as a hostplant in Cameroon, but we did not observe this species in West Africa, despite historical records in the Global Bioinformatics Information Facility (GBIF) (Fig. 2). Although this list of hostplants is likely

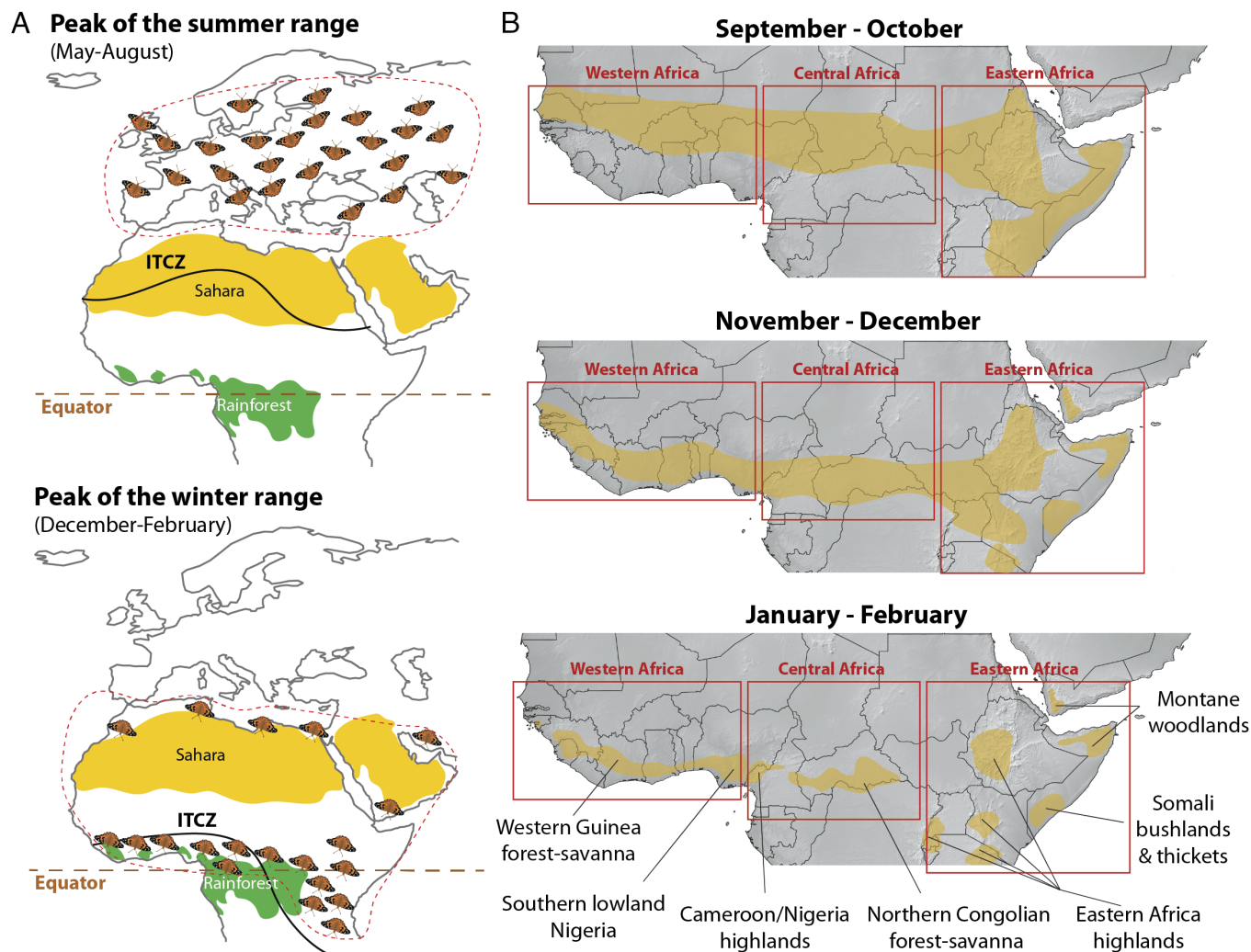


Fig. 5. Conceptual map representing the spatiotemporal occupancy transition of *Vanessa cardui*'s Palearctic-African migratory cycle (A) in the whole annual range and (B) in sub-Saharan Africa from September to February. Most individuals inhabit the Palearctic from February to September. The Intertropical Convergence Zone (ITCZ) displacement drives seasonal precipitation patterns in Africa, and hence vegetation growth. The fall and winter *V. cardui* succeeding generations in the sub-Sahara follow the retreat south of the ITCZ, which delimits the southern boundary of the cycle.

to be complemented in the future with new findings, it indicates a surprisingly low hostplant diversity. In addition, the phenological window for all herbaceous plants was narrow in the Western grasslands, likely allowing only one or two generations of *V. cardui* to breed, and thus confirming the ephemeral use of Western Africa grasslands by *V. cardui* to reproduce (23, 26).

Plant diversity and phenology is remarkably different in Eastern Africa and the Cameroonian highlands (Fig. 2). We report a higher diversity of hostplants, longer active phenologies and, in some cases, sufficient plant availability year-round. Such diversity and longer presence might be favored by the heterogeneity of the landscape and the altitudinal gradient, as well as by the seasonality regime, accounting for two rainy seasons in countries near the equator, such as Cameroon, Kenya, or Uganda. An exception is probably the Ethiopian highlands, where only one main rainy season might not be enough to avoid interseason droughts, even at a substantial altitude.

Since larval hostplants are essential for butterfly development, migratory patterns might be better understood in relation to host-plant phenological change over space. For migratory Afrotropical butterfly species such as the nymphalids *V. cardui*, *Danaus chrysippus*, and *Junonia orithya*, and the pierids *Belenois aurota* and *Catopsilia florella*, these patterns might be largely different. Even if different

species can be seen migrating in the same season, these may be seen flying in opposite directions (40, 41). While some species tend to migrate latitudinally, others seem to follow longitudinal bearings. For example, while *V. cardui* relies on the growth of fresh herbaceous plants, *B. aurota* feeds on *Boscia* evergreen trees and caperbushes (*Capparis* spp.), and the distinctive hostplant phenologies are expected to influence the migratory patterns of the butterflies. Among Afrotropical migratory butterflies, *V. cardui* is probably the only species that crosses the Sahara regularly, thus strongly relying on both tropical and temperate regions to complete their annual cycle. This demonstrates exceptional adaptability to a broad range of environments, but it may also indicate a tight link to the distribution/phenology patterns of their larval hostplants (29).

The Ecological Relevance of Delimiting Spatiotemporal Distributions in Migratory Insects. Here, we use presence records of immatures obtained from field surveys in combination with quantitative data for adults, immatures, and hostplants from a monitoring program to build a model that predicts the shifts of monthly distributions of an insect migrant. This approach underlines the importance of field data for both building and validating distribution models. This approach could be applicable to the study of any migratory insect, although obtaining a sufficient

number of observations of immatures may be difficult for some species. Both *V. cardui* (with larvae that create quite visible silken shelters) and the monarch (with aposematic larvae) are easier to detect than species with less conspicuous larvae.

Delimiting spatiotemporal distributions sheds light on the migratory ranges of the populations, which are poorly known in most migratory species. This has obvious implications in conservation management and can help identifying temporal hotspots that might be critical for the reproduction of particular generations along the annual cycle of migratory species. Reliable seasonal distributions may also allow predicting how potentially changing climatic conditions over the short term may affect these monthly distributions and the associated migratory movements. Understanding spatial and temporal gains and losses may allow determining possible factors triggering regular distributional shifts. For example, the immature phases of migratory dragonflies rely on water bodies and their seasonal desiccation may trigger migration (42, 43). In a similar way, migratory phytophagous insects rely on the availability of fresh plants, which are in some cases ephemeral (22, 29, 31). Understanding the extent of these temporal resources may give information about potential breeding stress of the insect population (if the resources are limited in space or time), or about the potential to rapidly increase population numbers in the form of outbreaks (if the temporal resources are large in space and extended in time).

Spatiotemporal distribution models, however, do not provide information on connectivity, another major challenge in insect migration ecology. Even if seasonal distributional shifts may be highlighted by distribution models, a direct link between an area lost and an area gained can only be tested with tools quantifying movement at the individual scale, such as stable isotopes (24, 44), pollen metabarcoding (25, 45), or radio tracking (46–48). At an evolutionary scale, phylogeography can also provide insights on historical connectivity by highlighting population genetic structure and thus recurrently connected/isolated populations.

Conclusion

Through this study, we help resolve a long-standing mystery of the longest migratory cycle known among butterflies: where are *V. cardui* butterflies during the European winter (from December to February)? We contribute original spatiotemporal occurrence and ecological data for this species in the Afrotropical region that is connected to the Palearctic, by long-term monitoring of phenological patterns, by documenting hostplants, and by reassessing monthly distribution areas. Overall, these data suggest a constraint in the spatial range that is suitable for *V. cardui* from December to February. Through field-based validation and systematic monitoring, we show that the use of occurrences of immatures for spatiotemporal niche modeling provides reliable predictions of suitable habitat and that this technique might be generally useful to predict seasonal distribution shifts of multigenerational insect migrants.

Methods

December-January Field Surveys and Year-Round Monitoring. Seven field expeditions were conducted across the Afrotropics during the months of December and January (2017 to 2020) to test the previously predicted breeding probability maps (4). These expeditions spanned a wide longitudinal range from West to East Africa: Cote d'Ivoire (January 2020), Ghana (January 2018), Benin (December 2019), Cameroon (January 2018), Uganda (December 2017), and Kenya (December 2017 and January 2019). Presence and absence of immatures were recorded across each site. Surveyed areas and inspected sites in each country are delineated in Fig. 1, and detailed methods are described in *SI Appendix*.

Monitoring data were systematically obtained between September 2018 and June 2020 across nine sites in six countries: Canary Islands (Spain), Senegal,

Benin, Cameroon, Ethiopia, and Kenya (*SI Appendix, Table S1* and *Dataset S2*). We collected presence and abundance data for adult and immature stages in each location. To monitor immatures, we selected sites that had abundant and contiguous larval hostplants and where we had previously documented immatures. To monitor adult activity, we selected hilltops near the breeding sites where adult males carry out territorial flights (22). Overall counts of immatures and adults were done in these sites. Visits to the sites were recurrently done approximately every fifteen days over the entire year. The overall number of surveys for the 15 monitoring sites (7 breeding sites and 8 hilltops) involved a total of 290 d of fieldwork. Detailed methods are described in *SI Appendix*.

Spatiotemporal Ecological Niche Modeling. A dataset of records of *V. cardui* immatures for the Palearctic-African migratory range was gathered, with the objective of performing spatiotemporal ENM (*SI Appendix, Fig. S1* and *Dataset S1*). The dataset combined three sources: i) the records of immatures obtained from field surveys and monitoring in this work, ii) the previously gathered dataset in ref. 4, and iii) a set of additional records from other seasons and/or regions within the range, obtained from the authors' own field surveys, from citizen scientists involved in the Butterfly Migration project (www.butterflymigration.org) and from new observations available in the platform iNaturalist. All records were compiled with associated metadata, including specific observation dates, geographical coordinates, life stages observed, and number of immatures. In total, the dataset included 1,268 breeding localities, being 637 (50.2%) new additions recorded between 2016 and 2020 and belonging to 35 countries. Among the new occurrences, 230 (36.1%) were from sub-Saharan Africa, 344 (54%) from the circum-Mediterranean region (Southern Europe, North Africa, Canary Islands, and the Middle East), and 63 (9.9%) from non-Mediterranean Europe. Many of the new occurrences (51.50%) were recorded between November and February, a period that was previously data deficient. Overall, this dataset greatly increases coverage for the winter months and for the African continent, thus representing a significant improvement for those seasons and regions, previously underrepresented.

We included absence points for modeling analyses with great caution. First, we used absences for sites where year-round monitoring was performed, fulfilling three criteria: i) no immatures were observed in that month, nor for the previous or succeeding month, ii) no hilltopping or other *V. cardui* activity was observed in the site or in nearby monitored spots, and iii) no fresh hostplant was observed at the monitoring and nearby areas. We believe that this is a reasonable validation of the absence points for the monitored sites in Africa, resulting in 14 absences in six months (*SI Appendix* and *Dataset S1*).

Additional absence points were used for the modeling, including a set of 400 random points above 50°N latitude for the winter months (corresponding to northern Europe from November to February), plus 10 field observations in desertic/arid areas, as in ref. 4. We here included an additional set of points corresponding to dense tropical rainforest in Africa, where *V. cardui* is not known to occur and no adults have been recorded in the literature. To be conservative, we selected seven pixels of fully covered canopy by visually inspecting satellite images, which were at the same time surrounded by pixels of the same type, overall spanning a total of approximately 490 km² of contiguous canopy. These pixels corresponded to locations in Gabon, Democratic Republic of Congo, Cameroon, Congo, and Equatorial Guinea. Similarly, we also selected 14 pixels in continuous sandy areas across the extent of the Sahara as permanently unsuitable, corresponding to locations in Mauritania, Mali, Algeria, Niger, Chad, Sudan, and Egypt. Pixel values for rainforest and desert included all months in randomized years.

Using the improved dataset, we built a spatiotemporal niche model to refine the inferred monthly predictions in ref. 4 (Fig. 4 and *SI Appendix, Fig. S2*), following the same approach. We used five climatic variables as time-partitioned predictors: precipitation (pre), frost day frequency (frs), daily mean temperature (tmp), vapor pressure (vap), and wet day frequency (wet). Monthly layers of these climatic variables for a time series between 1971 and 2020 were retrieved from the Climatic Research Unit time-series TS v. 4.05 at 0.5° resolution (49). We also used three land cover categories (open, closed, and bare land) that resulted from the recategorization of 22 variables obtained from Globcover v. 2.4 at 30 arc-seconds resolution (50), as in ref. 4. Land cover layers were resampled to match the resolution of the climatic variables.

Monthly *V. cardui* breeding distributions were estimated using an ensemble of forecasts of species distribution models with the R package biomod2 (51, 52).

Each spatial data point (immature occurrence) was linked to a particular month and year, and specific values of the predictor variables (bioclimatic and land cover) for each of these locations and times were retrieved. Forecasts included projections from five statistical models: generalized linear models, generalized additive models, multivariate adaptive regression splines, generalized boosted models, and random forests. We calibrated the models using a randomly selected 70% of occurrence data, and accuracy was evaluated against the remaining 30% of the data, using the area under the curve (AUC) (53) and the True Skill Statistic (TSS) (54). We replicated the analysis 10 times, and all calibrated models were after projected to a monthly set of variables resulting from averaging the time series (1971 to 2020), thus obtaining a single prediction per month and model. To summarize the results of all the models, we used an ensemble strategy with committee averaging based on model TSS values (greater than 0.35).

The modeling pipeline was run for two different observational datasets: i) for the entire dataset and projected to the whole range (Palearctic + Afrotropic) (1,268 occurrences/674 absences) (SI Appendix, Fig. S2) and ii) for the Afrotropical dataset alone (266 occurrences/199 absences) and projected to the Afrotropical range (Fig. 4). Additionally, in order to test for the importance of the Afrotropical data, we also obtained a model exclusively using the observational dataset from the Palearctic (1,002 occurrences/475 absences) and projected it to the whole range (SI Appendix, Fig. S5).

Data, Materials, and Software Availability. All study data are included in the article and/or SI Appendix.

ACKNOWLEDGMENTS. We thank Lisa Pokorny, Oriane Hidalgo, and Llorenç Sáez for help with plant identification. We thank Fatou Fall, Machihodou Baprapé, Ange Tobada, Sadam Toro, Stefano Scalercio, Beatrice Melone, Elena Plana, Perpetra Akite, Benny Shalmon, and Tzvi Avni for field support and for contributing observations. We also thank the Parc National de la Pendjari and the Parc National du W du Niger (African Parks and CENAGREF, Benin), the Mpala Research Centre, the Centre Suisse de Recherches Scientifiques en Côte d'Ivoire, the Kenya Wildlife Services, the Uganda Wildlife Authority, and the Congo Basin

Institute (Cameroon) for supporting this research. This work was funded by the National Geographic Society (grant WW1-300R-18); by the British Ecological Society (grant LRB16/1015); by the Research and Conservation Projects of the Fundació Barcelona Zoo; by the grant PID2020-117739GA-I00/MCIN/AEI/10.13039/501100011033 of the Spanish Ministry of Science and Innovation and the Spanish State Research Agency to G.T.; by the grant LINKA20399 from the Spanish National Research Council iLink program to G.T., C.P.B., N.E.P., and R.V.; by fellowship FPU19/01593 of the program Formación de Profesorado Universitario (FPU) to A.G.-B.; by the Turkana Basin Institute, National Geographic Society, and Whitley Fund for Nature to D.J.M.; and by grant 2018-00738 of the New Frontiers in Research Fund (Government of Canada) to G.T. and C.P.B.

Author affiliations: ^aInstitut Botànic de Barcelona, Consejo Superior de Investigaciones Científicas and Ajuntament de Barcelona, Barcelona, Catalonia 08038, Spain; ^bDepartment of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138; ^cLaboratory of Applied Biology and Ecology, Faculty of Science, University of Dschang, Dschang, Cameroon; ^dMpala Research Centre, Nanyuki 555-10400, Kenya; ^eSOS Savane - NGO, BP 145 Tanguiéta, Benin; ^fDepartment of Chemical Engineering and Applied Biology, Polytechnic Higher School of Dakar, BP 5085 Dakar, Senegal; ^gEgerton University, Egerton, Kenya; ^hAsociación Española para la Protección de las Mariposas y su Medio - Zerynthia, Sección Tenerife, E-26004 Logroño, Spain; ⁱDepartment of Restoration of Terrestrial and Marine Ecosystems, Public Authority of Agriculture Affairs and Fish Resources, 13075 Kuwait City, Kuwait; ^jEthiopian Biodiversity Institute, 30726 Addis Ababa, Ethiopia; ^kLaboratoire d'Ecologie Appliquée, Université d'Abomey-Calavi, BP 526 Cotonou, Benin; ^lThe Israeli Lepidopterist Society, Beit Arye 7194700, Israel; ^mDepartment of Biology, University of Ottawa, ON K1N 6N5 Ottawa, Canada; ⁿDipartimento di Biologia, Università degli Studi di Firenze, 50019 Sesto Fiorentino, Italy; ^oAfrican Butterfly Research Institute, Nairobi, Kenya; ^pDepartment of Earth and Environmental Sciences, University of Ottawa, ON K1N 6N5 Ottawa, Canada; ^qTurkana Basin Institute, Stony Brook University, NY 11794; ^rW. Szafer Institute of Botany, Polish Academy of Sciences, Kraków 31-512, Poland; and ^sInstitut de Biologia Evolutiva (Consejo Superior de Investigaciones Científicas and Universitat Pompeu Fabra), Barcelona, Catalonia 08003, Spain

Author contributions: G.T., N.E.P., D.J.M., T.S., M.M., and R.V. designed research; G.T., A.G.-B., V.N.K.T., I.N., F. Bahleman, K.K., K.M.N., D.P., M.A.J.M., A.K., E.O.A.G., M.K., D.B., M.S.R., R.L.-M., T.S., M.M., and R.V. performed research; S.C.C. and C.P.B. contributed new reagents/analytical tools; G.T., A.G.-B., R.L.-M., F. Benetello, and M.M. analyzed data; G.T., C.P.B., N.E.P., D.J.M., T.S., and R.V. acquired funding; and G.T. and R.V. wrote the paper.

1. C. B. Williams, *The Migration of Butterflies* (Oliver & Boyd, Edinburgh, 1930).
2. F. A. Urquhart, Found at last: The monarch's winter home. *Natl. Geogr. Mag.* **150**, 161–173 (1976).
3. T. C. Bonebrake, L. C. Ponision, C. L. Boggs, P. R. Ehrlich, More than just indicators: A review of tropical butterfly ecology and conservation. *Biol. Conserv.* **143**, 1831–1841 (2010).
4. M. Menchetti, M. Guéguen, G. Talavera, Spatio-temporal ecological niche modelling of multigenerational insect migrations. *Proc. Royal Soc. B.* **286**, 20191583 (2019).
5. J. M. Kass *et al.*, Biotic predictors with phenological information improve range estimates for migrating monarch butterflies in Mexico. *Ecography* **43**, 341–352 (2020).
6. T. Newbold *et al.*, Testing the accuracy of species distribution models using species records from a new field survey. *Oikos* **119**, 1326–1334 (2010).
7. B. J. Walls, A. L. Stigall, A field-based analysis of the accuracy of niche models applied to the fossil record. *Paleont. Cont.* **6**, 1–12 (2012).
8. A. Eskildsen *et al.*, Testing species distribution models across space and time: High latitude butterflies and recent warming. *Glob. Ecol. Biogeogr.* **22**, 1293–1303 (2013).
9. A. M. West, S. Kumar, C. S. Brown, T. J. Stohlgren, J. Bromberg, Field validation of an invasive species maxent model. *Ecol. Inform.* **36**, 126–134 (2016).
10. G. H. Pyke, P. R. Ehrlich, Biological collections and ecological/environmental research: A review, some observations and a look to the future. *Biol. Rev.* **85**, 247–266 (2010).
11. P. Acevedo, A. Jiménez-Valverde, P. Aragón, A. Namiir, "New developments in the study of species distribution" in *Current Trends in Wildlife Research Wildlife Research Monographs*, R. Mateo, B. Arroyo, J. Garcia, Eds. (Springer, 2016), vol. 1, pp. 151–175.
12. J. E. Ball-Damerow *et al.*, Research applications of primary biodiversity databases in the digital age. *PLoS One* **14**, e0215794 (2019).
13. C. H. Graham, S. Ferrier, F. Huettman, C. Moritz, A. T. Peterson, New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* **19**, 497–503 (2004).
14. G. Hu *et al.*, Mass seasonal bioflows of high-flying insect migrants. *Science* **354**, 1584–1587 (2016).
15. D. A. Satterfield, T. S. Sillett, J. W. Chapman, S. Altizer, P. P. Marra, Seasonal insect migrations: Massive, influential, and overlooked. *Front. Ecol. Environ.* **18**, 335–344 (2020).
16. B. Gao *et al.*, The 'migratory connectivity' concept, and its applicability to insect migrants. *Movement Ecol.* **8**, 48 (2020).
17. H. Dingle, *Migration: The Biology of Life on the Move* (Oxford University Press, 1996).
18. H. Dingle, V. A. Drake, What is migration? *Bioscience* **57**, 113–121 (1996).
19. N. Jonzén, E. Knudsen, R. D. Holt, B. E. Sæther, "Uncertainty and predictability: The niches of migrants and nomads" in *Animal Migration: A Synthesis*, E. J. Milner-Gulland, J. M. Fryxell, A. R. E. Sinclair, Eds. (Oxford University Press, 2011), pp. 91–109.
20. A. García-Berro *et al.*, Migratory behaviour is positively associated with genetic diversity in butterflies. *Mol. Ecol.* **32**, 560–574 (2023).
21. C. N. Meynard *et al.*, Climate-driven geographic distribution of the desert locust during recession periods: Subspecies' niche differentiation and relative risks under scenarios of climate change. *Glob. Chang. Biol.* **23**, 4739–4749 (2017).
22. J. W. Chapman, D. R. Reynolds, K. Wilson, Long-range seasonal migration in insects: Mechanisms, evolutionary drivers and ecological consequences. *Ecol. Lett.* **18**, 287–302 (2015).
23. G. Talavera, R. Vila, 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.* **120**, 274–285 (2017).
24. G. Talavera, C. Bataille, D. Benyamini, M. Gascoigne-Pees, R. Vila, Round-trip across the Sahara: Afrotropical Painted Lady butterflies recolonize the Mediterranean in early spring. *Biol. Lett.* **14**, 20180274 (2018).
25. T. Suchan, G. Talavera, L. Sáez, M. Ronikier, R. Vila, Pollen metabarcoding as a tool for tracking long-distance insect migrations. *Mol. Ecol. Res.* **19**, 149–162 (2019).
26. C. Stefanescu, D. X. Soto, G. Talavera, R. Vila, K. A. Hobson, Long-distance autumn migration across the Sahara by painted lady butterflies: Exploiting resource

39. P. R. Ackery, Hostplants and classification: A review of nymphalid butterflies. *Biol. J. Linn. Soc.* **33**, 95–203 (1988).
40. B. G. Milne, D. L. Hill, A migration of *Belenois aurota* F., and *Catopsilia florella* L. (Lepidoptera: Pieridae) in the Transvaal. *J. Entomol. Soc. S. Afr.* **21**, 37–42 (1964).
41. J. E. H. Flux, Migration of the butterflies *Glycesthia aurota*, *Catopsilia florella* and *Crenis occidentatum* in East Africa in 1967–1968. *J. E. Afr. Nat. Hist. Soc. Nat. Mus.* **28**, 27–29 (1970).
42. P. S. Corbet, *Dragonflies: Behaviour and Ecology of Odonata* (Harley Books, Colchester, UK, 1999).
43. C. Deacon, M. J. Samways, J. S. Pryke, Determining drivers of dragonfly diversity patterns and the implications for conservation in South Africa. *Biol. Conserv.* **245**, 108548 (2020).
44. M. S. Reich, D. T. T. Flockhart, D. R. Norris, L. Hu, C. P. Bataille, Continuous-surface geographic assignment of migratory animals using strontium isotopes: A case study with monarch butterflies. *Methods Ecol. Evol.* **12**, 2445–2457 (2021).
45. H. Jia *et al.*, Windborne migration amplifies insect-mediated pollination services. *eLife* **11**, e76230 (2022).
46. D. W. Kissling, D. E. Pattemore, M. Hagen, Challenges and prospects in the telemetry of insects. *Biol. Rev.* **89**, 511–530 (2014).
47. S. M. Knight, G. M. Pitman, D. T. T. Flockhart, D. R. Norris, Radio-tracking reveals how wind and temperature influence the pace of daytime insect migration. *Biol. Lett.* **15**, 20190327 (2019).
48. M. H. M. Menz *et al.*, 2022 Individual tracking reveals long-distance flight-path control in a nocturnally migrating moth. *Science* **377**, 764–768 (2022).
49. I. Harris, T. J. Osborn, P. Jones, D. Lister, Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* **7**, 109 (2020).
50. O. Arino *et al.*, *Global Land Cover Map for 2009 (GlobCover 2009)* (European Space Agency ESA, Pangaea, & Université Catholique de Louvain (UCL), 2012).
51. W. Thuiller, B. Lafourcade, E. Engler, M. B. Araújo, BIOMOD—A platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373 (2009).
52. W. Thuiller, D. Georges, R. Engler, biomod2: Ensemble platform for species distribution modelling. *Ecography* **32**, 369–373 (2014).
53. K. A. Swets, Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–1293 (1988).
54. O. Allouche, A. Tsoar, R. Kadmon, Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232 (2006).