

## Research



**Cite this article:** Pitteloud C *et al.* 2017

Climatic niche evolution is faster in sympatric than allopatric lineages of the butterfly genus *Pyrgus*. *Proc. R. Soc. B* **284**: 20170208.  
<http://dx.doi.org/10.1098/rspb.2017.0208>

Received: 6 February 2017

Accepted: 14 March 2017

**Subject Category:**

Evolution

**Subject Areas:**

evolution, ecology

**Keywords:**

climatic niche, macro-evolutionary processes, parametric biogeography, next-generation sequencing

**Authors for correspondence:**

Camille Pitteloud

e-mail: [camille.pitteloud@usys.ethz.ch](mailto:camille.pitteloud@usys.ethz.ch)

Nadir Alvarez

e-mail: [nadir.alvarez@unil.ch](mailto:nadir.alvarez@unil.ch)

†N.A., L.P. and S.B. supervised this work together and are considered as joint last authors.

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3726859>.

# Climatic niche evolution is faster in sympatric than allopatric lineages of the butterfly genus *Pyrgus*

Camille Pitteloud<sup>1,2,3</sup>, Nils Arrigo<sup>1</sup>, Tomasz Suchan<sup>1,4</sup>, Alicia Mastretta-Yanes<sup>5</sup>, Roger Vila<sup>6</sup>, Vlad Dincă<sup>6,7</sup>, Juan Hernández-Roldán<sup>6,8</sup>, Ernst Brockmann<sup>9</sup>, Yannick Chittaro<sup>10</sup>, Irena Kleckova<sup>11</sup>, Luca Fumagalli<sup>1</sup>, Sven Buerki<sup>12,†</sup>, Loïc Pellissier<sup>2,3,†</sup> and Nadir Alvarez<sup>1,†</sup>

<sup>1</sup>Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

<sup>2</sup>Institute of Terrestrial Ecosystems, ETH Zürich, Zurich, Switzerland

<sup>3</sup>Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

<sup>4</sup>W. Szafer Institute of Botany, Polish Academy of Sciences, Kraków, Poland

<sup>5</sup>CONACYT Research Fellow assigned to CONABIO, Comisión Nacional para el conocimiento y uso de la Biodiversidad, Mexico City, Mexico

<sup>6</sup>Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain

<sup>7</sup>Center for Biodiversity Genomics, Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada

<sup>8</sup>Departamento de Biología, Universidad Autónoma de Madrid, Madrid, Spain

<sup>9</sup>Laubacher Str. 4, Lich, Hessen, Germany

<sup>10</sup>Centre Suisse de Cartographie de la Faune, Neuchâtel, Switzerland

<sup>11</sup>Biology Centre of the Czech Academy of Sciences, Ceske Budejovice, Czech Republic

<sup>12</sup>Department of Life Sciences, Natural History Museum, Cromwell Road, London, UK

RV, 0000-0002-2447-4388; NA, 0000-0002-0729-166X

Understanding how speciation relates to ecological divergence has long fascinated biologists. It is assumed that ecological divergence is essential to sympatric speciation, as a mechanism to avoid competition and eventually lead to reproductive isolation, while divergence in allopatry is not necessarily associated with niche differentiation. The impact of the spatial context of divergence on the evolutionary rates of abiotic dimensions of the ecological niche has rarely been explored for an entire clade. Here, we compare the magnitude of climatic niche shifts between sympatric versus allopatric divergence of lineages in butterflies. By combining next-generation sequencing, parametric biogeography and ecological niche analyses applied to a genus-wide phylogeny of Palaearctic *Pyrgus* butterflies, we compare evolutionary rates along eight climatic dimensions across sister lineages that diverged in large-scale sympatry versus allopatry. In order to examine the possible effects of the spatial scale at which sympatry is defined, we considered three sets of biogeographic assignments, ranging from narrow to broad definition. Our findings suggest higher rates of niche evolution along all climatic dimensions for sister lineages that diverge in sympatry, when using a narrow delineation of biogeographic areas. This result contrasts with significantly lower rates of climatic niche evolution found in cases of allopatric speciation, despite the biogeographic regions defined here being characterized by significantly different climates. Higher rates in allopatry are retrieved when biogeographic areas are too widely defined—in such a case allopatric events may be recorded as sympatric. Our results reveal the macro-evolutionary significance of abiotic niche differentiation involved in speciation processes within biogeographic regions, and illustrate the importance of the spatial scale chosen to define areas when applying parametric biogeographic analyses.

## 1. Introduction

How speciation mechanisms shape biodiversity on Earth is a long-standing question in biology [1–6]. Because speciation depends on the response of species to abiotic conditions and biotic interactions, understanding diversification processes

relates directly to species ecology [7]. While divergent selection pressures are considered as the major drivers of lineage differentiation, the importance of shifts in ecological niche may vary according to the geographical mode of divergence.

In sympatric speciation, reproductive isolation is typically caused by differential adaptation of populations to contrasting environmental conditions [8–11], resulting in ecological speciation [12–15]. Sympatric speciation thus relies on adaptation to distinct ecological niches, which creates barriers to gene flow among populations and so allows lineage divergence [5,16]. In contrast, it is often assumed that the geographical isolation of populations in allopatry enables genetic drift to produce differentiation of sister lineages [17,18]. However, even though allopatric speciation has often been shown to be neutral, it can potentially be associated with ecological divergence [19]—driven by natural selection—when the geographical areas occupied by the diverging lineages show distinct environments [20–22].

Since they are based on different geographical contexts, sympatric speciation is also considered to differ from allopatry because it can directly arise as a consequence of competition for resources [23]. This hypothesis is supported by several models [24], as well as by cases of recently diverging lineages [25,26] and ongoing ecological speciation taking place under natural conditions [27–29]. Over small spatial scales, habitat heterogeneity in the form of climatic variation or differences in community composition may initiate niche shifting in response to competition. Variation in the performance of populations under contrasting environmental conditions may ultimately promote divergence, as shown along a depth gradient in fishes [30].

In contrast to sympatric processes, many events of speciation in allopatry are not associated with any form of ecological divergence [31], resulting in a strong signal of niche conservatism [32,33]. Even when speciation is partly driven by ecological divergence, ecological niche shifts are nonetheless expected to be less frequent than in sympatry. Because physical separation does not allow the differential action of competition, this driver of divergence is absent in allopatry. The alternative view would be that in allopatry we would expect a greater possibility of distinct opportunities available, as we would probably expect both abiotic and biotic conditions to be more distinct than in sympatry.

While ecological niche shifts have been demonstrated in many studies using single-taxon approaches, the examination of the signature of ecological speciation at larger evolutionary and geographical scales remains sparse in the literature [34–37]. This study aims at improving our understanding of the geographical context of speciation associated with ecological niche shifts, at a genus-wide evolutionary level and over large spatial scales.

Using the Palaearctic butterfly genus *Pyrgus*, we test the hypothesis that the evolutionary rate of the abiotic component of the ecological niche is higher in cases of sympatric versus allopatric speciation. We investigate variation in the evolutionary rates of eight climatic variables for sympatric and allopatric speciation events using increasingly larger delineations of geographical areas (e.g. mountain massifs) and comparing strict and broad definitions of speciation modes to test the influence of geographical areas and speciation modes. We propose an integrative approach combining next-generation sequencing of the mitochondrial genome and ribosomal DNA, with phylogenetic, biogeographic and ecological analyses. We focus on an almost exhaustive sampling of the genus of skipper

butterflies *Pyrgus* (Hübner, 1819; family: Hesperidae, subfamily: Pyrginae, tribe: Pyrgini; [38]), represented by ca 37 species in the Palaearctic.

Climatic factors associated with temperature and precipitation are critical determinants of ecological niche establishment in insects [39,40]. Thus, it is expected that such ecological factors should be associated with differences in butterfly life-history traits such as growth, development and survival at adult or pre-imaginal stages [41,42]. Whereas biotic components related with host-plant use are typically considered as key ecological factors to explain lineage divergence in phytophagous insects [43,44], this is not the case in *Pyrgus* butterflies as most of the species are generalists on *Potentilla* spp. or on various host plants [45].

Here, we targeted all divergent pairs of sister lineages (more than 90% of the Palaearctic taxa being included in our study), that evolved within contrasting geographical contexts, to test if evolutionary rates of the climatic niche are greater when divergence produces lineages with spatially overlapping ranges, than when divergence occurs between regions.

## 2. Material and methods

### (a) Sampling and molecular analyses

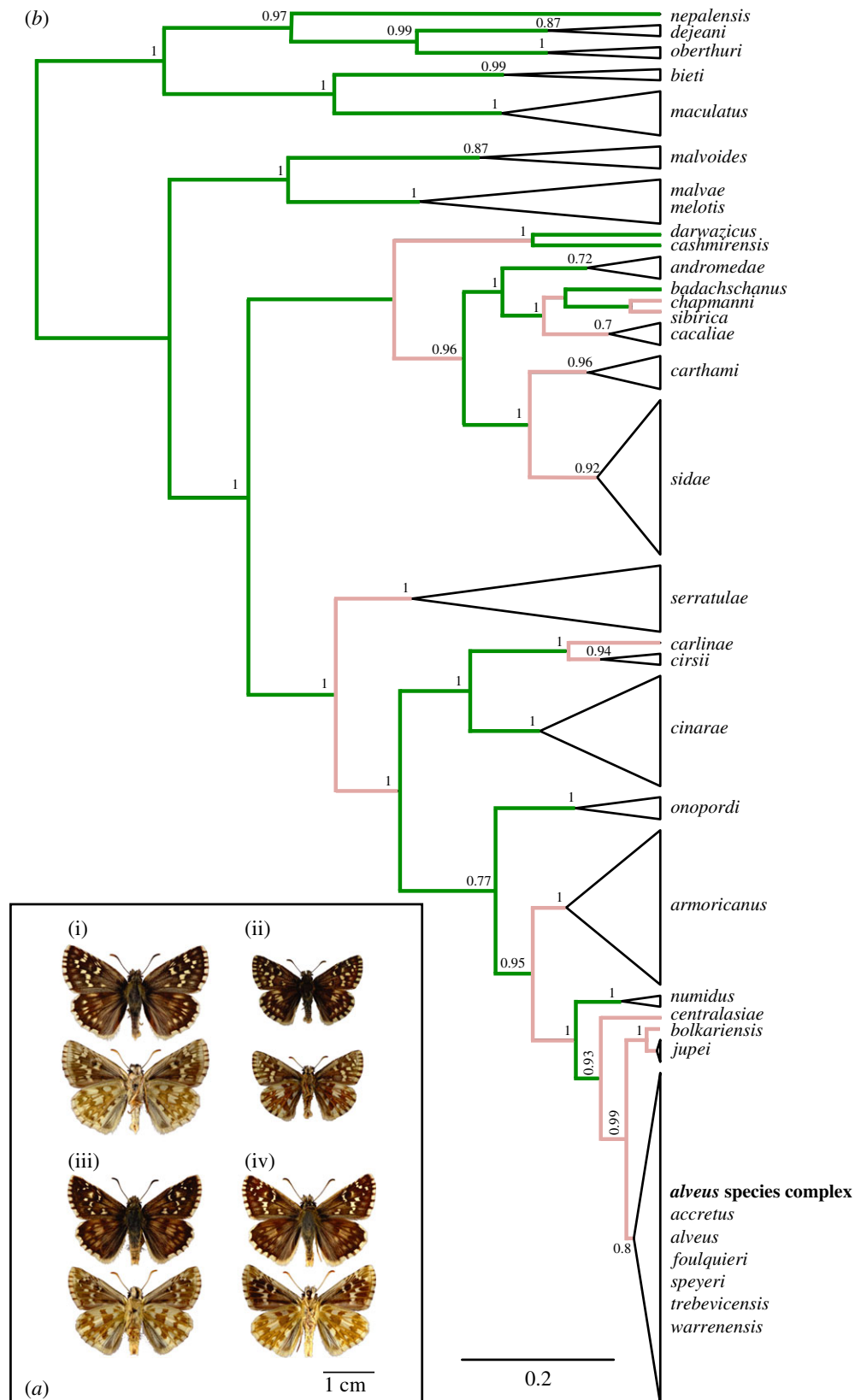
#### (i) Sampling and species determination

Sampling includes 36 of the 37 accepted *Pyrgus* species in the Palaearctic. Specimens were obtained through either field collection, museums or private donations. Preservation conditions of samples varied according to their origin (see electronic supplementary material, appendix S1, table S1). Specimens from museums or private collections were all dry-pinned and collected between 1929 and 2012. Field samples were collected from 1993 to 2013 and stored in 90% EtOH at  $-18^{\circ}\text{C}$ . The sampling was broadened with published and unpublished COI sequences from additional specimens to constitute a final molecular dataset of 132 specimens. For all the studied species, we selected several specimens from different locations across their distribution. *Spialia sertorius* (Hoffmannsegg, 1804, tribe: *Carcharodini*), *Erynnis montanus* (Bremer, 1861, tribe: *Erynnini*, GenBank accession number KC659955), *Eagris sabadius* (Gray, 1832, tribe: *Tagiadini*, GenBank accession number FJ817723), and *Celaenorhinus humbloti* (Mabille, 1884, tribe: *Celaenorhinini*, GenBank accession number FJ817832.1) were used as outgroup taxa in the phylogenetic reconstructions [46]. Species identification was performed following de Jong [47], based on morphology and upon examination of genitalia when necessary, as *Pyrgus* species are relatively undifferentiated in overall appearance (figure 1a).

Gathering of species occurrences required for biogeographic and ecological niche analyses was carried out based on exhaustive species' range distributions. We used all available precise coordinates from the samples included in this study, in combination with occurrence data from the Butterfly Tissue and DNA Collection at the Butterfly Diversity & Evolution laboratory (Institute of Evolutionary Biology, Barcelona, Spain) and from the Global Biodiversity Information Facility [48]. For taxa distributed in Asia, we used all known accessible databases and monographies (see electronic supplementary material, appendix S1, figure S1). We considered only occurrences with a precision of at least 5 km (see below), yielding a total of 16 177 records with a median value of 146 occurrence points per species.

#### (ii) DNA extraction and next-generation sequencing

DNA was extracted from butterfly legs. Museum specimens were extracted using the QIAamp DNA Micro kit (Qiagen, Hombrechtikon, Switzerland) whereas extraction of fresh samples



**Figure 1.** (a) Illustration of *Pyrgus* species. (i) *Pyrgus carthami*, (ii) *Pyrgus malvae*, (iii) *Pyrgus alveus*, (iv) *Pyrgus foulquieri*. Most *Pyrgus* species are only characterized by slight variations in wing patterns. Genitalia analysis is often required to confirm morphological determination, which nonetheless remains particularly difficult within the *alveus* species complex characterized by several taxonomic ambiguities. (b) Chronogram with tips cartooned to the species level, displaying phylogenetic relationships among *Pyrgus* taxa. Colours on branches indicate inferred sympatric (pink) versus allopatric (green) speciation events used for the climatic niche analyses in the *sensu lato* speciation mode assignment considering 15 biogeographic areas. Node supports higher than 0.5 are displayed above branches.

was automated using a high-throughput DNA-extraction robot (Biosprint 96 workstation—Qiagen, Hombrechtikon, Switzerland).

Mitochondrial sequences of nine collection samples were obtained by direct multiplex sequencing (DMPS) following a

modified protocol of Stiller *et al.* [49] and Meyer & Kircher [50]. A detailed description of the procedure is available in electronic supplementary material, appendix S2. Sequencing was performed with an Illumina MiSeq 2 × 150 bp protocol (Lausanne Genomic Technologies Facility, Switzerland).



Mitochondrial genomes and nuclear ribosomal DNA regions were analysed in 56 specimens: shotgun libraries for 24 fresh *Pyrgus* specimens were prepared using the Nextera DNA Sample-Prep-kit (Illumina, San Diego, USA) and further sequenced by Illumina HiSeq 2 × 100 bp (Fasteris, Geneva, Switzerland); 32 Asian and rare European specimens, mostly from private collections, were shotgun-sequenced using a modified version of Tin *et al.* [51]—the main modification includes a phosphorylation of the denatured extracted DNA allowing the ligation of the same P1 adaptors used for the DMPS technique (electronic supplementary material, appendix S2)—and sequenced by Illumina MiSeq 2 × 300 bp (Lausanne Genomic Technologies Facility, Switzerland). As historical shotgun libraries are more prone to sequencing errors (historical samples are usually characterized by fragmented DNA, i.e. less than 300 bp [52,53]), a large overlap between paired sequenced reads allowed for more accurate sequence assessment.

Both shotgun and DMPS raw reads were demultiplexed and cleaned with Trimmomatic [54]. DMPS libraries were assembled with MIRA 4 [55], using the complete mitochondrial genome of the closely related species *Erynnis montanus* (GenBank accession number KC659955) as a reference. For each specimen, contigs were aligned independently using CLUSTAL W [56]. Alignments were merged together with COI sequences by pairwise profile alignment and manually checked with BioEdit 7.0.4.1 [57]. For HiSeq and MiSeq shotgun libraries, *de novo* assemblies were performed using SPAdes [58]. The nuclear reference required for alignment was built using the longest SPAdes contigs that matched along ribosomal DNA using BLAST [59]; the latter included 18S ITS1, 5.8S, ITS2 and 28S regions. The *Erynnis montanus* complete mitochondrial genome served again as reference for mitochondrial alignment. Contigs were individually mapped against references using the Geneious mapper algorithm [60] with the medium sensitivity option set with two iterations. Multiple alignment was performed using MAFFT 7.017 [61] with the G-INS-I algorithm, 200PAM/k = 2 scoring matrix, gap open penalty of 1.53 and offset value of 0.123.

### (b) Genome annotation

The mitochondrial genomes as well as the nuclear ribosomal DNA regions were aligned and subsequently annotated using respectively the complete and fully annotated sequences of the mitochondrial genome (mtDNA) of *Erynnis montanus* (GenBank number KC659955) and the nuclear ribosomal DNA region of *Papilio xuthus* (GenBank number AB674749) as references. The alignment and annotation of the mitochondrial and nuclear DNA regions were done in Geneious v. 8.1.7 (Biomatters, Auckland, New Zealand) using MUSCLE [62] to conduct the alignments and the implemented Annotate and Predict functions (a similarity threshold of 80% was applied to predict and transfer the annotations). The annotations were subsequently manually checked and the reading frames adjusted if necessary. The full matrices comprising shotgun-based rDNA and mtDNA—also encompassing DMPS and COI sequences—were produced by pairwise profile alignment using ClustalX [63] and manually corrected with BioEdit 7.0.4.1 [57]. Characters of phylogenetic relevance were selected using BMGE 1.1 [64], yielding a matrix of 20 508 nucleotides—6174 bp and 14 334 bp for the ribosomal DNA and mitochondrial genome, respectively. The two aligned and fully partitioned supermatrices (mitochondrial and nuclear alignments) were concatenated and used to conduct the phylogenetic inferences.

### (c) Phylogenetic and dating inferences

The accession of *Erynnis montanus* was used as the most external outgroup taxon. Partitioned maximum-likelihood (ML) analyses were conducted using RAxML [65,66]. We used the RAxML-HPC v.8 tool implemented on the CIPRES portal (<http://www.phylo.org/>). Analyses were fully partitioned by providing files containing the set of partitions and were run using the GTRCAT model applied

to each partition, as well as 25 rate categories as suggested by Stamatakis [65]. Node support was performed using aBayes calculation in PhyML [67,68], for which the GTR + G analysis on one single partition produced a topology identical to the one produced by RAxML. Finally, because the biogeographic and ecological niche analyses require an ultrametric phylogenetic tree [69], we first performed a dated phylogeny using BEAST 1.8.3 (<http://beast.bio.ed.ac.uk/>) based on the COI alignment, by applying a classic mitochondrial DNA molecular clock of 2.3% divergence per million year [70]; second the inferred root date was used on the combined partitioned best RAxML tree using a penalized likelihood approach, as implemented in the *Rape* package [71]. The smoothing value ( $\lambda = 100$ ) was established using the cross-validation routine implemented in the *ape* package and the relaxed model of substitution rate variation among branches was applied (see *ape* function *chronos*).

### (d) Biogeographic inferences

In this study, the biogeographic scenario of Palaearctic species of *Pyrgus* will serve as a basis to estimate and compare the evolutionary rates between splits in allopatry versus sympatry along the climatic niche. To infer the biogeographic history of Palaearctic *Pyrgus* species, we defined biogeographic areas based on natural geographical boundaries and the distribution of the genus as follows.

In order to examine putative effects of the spatial scale at which sympatry is defined, we considered three biogeographic divisions of the Palaearctic, ranging from narrow to broad area definitions (electronic supplementary material, appendix S1, figure S1). We first used 11 areas that were defined as follows: (A) North Africa and the Iberian peninsula (including the Pyrenees), (B) Western Europe including the Alps and surrounding territories limited by the North Sea, the Balkans and including the Italian pre-Alps, (C) Italy, Sardinia, Corsica and Sicily, (D) Eastern Europe including the Balkans and the Carpathians, (E) Anatolia, the Caucasus, Crimea, Southwestern Russia and Mesopotamia, (F) Ural Mountains, (G) Fennoscandinavia and Baltic regions, (H) Central Asia as well as Altai and Cashmere mountain massifs, (I) Central Himalayan range, Eastern Tibet and Sichuan (China), (K) Western Manchuria (China), Mongolia and Siberia (Russia), and (J) South-East Russia and Japan. A finer partitioning resulted in 15 divisions of the Palaearctic, while the broader delineation recasts areas into seven regions. For a detailed description of area partitioning in 15 and seven areas, please refer to the electronic supplementary material, appendix S1, figure S1.

Biogeographic scenarios were inferred using the dispersal–extinction–cladogenesis (DEC) likelihood model [72,73], extended from dispersal–vicariance analysis [74]. The inference of ancestral areas along a phylogenetic tree considers two processes: (i) anagenetic (internode) evolution and (ii) cladogenetic range evolution (at nodes) [73]. Anagenetic evolution is governed by a transition matrix (Q-matrix). Q-matrices, designed independently for each biogeographic division of the Palaearctic, were adjusted to reflect area connectivity by assigning high dispersal probabilities to  $10^0$  (=1) when areas were adjacent and lower dispersal probabilities to  $10^{-3}$  (=0.001) when they were not, with a few cases where they were set to  $10^{-1.5}$  (= 0.032) for intermediate cases where areas were not strictly adjacent (electronic supplementary material, appendix S1, table S2). We did not consider the intraspecific level because coalescent processes may prevent lineage sorting and the formation of well-segregated groups. Cladogenetic range processes at each node were assessed by performing DEC analysis. Analyses were performed using the RASP software [75] providing the ancestral area and its probability of assignment at each node.

Based on the DEC ancestral areas reconstructions and the corresponding events route, we attributed speciation modes (i.e. sympatry versus allopatry) for each node leading to pairs of sister

species (i.e. at the level of the most recent common ancestors of these sister lineages). We firstly applied a broad definition of sympatry, including dispersal events during the subsequent anagenetic process and cases of more complex geographical context (e.g. divergence events implying more than one common area between sister species). Following Buerki *et al.* [76], instances of peripheral isolation were registered under allopatry. A more stringent definition of sympatry was also applied by selecting only the nodes unambiguously showing a split either in sympatry or in allopatry. Climatic niche analyses were conducted for broad and strict definitions of sympatry and for the three partitioning types of biogeographic areas.

### (e) Estimation and comparison of evolutionary rates along climatic dimensions of the ecological niche

Based on species occurrences gathered from all known records of *Pyrgus*, we extracted climatic data using environmental layers from Hijmans *et al.* [77] on <http://www.worldclim.org>, with a grid size of 5 arc-minutes resolution. We selected eight BioClim variables among the 19 available. Related to temperature and precipitation, these variables were chosen according to their ability to describe basic general features of the environment important for butterfly ecology [78,79]. Temperature was represented by the annual mean temperature (Bio1), the maximal temperature of the warmest month (Bio5), the minimal temperature of the coldest month (Bio6), the mean temperature of the warmest quarter (Bio10) and the mean temperature of the coldest quarter (Bio11). Precipitation was described by the annual precipitation (Bio12), the precipitation of the warmest quarter (Bio18) and the precipitation of the coldest quarter (Bio19).

Note that we used here a species-based rather than a tip-based approach (i.e. with splits of interest selected upstream of the intraspecific level) to comply with requirements of algorithms that compute evolutionary rates along phylogenies [80] and to increase the sampling for a more accurate description of the ecological niche. Following climatic data extraction of specimen localities, we established ecological-niche evolutionary rates using phylogenies with one single-tip per species, each of them summarized by mean values for each of the eight BioClim variables. To account for branch-length uncertainty, we randomly sampled from the dated phylogenetic tree a total of 100 single-tipped-species trees with identical topology but different relative branch lengths.

We compared evolutionary rates for climatic factors (each of the eight BioClim variables) along branches of those 100 trees using a maximum-likelihood method implemented in the brownieREML function of the Phytools R package [81]. This function uses restricted maximum likelihood to fit the Brownian rate variation model [69] in order to compare different rates of continuous trait evolution and test whether a single-rate model is favoured over a multiple rates model that allows multistate trait evolution. In the present analysis, the alternative model considers two evolutionary rates expected to differ between pairs of sister species that evolved in sympatry versus allopatry, as previously registered according to the biogeographic analyses. The multiple continuous traits of evolution correspond in this case to differential evolutionary rates of ecological niche and are reflected by the different climatic factors. Log-likelihood values of the estimated parameters were used to calculate the Akaike information criterion (AIC), which allowed selection of the best model (i.e. single versus multiple rate model). When AIC indicated a significantly better fit of the multiple versus simple model to the data ( $p$  value < 0.05), we extracted the values of the estimated evolutionary rates (sigma2) and identified rate shifts for both speciation modes (i.e. sympatry versus allopatry). Significance levels were assessed by randomly assigning allopatric versus sympatric speciation modes to the pairs of sister species and repeating the Brownie analysis, in order to generate evolutionary rates expected under a null-hypothesis. Based on the results obtained in those

randomized datasets, the ratio of evolutionary rates in sympatry over allopatry was computed and compared to the ratio of the empirical estimates using Wilcoxon sign-rank tests and applying Bonferroni correction to account for multiple testing.

To test whether there was significant climatic variation between biogeographic areas we performed a principal component analysis (PCA) on Bioclim extracted climatic data with the ade4 R package [82] and then compared the differences among the means on each of the two first PCA axes using one-way ANOVA and post-hoc Tukey test, with the biogeographic region as a grouping factor.

## 3. Results

### (a) Alignment, annotation and phylogenetic inferences

All of the protein-coding genes (13) and rRNA regions (two: 12S and 16S) of the mtDNA genome were included in the supermatrix, but we only managed to successfully annotate 12 of the 22 tRNA regions (54.5%). The aligned nuclear ribosomal DNA matrix included the following five partitions: 28S rRNA, ITS1, 5.8S rRNA, ITS2 and 18S rRNA. The combined supermatrix thus contained 32 partitions and 20 508 characters (14 334 from the mtDNA genome and 6174 from the nuclear ribosomal region; see below). We did not find any incongruence between the mtDNA and ribosomal DNA topologies (when considering aBayes supports higher than 0.7); however, the combined partitioned inference mainly reflects the information provided by the mtDNA genome, since the inference from the nuclear ribosomal DNA region shows little support (data not shown).

We obtained mtDNA and rDNA sequences for 56 of the 58 shotgun-sequenced specimens—assemblies were not satisfactory for the representatives of *P. alpinus* and *P. centaureae*. Our dataset therefore comprised samples from 34 of the 37 Palaearctic species described and consisted of 132 specimens (including five outgroups), with variable sequence lengths ranging from 121 bp to 20 508 bp. DMPS of historical and fresh samples yielded sequences from 121 bp to 7124 bp, MiSeq shotgun libraries of historical and fresh samples from 542 bp to 15 597 bp and HiSeq shotgun libraries of fresh samples from 19 397 bp to 20 508 bp. The median number of nucleotides analysed per specimen (excluding those for which only COI was available) was 11 744.

The phylogeny largely retrieves relationships described by de Jong [47]. The monophyly of all species was supported with aBayes support higher than 0.7 and 30 among the 52 nodes describing among-species relationships within *Pyrgus* (i.e. 58%) exhibited support higher than 0.95 (figure 1b). Exceptions are: (i) nine species represented by a single specimen, (ii) seven taxa belonging to the *P. alveus* complex, a group with debated species circumscriptions, and (iii) *P. melotis*, a species sharing an identical COI sequence with specimens from *P. malvae*. The position of the outgroups was confirmed with a support of 0.87 and the dating analysis estimated the origin of the *Pyrgus* genus at 4.04 million years ago (see electronic supplementary material, appendix S1, figure S2).

### (b) Biogeographic inferences

The Lagrange biogeographic scenario that takes into account phylogenetic and branch length uncertainty is given in the electronic supplementary material, appendix S1, table S3. Both allopatric and sympatric nodes used in this study were relatively well distributed across the tree depth. Attribution of speciation modes based on the broad definition of sympatry

results in 16, 18 and 16 speciation events in sympatry and 10, 8 and 10 in allopatry for the partitioning of areas into seven, 11 and 15 regions, respectively (see an illustration of speciation mode assignment in figure 1*b*). Sympatric speciation events under the strict definition were inferred at 12, seven and six nodes and allopatric speciation was invoked at four, four and four nodes for biogeographic divisions into seven, 11 and 15 regions, respectively (electronic supplementary material, appendix S1, table S3).

### (c) Estimates and comparison of ecological niche evolutionary rates

Based on the AIC criterion, analyses of the evolutionary rate of the climatic niche made with the brownieREML R function significantly favoured the multiple model that considers different rates of evolution to the single Brownian-motion model. Results of the comparison of the evolutionary rates are shown in table 1, for both definitions of speciation mode assignment and the three definitions of biogeographic areas. When evolutionary rates are inferred from area partitioning that considers 11 or 15 regions, the outcome is consistent across the majority of the climatic variables individually tested and exhibits a significant *p*-value in most analyses. When using the strict assignment of sympatry, the multiple model applied on Bio6 and Bio11 was not significantly different from the single model. Bio5 and Bio10 appeared to have higher evolutionary rates in allopatry than in sympatry. When using seven areas in biogeographic inferences, evolutionary rates of the ecological niche were always higher for the case of allopatric speciation.

Climatic differences between the biogeographic areas are illustrated in electronic supplementary material, appendix S1, figure S3. The two first axes of the PCA explained 60.6% and 30.0% of the variance, respectively. The first axis was most strongly correlated with the following variables: Bio1 ( $r = 0.896$ ), Bio5 ( $r = 0.934$ ) and Bio10 ( $r = 0.946$ ). The second axis was mostly explained by Bio6 ( $r = 0.848$ ) and Bio11 ( $r = 0.785$ ). Results are given for the three definitions of biogeographic areas for the first PCA axis. The percentages of significant *p*-values from Tukey tests for pairwise comparisons of means were 42.9%, 60.0% and 72.4% for the division into seven, 11 and 15 areas, respectively.

## 4. Discussion

The signature of shifts along abiotic ecological dimensions during lineage divergence has largely been left unexplored at the macro-evolutionary scale (but see [34–37]). Ecological niche shift has been shown to be at work in a large number of cases of ongoing speciation. Alternatively, ecological niche assortment—an extension of Losos' size assortment concept [83] to the multidimensional ecological niche—could also allow lineages to successfully colonize contrasted habitats in sympatry. Despite the fact that our experimental design does not investigate potential consequences of ecological niche assortment at the community level, it demonstrates that the footprint of ecological niche differentiation is visible at a wide phylogenetic scale.

Here, we provide evidence, based on the analysis of Palearctic *Pyrgus* butterflies, that the evolutionary rate of abiotic dimensions of the ecological niche is usually larger

when lineages diverge in regional sympatry than in allopatry, at least when defining relatively narrow geographical areas at which sympatric versus allopatric speciation processes act (see table 1); an outcome that fits expectations of the competitive displacement theory [8] or the fitness advantage of specialization to distinct conditions [84]. However, our results point to the fact that the scale employed to define geographical areas for producing biogeographic scenarios may have a strong impact on the results. When considering broad area definitions (as in the case of our seven-areas dataset), allopatric events show a higher rate of evolution of the climatic niche than sympatric events. This is actually not surprising, as broader areas would less finely translate real geographical barriers. This may be confirmed by the low percentage of significant *p*-values in the Tukey tests for multiple means comparisons of climatic components among areas (42.9%), which is the lowest among all three datasets, indicating that with a seven-area definition, areas are less dissimilar climatically than when applying a more narrow definition (see electronic supplementary material, appendix S1, figure S3*a*). As a consequence, speciation events that would be classified as allopatric under a narrower definition of areas, would in that case be recorded as sympatric. Because branches of those splits usually have lower rates of evolution, an incorrect assignment will decrease the average values of evolutionary rates for sympatric speciation events, and blur the distinction between the two modes.

Not only does the number of areas considered have an impact on the biogeographic scenario at work, but so does the way we define sympatry. Indeed, considering a strict definition of sympatry leads to a strong decrease in the number of nodes of interest, which may then produce erroneous results because of the stochasticity associated with small numbers of observations. Although sympatric speciation was still associated, generally, with a higher evolutionary rate of the climatic niche for both the 11-area and 15-area datasets, evolutionary rate estimates were higher in allopatry for one out of eight climatic variables for the 11-area dataset. We thus suggest that the definition of biogeographic areas and the number of nodes taken into account should be considered with caution to guarantee the robustness of the analyses and the subsequent interpretation of results. We therefore recommend that studies aiming at applying the same methods to other biota should (i) favour a fine geographical division of biogeographic areas consistent with dispersal characteristics of the study species and the geographical barriers of the study area and (ii) use phylogenies with a high number of nodes in order to reduce stochasticity and, eventually, to restrict node selection to those that display high phylogenetic support and high probability of ancestral areas assignment. In addition, biogeographic inference methods using more than a single input phylogenetic tree (e.g. Bayes–Lagrange, S-DIVA as implemented in RASP [75]) may account for uncertainty in topologies, branch lengths, and therefore in ancestral areas assignment.

Notwithstanding with those potential limitations, the general pattern of our result is in line with a large number of specific case-studies [25–28,85] but also evidences the substantial role played by niche shifting in sympatry at the macro-evolutionary level. At a wider taxonomical scale, it also complements the findings from one of the few works on diversification of lepidopteran lineages, in which allopatric divergence without niche shift is shown to be commonly



**Table 1.** Summary of estimates and comparisons of the evolutionary rates of climatic niche. Results of Brownie analyses are shown for each of the eight BioClim climatic variables considered each biogeographic area definition of the Palaearctic (i.e. seven, 11 and 15 regions), and both broad and strict speciation mode assignments. The *p*-values of the Wilcoxon sign-rank test that compare if observed rate estimates are significantly different from random rate estimates, are shown after multiple testing correction. Evolutionary rates of ecological niche  $\pm$  s.e. after log transformation are indicated when the *p*-value of the Wilcoxon test is significant. The mode of speciation with the higher rate is indicated with an italic evolutionary rate estimate. This analysis was performed for all nodes, irrespective of their probability of ancestral area(s); the number of nodes used is indicated.

	7 areas			11 areas			15 areas		
	evolutionary rate estimates			evolutionary rate estimates			evolutionary rate estimates		
	<i>p</i> -value	sympatry	allopatry	<i>p</i> -value	sympatry	allopatry	<i>p</i> -value	sympatry	allopatry
<i>sensu lato</i> speciation mode	$N_{\text{nodes}}$ : sympatry = 16, allopatry = 10			$N_{\text{nodes}}$ : sympatry = 18, allopatry = 8			$N_{\text{nodes}}$ : sympatry = 16, allopatry = 10		
annual mean temperature (Bio1)	<0.001	3.92 $\pm$ 0.24	6.67 $\pm$ 0.68	<0.001	6.68 $\pm$ 0.68	3.84 $\pm$ 0.23	<0.001	6.68 $\pm$ 0.68	3.68 $\pm$ 0.25
maximum temperature of warmest month (Bio5)	<0.001	3.92 $\pm$ 0.20	6.71 $\pm$ 0.68	<0.001	6.71 $\pm$ 0.68	3.83 $\pm$ 0.21	<0.001	6.71 $\pm$ 0.68	3.99 $\pm$ 0.27
minimum temperature of coldest month (Bio6)	<0.001	4.35 $\pm$ 0.29	7.02 $\pm$ 0.70	<0.001	7.02 $\pm$ 0.70	4.58 $\pm$ 0.33	<0.001	7.02 $\pm$ 0.70	4.28 $\pm$ 0.30
mean temperature of warmest quarter (Bio10)	<0.001	3.72 $\pm$ 0.23	6.57 $\pm$ 0.67	<0.001	6.57 $\pm$ 0.67	3.53 $\pm$ 0.22	<0.001	6.57 $\pm$ 0.67	3.72 $\pm$ 0.27
mean temperature of coldest quarter (Bio11)	<0.001	4.30 $\pm$ 0.28	6.90 $\pm$ 0.69	<0.001	6.90 $\pm$ 0.68	4.40 $\pm$ 0.27	<0.001	6.90 $\pm$ 0.69	4.02 $\pm$ 0.28
annual precipitation (Bio12)	<0.001	5.75 $\pm$ 0.41	8.30 $\pm$ 0.83	<0.001	8.30 $\pm$ 0.83	5.90 $\pm$ 0.44	<0.001	8.29 $\pm$ 0.83	5.79 $\pm$ 0.42
precipitation of warmest quarter (Bio18)	<0.05	5.16 $\pm$ 0.37	7.52 $\pm$ 0.75	<0.001	7.52 $\pm$ 0.74	4.84 $\pm$ 0.40	<0.001	7.52 $\pm$ 0.75	4.81 $\pm$ 0.37
precipitation of coldest quarter (Bio19)	<0.001	4.56 $\pm$ 0.34	7.00 $\pm$ 0.71	<0.001	7.00 $\pm$ 0.71	4.71 $\pm$ 0.31	<0.001	7.00 $\pm$ 0.71	4.70 $\pm$ 0.34
<i>sensu stricto</i> speciation mode	$N_{\text{nodes}}$ : sympatry = 12 versus allopatry = 4			$N_{\text{nodes}}$ : sympatry = 7 versus allopatry = 4			$N_{\text{nodes}}$ : sympatry = 6 versus allopatry = 4		
annual mean temperature (Bio1)	<0.001	3.80 $\pm$ 0.26	6.28 $\pm$ 0.64	<0.001	3.89 $\pm$ 0.19	3.63 $\pm$ 0.32	<0.001	3.89 $\pm$ 0.21	3.47 $\pm$ 0.20
maximum temperature of warmest month (Bio5)	<0.001	4.02 $\pm$ 0.28	6.31 $\pm$ 0.64	<0.001	4.08 $\pm$ 0.28	1.26 $\pm$ 0.14	<0.001	3.75 $\pm$ 0.20	1.45 $\pm$ 0.22
minimum temperature of coldest month (Bio6)	<0.001	3.82 $\pm$ 0.26	6.61 $\pm$ 0.66	0.09 <sup>NS</sup>	N/A	N/A	0.46 <sup>NS</sup>	N/A	N/A
mean temperature of warmest quarter (Bio10)	<0.01	3.87 $\pm$ 0.27	6.17 $\pm$ 0.62	<0.001	3.96 $\pm$ 0.24	2.15 $\pm$ 0.01	<0.001	3.78 $\pm$ 0.21	2.16 $\pm$ 0.02
mean temperature of coldest quarter (Bio11)	<0.01	3.79 $\pm$ 0.25	6.49 $\pm$ 0.65	<0.001	3.85 $\pm$ 0.13	4.35 $\pm$ 0.40	0.07 <sup>NS</sup>	N/A	N/A
annual precipitation (Bio12)	<0.001	5.78 $\pm$ 0.46	7.90 $\pm$ 0.79	<0.001	5.84 $\pm$ 0.41	4.63 $\pm$ 0.34	<0.001	5.92 $\pm$ 0.44	4.61 $\pm$ 0.26
precipitation of warmest quarter (Bio18)	<0.001	5.25 $\pm$ 0.40	7.10 $\pm$ 0.71	<0.001	5.47 $\pm$ 0.41	5.11 $\pm$ 0.50	<0.001	5.56 $\pm$ 0.44	4.49 $\pm$ 0.29
precipitation of coldest quarter (Bio19)	<0.001	4.55 $\pm$ 0.33	6.60 $\pm$ 0.67	<0.001	4.60 $\pm$ 0.29	2.81 $\pm$ 0.08	<0.001	4.15 $\pm$ 0.20	2.80 $\pm$ 0.02

underestimated [86]. Although one might consider that allopatric cladogenesis may require niche shifts to accommodate the contrasting climatic conditions found in the occupied areas (see electronic supplementary material, appendix S1, figure S3b and S3c), these results reveal the importance of abiotic niche differentiation in sympatric speciation.

This outcome contrasts with recent work on the geographical context of plant speciation, which has acknowledged that ecological divergence did not vary as a function of range overlap between sister lineages [23,87]. On the contrary, our hypothesis is particularly consistent with conclusions drawn by Bush & Smith [88] in a study where they evidence the role of competition in sympatry as a leading force for niche shifting. We suggest that similar mechanisms played an important role in shaping the evolutionary history of *Pyrgus* butterflies by promoting high rates of niche shifting in sympatric sister species. Our results represent compelling evidence that the footprints of micro-evolutionary processes driven by abiotic ecological components during sympatric versus allopatric speciation are reflected at the macro-evolutionary scale.

## References

- Wright S. 1932 The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In *Proceedings of the 6th International Congress on Genetics*, pp. 356–366.
- Dobzhansky T, Dobzhansky TG. 1937 *Genetics and the origin of species*. New York, NY: Columbia University Press.
- Mayr E. 1942 *Systematics and the origin of species, from the viewpoint of a zoologist*. New York, NY: Harvard University Press.
- Schluter D. 2000 *The ecology of adaptive radiation*. Oxford, UK: OUP.
- Coyne JA, Orr HA. 2004 *Speciation*. Sunderland, MA: Sinauer Associates.
- Grant PR, Rosemary Grant B. 2011 *How and why species multiply: the radiation of Darwin's finches*. Princeton, NJ: Princeton University Press.
- Schluter D. 2001 Ecology and the origin of species. *Trends Ecol. Evol.* **16**, 372–380. (doi:10.1016/S0169-5347(01)02198-X)
- Safran RJ, Nosil P. 2012 Speciation: the origin of new species. *Nat. Educ. Knowl.* **3**, 17.
- Schluter D. 2009 Evidence for ecological speciation and its alternative. *Science* **323**, 737–741. (doi:10.1126/science.1160006)
- Faria R *et al.* 2014 Advances in ecological speciation: an integrative approach. *Mol. Ecol.* **23**, 513–521. (doi:10.1111/mec.12616)
- Feder JL, Egan SP, Nosil P. 2012 The genomics of speciation-with-gene-flow. *Trends Genet.* **28**, 342–350. (doi:10.1016/j.tig.2012.03.009)
- Berlacher SH, Feder JL. 2002 Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu. Rev. Entomol.* **47**, 773–815. (doi:10.1146/annurev.ento.47.091201.145312)
- Rundle HD, Nosil P. 2005 Ecological speciation. *Ecol. Lett.* **8**, 336–352. (doi:10.1111/j.1461-0248.2004.00715.x)
- Pfennig KS, Pfennig DW. 2009 Character displacement: ecological and reproductive responses to a common evolutionary problem. *Q. Rev. Biol.* **84**, 253–276. (doi:10.1086/605079)
- Nosil P. 2012 *Ecological speciation*. Oxford, UK: Oxford University Press.
- Hartl DL, Clark AG, Clark AG. 1997 *Principles of population genetics*. Sunderland, MA: Sinauer associates.
- Ng J, Glor RE. 2011 Genetic differentiation among populations of a Hispaniolan trunk anole that exhibit geographical variation in dewlap colour: Genetic differentiation in *Anolis distichus*. *Mol. Ecol.* **20**, 4302–4317. (doi:10.1111/j.1365-294X.2011.05267.x)
- Prunier JG, Dubut V, Chikhi L, Blanchet S. 2016 Isolation-by-drift: quantifying the respective contributions of genetic drift and gene flow in shaping spatial patterns of genetic differentiation. *BioRxiv*. (doi:10.1101/031633)
- McCormack JE, Zellmer AJ, Knowles LL. 2010 Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation? Insights from tests with niche models. *Evolution* **64**, 1231–1244.
- Lande R. 1976 Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334. (doi:10.2307/2407703)
- Lande R. 1980 Genetic variation and phenotypic evolution during allopatric speciation. *Am. Nat.* **116**, 463–479. (doi:10.1086/283642)
- Mallet J, Meyer A, Nosil P, Feder JL. 2009 Space, sympatry and speciation. *J. Evol. Biol.* **22**, 2332–2341. (doi:10.1111/j.1420-9101.2009.01816.x)
- Brown WL, Wilson EO. 1956 Character displacement. *Syst. Zool.* **5**, 49–64. (doi:10.2307/2411924)
- Bolnick DI. 2007 Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **38**, 459–487. (doi:10.1146/annurev.ecolsys.38.091206.095804)
- Grant PR. 1972 Convergent and divergent character displacement. *Biol. J. Linn. Soc. Lond.* **4**, 39–68. (doi:10.1111/j.1095-8312.1972.tb00690.x)
- Iyengar VK, Castle T, Mullen SP. 2014 Sympatric sexual signal divergence among North American *Calopteryx* damselflies is correlated with increased intra- and interspecific male–male aggression. *Behav. Ecol. Sociobiol.* **68**, 275–282. (doi:10.1007/s00265-013-1642-2)
- Prosser JI *et al.* 2007 The role of ecological theory in microbial ecology. *Nat. Rev. Microbiol.* **5**, 384–392. (doi:10.1038/nrmicro1643)
- Wagner CE, McCune AR, Lovette IJ. 2012 Recent speciation between sympatric Tanganyikan cichlid colour morphs. *Mol. Ecol.* **21**, 3283–3292. (doi:10.1111/j.1365-294X.2012.05607.x)
- Hernández-Roldán JL, Dapporto L, Dincă V, Vicente JC, Hornett EA, Šíchová J, Lukhtanov VA, Talavera G, Vila R. 2016 Integrative analyses unveil speciation linked to host plant shift in *Spialia* butterflies. *Mol. Ecol.* **25**, 4267–4284. (doi:10.1111/mec.13756)
- Rennison DJ, Owens GL, Heckman N, Schluter D, Veen T. 2016 Rapid adaptive evolution of colour vision in the threespine stickleback radiation. *Proc. R. Soc. B* **283**, 20160242. (doi:10.1098/rspb.2016.0242)
- Turelli M, Barton NH, Coyne JA. 2001 Theory and speciation. *Trends Ecol. Evol.* **16**, 330–343. (doi:10.1016/S0169-5347(01)02177-2)
- Wiens JJ *et al.* 2010 Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol. Lett.* **13**, 1310–1324. (doi:10.1111/j.1461-0248.2010.01515.x)
- Peterson AT. 2011 Ecological niche conservatism: a time-structured review of evidence. *J. Biogeogr.* **38**, 817–827. (doi:10.1111/j.1365-2699.2010.02456.x)



34. Feder JL, Chilcote CA, Bush GL. 1988 Genetic differentiation between sympatric host races of the apple maggot fly *Rhagoletis pomonella*. *Nature* **336**, 61–64. (doi:10.1038/336061a0)
35. Florio AM, Ingram CM, Rakotondravony HA, Louis EE, Raxworthy CJ. 2012 Detecting cryptic speciation in the widespread and morphologically conservative carpet chameleon (*Furcifer lateralis*) of Madagascar. *J. Evol. Biol.* **25**, 1399–1414. (doi:10.1111/j.1420-9101.2012.02528.x)
36. Blair ME, Sterling EJ, Dusch M, Raxworthy CJ, Pearson RG. 2013 Ecological divergence and speciation between lemur (*Eulemur*) sister species in Madagascar. *J. Evol. Biol.* **26**, 1790–1801. (doi:10.1111/jeb.12179)
37. Merrill RM, Naisbit RE, Mallet J, Jiggins CD. 2013 Ecological and genetic factors influencing the transition between host-use strategies in sympatric *Heliconius* butterflies. *J. Evol. Biol.* **26**, 1959–1967. (doi:10.1111/jeb.12194)
38. Warren AD, Ogawa JR, Brower AVZ. 2008 Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperidae (Lepidoptera: Hesperioidea). *Cladistics* **24**, 642–676. (doi:10.1111/j.1096-0031.2008.00218.x)
39. Hodkinson ID. 2005 Terrestrial insects along elevation gradients: species and community responses to altitude. *Biol. Rev. Camb. Philos. Soc.* **80**, 489–513. (doi:10.1017/S1464793105006767)
40. Sømme L, Zachariassen KE. 1981 Adaptations to low temperature in high altitude insects from Mount Kenya. *Ecol. Entomol.* **6**, 199–204. (doi:10.1111/j.1365-2311.1981.tb00606.x)
41. Sehna F. 1991 Effects of cold on morphogenesis. In *Insects at low temperature* (eds RE Lee Jr, DL Denlinger), pp. 149–173. New York, NY: Springer.
42. Pearson RG, Dawson TP. 2003 Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* **12**, 361–371. (doi:10.1046/j.1466-822X.2003.00042.x)
43. Ehrlich PR, Raven PH. 1964 Butterflies and plants: a study in coevolution. *Evolution* **18**, 586–608. (doi:10.2307/2406212)
44. Kergoat GJ, Alvarez N, Hossaert-McKey M, Faure N, Silvain J-F. 2005 Parallels in the evolution of the two largest New and Old World seed-beetle genera (Coleoptera, Bruchidae). *Mol. Ecol.* **14**, 4003–4021. (doi:10.1111/j.1365-294X.2005.02702.x)
45. Working Group on Lepidopterists. 1999 *Les papillons et leurs biotopes*. Tome 2. Cheseaux-Noréaz: Switzerland: Pro Natura.
46. Maddison DR, Schulz K-S, Maddison WP. 2007 The tree of life web project. *Zootaxa* **1668**, 1–766.
47. Jong R. 1972 Systematics and geographic history of the genus *Pyrgus* in the Palaearctic region (Lepidoptera, Hesperidae). *Tijdschr. Entomol.* **115**, 1–121.
48. Telenius A. 2011 Biodiversity information goes public: GBIF at your service. *Nord. J. Bot.* **29**, 378–381. (doi:10.1111/j.1756-1051.2011.01167.x)
49. Stiller M, Knapp M, Stenzel U, Hofreiter M, Meyer M. 2009 Direct multiplex sequencing (DMPS)—a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. *Genome Res.* **19**, 1843–1848. (doi:10.1101/gr.095760.109)
50. Meyer M, Kircher M. 2010 Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448. (doi:10.1101/pdb.prot5448)
51. Tin MM-Y, Economo EP, Mikheyev AS. 2014 Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. *PLoS ONE* **9**, e96793. (doi:10.1371/journal.pone.0096793)
52. Mandrioli M, Borsatti F, Mola L. 2006 Factors affecting DNA preservation from museum-collected lepidopteran specimens. *Entomol. Exp. Appl.* **120**, 239–244. (doi:10.1111/j.1570-7458.2006.00451.x)
53. Lee PLM, Prys-Jones RP. 2008 Extracting DNA from museum bird eggs, and whole genome amplification of archive DNA. *Mol. Ecol. Resour.* **8**, 551–560. (doi:10.1111/j.1471-8286.2007.02042.x)
54. Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. (doi:10.1093/bioinformatics/btu170)
55. Chevreux B, Wetter T, Suhai S. 1999 Genome sequence assembly using trace signals and additional sequence information. Computer science and biology: proceedings of the German Conference on Bioinformatics, **99**, 45–56.
56. Thompson JD, Higgins DG, Gibson TJ. 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680. (doi:10.1093/nar/22.22.4673)
57. Hall TA. 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acid S.*, **41**, 95–98.
58. Bankevich A *et al.* 2012 SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477. (doi:10.1089/cmb.2012.0021)
59. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. (doi:10.1016/S0022-2836(05)80360-2)
60. Kearse M *et al.* 2012 Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649. (doi:10.1093/bioinformatics/bts199)
61. Katoh K, Misawa K, Kuma K-I, Miyata T. 2002 MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066. (doi:10.1093/nar/gkf436)
62. Edgar RC. 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797. (doi:10.1093/nar/gkh340)
63. Larkin MA *et al.* 2007 Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948. (doi:10.1093/bioinformatics/btm404)
64. Criscuolo A, Gribaldo S. 2010 BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210. (doi:10.1186/1471-2148-10-210)
65. Stamatakis A. 2006 RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi:10.1093/bioinformatics/btl446)
66. Stamatakis A, Hoover P, Rougemont J. 2008 A rapid bootstrap algorithm for the RAXML Web servers. *Syst. Biol.* **57**, 758–771. (doi:10.1080/10635150802429642)
67. Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O. 2011 Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst. Biol.* **60**, 685–699. (doi:10.1093/sysbio/syr041)
68. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321. (doi:10.1093/sysbio/syq010)
69. O'Meara BC, Ané C, Sanderson MJ, Wainwright PC. 2006 Testing for different rates of continuous trait evolution using likelihood. *Evolution* **60**, 922–933. (doi:10.1111/j.0014-3820.2006.tb01171.x)
70. Brower AV. 1994 Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl Acad. Sci. USA* **91**, 6491–6495. (doi:10.1073/pnas.91.14.6491)
71. Paradis E, Claude J, Strimmer K. 2004 APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290. (doi:10.1093/bioinformatics/btg412)
72. Ree RH, Smith SA. 2008 Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* **57**, 4–14. (doi:10.1080/10635150701883881)
73. Buerki S, Forest F, Alvarez N, Nylander JAA, Arrigo N, Sanmartín I. 2011 An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *J. Biogeogr.* **38**, 531–550. (doi:10.1111/j.1365-2699.2010.02432.x)
74. Ronquist F, Cannatella D. 1997 Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* **46**, 195–203. (doi:10.1093/sysbio/46.1.195)
75. Yu Y, Harris AJ, Blair C, He X. 2015 RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Mol. Phylogenet. Evol.* **87**, 46–49. (doi:10.1016/j.ympev.2015.03.008)
76. Buerki S, Forest F, Stadler T, Alvarez N. 2013 The abrupt climate change at the Eocene-Oligocene boundary and the emergence of Southeast Asia triggered the spread of sapindaceous lineages. *Ann. Bot.* **112**, 151–160. (doi:10.1093/aob/mct106)
77. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005 Very high resolution interpolated climate

- surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978. (doi:10.1002/joc.1276)
78. Mercader RJ, Scriber JM. 2008 Asymmetrical thermal constraints on the parapatric species boundaries of two widespread generalist butterflies. *Ecol. Entomol.* **33**, 537–545. (doi:10.1111/j.1365-2311.2008.01001.x)
79. Oksanen J *et al.* 2013 vegan: Community ecology package. R package version 2.0-9. See <https://cran.r-project.org>, <https://github.com/vegandevs/vegan>.
80. Eastman JM, Alfaro ME, Joyce P, Hipp AL, Harmon LJ. 2011 A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution* **65**, 3578–3589. (doi:10.1111/j.1558-5646.2011.01401.x)
81. Revell LJ. 2012 phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223. (doi:10.1111/j.2041-210X.2011.00169.x)
82. Dray S *et al.* 2007 The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* **22**, 1–20. (doi:10.18637/jss.v022.i04)
83. Losos JB. 1990 The evolution of form and function: morphology and locomotor performance in West Indian *Anolis* lizards. *Evolution* **44**, 1189–1203. (doi:10.2307/2409282)
84. Futuyma DJ, Moreno G. 1988 The evolution of ecological specialization. *Annu. Rev. Ecol. Syst.* **19**, 207–233. (doi:10.1146/annurev.es.19.110188.001231)
85. Borer M, van Noort T, Arrigo N, Buerki S, Alvarez N. 2011 Does a shift in host plants trigger speciation in the Alpine leaf beetle *Oreina speciosissima* (Coleoptera, Chrysomelidae)? *BMC Evol. Biol.* **11**, 310. (doi:10.1186/1471-2148-11-310)
86. Imada Y, Kawakita A, Kato M. 2011 Allopatric distribution and diversification without niche shift in a bryophyte-feeding basal moth lineage (Lepidoptera: Micropterigidae). *Proc. R. Soc. B* **278**, 3026–3033. (doi:10.1098/rspb.2011.0134)
87. Dieckmann U, Doebeli M. 1999 On the origin of species by sympatric speciation. *Nature* **400**, 354–357. (doi:10.1038/22521)
88. Bush GL, Smith JJ. 1998 The genetics and ecology of sympatric speciation: a case study. *Res. Popul. Ecol.* **40**, 175–187. (doi:10.1007/BF02763403)
89. Pitteloud C *et al.* 2017 Data from: Climatic niche evolution is faster in sympatric than allopatric lineages of the butterfly genus *Pyrgus*. Dryad Digital Repository. (<http://dx.doi.org/10.5061/dryad.sc613>)