**Anna K. Hundsdoerfer, Senckenberg Natural History Collections Dresden, Königsbrücker Landstr. 159, D-01109 Dresden, Germany, Tel. +49-351-7958414437, Fax. +49-351-7958414327, anna.hundsdoerfer@senckenberg.de**

# Museum archives revisited: Central Asiatic hawkmoths reveal exceptionally high late Pliocene species diversification (Lepidoptera, Sphingidae)

**ANNA K. HUNDSDOERFER [a§], MARTIN PÄCKERT [a,b], CHRISTIAN KEHLMAIER [a], PATRICK STRUTZENBERGER [a], IAN J. KITCHING [c]**

[a] Senckenberg Natural History Collections Dresden, Königsbrücker Landstr. 159, D-01109 Dresden, Germany.

[b] Biodiversity and Climate Research Centre (BiK-F), Senckenberganlage 25, D-60325 Frankfurt am Main, Germany.

[c] Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, U.K.

[§]Corresponding author.

# Abstract

Hundsdoerfer, A. K. (2016) Zoological Scripta, 00, 000-000.

Three high elevation *Hyles* species of Central Asia have proven difficult to sample and thus only a limited number of specimens are available for study. Ancient DNA techniques were applied to sequence two mitochondrial genes from 'historic' museum specimens of *H. gallii*, *H. renneri* and *H. salangensis* to elucidate the phylogenetic relationships of these species. This approach enabled us to include the holotypes and/or allotypes and paratypes. The status of *H. salangensis* as a species endemic to a mountain range north of Kabul in Afghanistan is confirmed by this study. It is most closely related to *H. nicaea* and *H. gallii*, and quite distant from the clade comprising the species from *H. vespertilio* to *H. tithymali*, despite this group and *H. salangensis* both completely lacking an arolium on the pretarsus. Our results show that the samples assigned to *H. renneri* and *H. livornica tatsienluica* are conspecific and so we reinstate *Hyles tatsienluica* **stat. nov.** as the valid name for this species and synonymize *Hyles renneri* **syn. nov.** with it. This study shows that the distribution range postulated for *H. tatsienluica* extends from Nepal well into the mountains of South Western China. The distribution ranges of *H. livornica* and *H. tatsienluica* overlap and thus assignment of specimens to these two species cannot be made based on locality data alone. The study confirms the previously proposed synonymies of *H. nepalensis*, *H. g. intermedia* and *H. g. tibetanica* with *H. gallii*. Extensive species sampling (over 80% of *Hyles* species) in this study allowed additional analyses. The dated phylogeny reveals the global *Hyles* hawkmoth radiation to be much more recent than previously thought: it began in the Late Miocene and culminated in a Pleistocene burst of diversification in the Northern Hemisphere. Ancestral ranges of basal nodes were reconstructed as highly equivocal, but the Neotropics has the highest probability

in the two oldest nodes. Although the origins of the Madagascan and Australian species also remain ambiguous, a large crown clade of fifteen species was reconstructed to have originated in the Palaearctic. The wide distribution ranges of the two migratory species, *H. livornica* and *H. gallii*, appear to blur any traces of the biogeographic origin of the clades containing these species. Specialization in larval host plant use onto particular plant families from the ancestral condition "polyphagous" may have led to an increased rate of speciation and phylogenetic diversification in three subgroups of *Hyles* (the Hawaiian clade, the *H. centralasiae*-group and the *Hyles euphorbiae*-complex).

*Anna K. Hundsdoerfer, Senckenberg Natural History Collections Dresden, Königsbrücker Landstr. 159, D-01109 Dresden, Germany*

*E-mail anna.hundsdoerfer@senckenberg.de*

**Keywords**

*Hyles*, type material, phylogeny, dating, ancestral range, biogeography.

# Introduction

The hawkmoth genus *Hyles* Hübner, 1819 (Lepidoptera: Sphingidae) has a global distribution with representatives on all continents (except Antarctica) and many major islands (e.g., Madagascar). The 32 species (following the 'Sphingidae Taxonomic Inventory'; Kitching 2016) are generally identified with the help of colour images using characters from the pattern and colour of the wings and abdomen. However, the taxonomy based on these morphological features is highly contentious with disagreement over both the number of species and how they are related (e.g., Danner et al. 1998; Harbich 2009; Hundsdoerfer et al. 2005, 2009; Meerman & Smid 1988; Speidel & Hassler 1989). A number of studies have been devoted to elucidating the phylogeny of the genus using molecular methods (Hundsdoerfer et al. 2005, 2009). For this study, we augmented the taxon sampling to over 80% (28) of the 32 species listed in Kitching (2016) and have now included all species we consider necessary for an adequate representation of their diverse biogeographic origins, enabling a Bayesian (and likelihood) evolutionary analysis of the ancestral biogeography. Of the species not included, based on morphology alone, *Hyles wilsoni* (Rothschild, 1894) from Hawaii would simply add another branch to the Hawaiian radiation. The other four, *H. apocyni* (Shchetkin, 1956), *H. chamyla* (Denso, 1913), *H. exilis* Derzhavets, 1979 (= *H. chuvilini* Eitschberger, Danner & Surholt, 1998) and *H. nervosa* (Rothschild & Jordan, 1903) would add another four Palaearctic species and further emphasise the importance of the Central Palaearctic for the biogeography of this genus.

### *Biogeography*

The current best phylogenetic hypothesis for the genus is based on data from one nuclear and three mitochondrial genes (Hundsdoerfer et al. 2009). From this, a biogeographic hypothesis was proposed that the genus originated in the Neotropics, given the oldest splits occur between lineages in this part of the world. Australia and Madagascar were postulated to have been colonised from there, the endemic species of these two regions, *H. livornicoides* (Lucas, 1780) and *H. biguttata* (Walker, 1856) respectively, branching off next. However, the routes are uncertain due to the lack of reliable dates for these phylogenetic splits to correlate with positions of land masses and the formation of the Antarctic ice shield. Hawkmoths are rather large and strong flying and should have been able to cross stretches of ocean, especially if assisted by favourable winds. They may also have used oceanic islands as stepping-stones, or at least as resting places, even if these are not inhabited by *Hyles* today. The subsequent colonisation of Hawaii could also have taken place in this manner, for example, via the Fijian islands from the south or via the Bering Strait from the north. Following this hypothesis, the Palaearctic would have been colonised from the East (Hundsdoerfer et al. 2005, 2009).

### *Systematics and taxonomy*

Danner et al. (1998) provided the first explicit subdivision of *Hyles*, recognizing eight subgenera on the basis of perceived differences in the colour patterns of the wings and larvae. Of particular relevance to the present study are the three subgenera *Celerio* Agassiz, 1846, *Danneria* Eitschberger & Zolotuhin, 1998, and the nominotypical subgenus *Hyles*. The Neotropical species, *Hyles lineata* (Fabricius, 1775), the Palaearctic/Afrotropical *H. livornica* (Esper, 1892) and the Australian *H.*

*livornicoides* all have similar wing patterns that include conspicuous white highlighting of the veins, quite different from the other members of the genus, and so were placed in subgenus *Danneria*, together with a newly described species, *H. renneri* Eitschberger, Danner & Surholt, 1998. Similarity in wing pattern, together with larval monophagy on *Euphorbia*, also led Danner et al. (1998) to place *H. euphorbiae* (Linnaeus, 1758), *H. robertsi* (Butler, 1880), *H. nicaea* (von Prunner, 1798) and *H. stroehlei* Eitschberger, Danner & Surholt, 1998) together in the nominotypical subgenus *Hyles*. They also included another taxon of the Palaearctic *Hyles-euphorbiae* complex (HEC) that they raised to species status, *H.* (*euphorbiae*) *conspicua* (Rothschild & Jordan, 1903), and also, tentatively, *H. salangensis* (Ebert, 1969) (see below). Danner et al. (1998) also resurrected the subgenus *Celerio* for *H. gallii* (von Rottemburg, 1775), *H. nepalensis* (Daniel, 1961) (a subspecies of *H. gallii* that these authors raised to species status) and, again tentatively, *H. zygophyllii* (Ochsenheimer, 1808). Their taxonomic actions were challenged by Kitching & Cadiou (2000), who morphologically revised the taxonomic status of several of the taxa that Danner et al. (1998) had raised to species status. Subsequently, phylogenetic reconstructions based on molecular data (Hundsdoerfer et al. 2005, 2009) provided a more robust hypothesis of the phylogenetic relationships within the genus. These studies showed that neither wing pattern similarity nor commonality of larval host plants necessarily reflect phylogenetic relationships within the genus *Hyles*, and consequently refuted the taxonomic validity of Danner et al.'s (1998) subgeneric concepts of *Danneria*, *Hyles* and *Celerio*. The currently valid taxonomy is maintained and updated in the 'Sphingidae Taxonomic Inventory' (Kitching 2016).

***Phylogenetic brain-teasers - the challenge of this study***

For a variety of reasons (e.g., inaccessibility of habitat, presence of human conflict zones), it has been difficult (and will presumably remain so) to obtain fresh tissue samples of several contentious Central Palaearctic *Hyles* species for molecular analyses. To circumvent this difficulty, we analysed tissue samples obtained from 'historic' dry museum specimens to try to resolve the validity and phylogenetic relationships of three enigmatic high elevation *Hyles* species: *H. salangensis, H. renneri* and *H. nepalensis*.

The first of these enigmatic species, *H. salangensis,* appears to be known reliably only from the original type material, 15 males collected in the 1960s from the Salang Pass, north of Kabul, Afghanistan (Ebert 1969) and a single male in NHMUK (NHMUK specimen BMNH #812472; depository abbreviations are given in legend of Table 1) captured in 1975. *Hyles salangensis* has a wing pattern similar to that of several other species of *Hyles*, such as *H. annei* (Guérin-Méneville, 1839), *H. centralasiae* (Staudinger, 1887), *H. euphorbiae*, *H. robertsi* (Butler, 1880) and *H. stroehlei*, but as these species have proven not to be closely related to each other, it is difficult to use wing pattern to determine the phylogenetic position of *H. salangensis*. Features of the females and larvae cannot be used as these life stages are presently unknown, as is the species' ecology. However, Rothschild & Jordan (1903) considered the degree of development of the arolium, an adhesive pad on the pretarsus between the claws that helps moths cling to smooth surfaces, to be important in the classification of Sphingidae and often used its presence or absence to differentiate genera within the subfamilies Sphinginae and Smerinthinae. However, in the genus *Hyles*, the arolium can be fully developed, vestigial or completely absent in different species, and the

condition of this structure in *H. salangensis* could provide evidence to help resolve the phylogenetic affinities of the species. We also sequenced historic DNA from a leg of the holotype, collected in 1966 and deposited in the LSNK, and from each of seven paratypes, collected in 1965 and deposited in the Vartian collection of the NHMV.

The second enigma, *H. renneri*, was described from high elevation sites in Nepal. It was differentiated from *H. livornica* on the basis of a much darker overall appearance and small differences in the genital morphology in both sexes (Danner et al. 1998), although the wing and abdominal patterns of the two species are almost identical. However, as it is known that changes in colour intensity can be induced by changing the temperature experienced during the pupal stage (de Freina 1994), with lower temperatures producing darker moths, and also that colours fade in old museum specimens, the status of this taxon as a species is thus questionable.

The material studied for the original description included over 100 individuals collected from May to July in 1973, 1995 and 1996 in numerous localities in Nepal, and is mainly deposited in the ZSM and MWM. The species was considered to be endemic to Nepal, with localities being mostly in the Annapurna and Ganesh Himal at elevations between 1600 m and 4650 m, where it was said to replace *H. livornica*. Danner et al. (1998) wrote that its restriction to high elevations and insular distribution indicated that the species could be assumed to be non-migratory. In contrast, *H. livornica* is well known as a strong-flying, migratory species that is widely distributed in Africa, southern Europe and Asia. They further stated that as *H. livornica* occurs in eastern Nepal, the two species could occur sympatrically. However, it is far from clear which species actually occur(s) in eastern Nepal. The

records cited by Danner et al. (1998) derive from Haruta (1994; 1995), and the seven localities given therein range in altitude from 1957 m to 4200 m. Furthermore, one of the *H. livornica* localities, Muktinath (3545 m), is in the Annapurna, at the centre of the distributional range described for *H. renneri*. Thus, this record should probably be attributed to *H. renneri* rather than to *H. livornica*. Only a single specimen was illustrated by Haruta (1994) and being neither obviously very dark nor very pale, it is difficult to attribute it to either of the species.

Another taxon, *H. livornica tatsienluica* (Oberthür, 1916) from Ta-tsien-lou, "Tibet" (now Kangding, Sichuan) in southwestern China also has a dark overall appearance. However, apparently without having studied any material first hand, Danner et al. (1998) treated it as a junior synonym of *H. livornica*. We included two specimens in this study because, contra Danner et al. (1998), its wing coloration is just as dark as that of the holotype, allotype- and paratype specimens of *H. renneri* collected at high elevation (2000–4000 m) in Nepal.

Kitching & Cadiou (2000) took a contrary view to Danner et al. (1998) , arguing that *H. renneri* was not distinct from *H. livornica* because although moths of the former were dark, similar dark specimens could be found elsewhere within the range of *H. livornica* (including several *H. l. tatsienluica* in NHMUK), and thus the colour difference was not diagnostic. Furthermore, from a comparison of the genitalia they considered that the purported diagnostic features in those structures did not hold either. The geographic and altitudinal range of *H. renneri* was later greatly increased when Eitschberger (1999b) recorded a male from Gangu County, Gansu, China, at an elevation of a mere 900 m. This locality is over 2000 km from central Nepal, in a

straight line across the Tibetan Plateau. Consequently, the distributional ranges and possible contact zones of the two species in the Himalaya and surrounding areas to the north and east cannot be considered to be known in any detail yet. Even the question of whether *H. livornica* occurs in Nepal remains open.

The third species whose status is open to question is *H. nepalensis*. Originally described as a subspecies of *H. gallii*, and based on only two specimens from Nepal, it was raised to species status by Danner et al. (1998). The main diagnostic characters were again a much darker overall appearance, compared to *H. gallii*, a reduction of the white patch on the upperside of the hindwing and a stronger white dorsal line on the abdomen. Danner et al. (1998) listed four moths of this taxon from Nepal deposited in the private collection of one of the authors (EMEM), but illustrated a further three from the ZSM (apparently not the types). Danner et al. (1998) gave the distribution of *H. nepalensis* as (our translation from German): "Only known from Nepal and immediately adjacent areas of Tibet, where moths were caught at different localities between 2750 and 3600 m between the 21st of May and 21st of July". However, following an assessment of the variation in adult and larval coloration, and genital structure, Kitching & Cadiou (2000) concluded there was no support for separate species, or even subspecies, status for *H. nepalensis* and synonymized it with *H. gallii*. Moths of *H. gallii* have been reported flying into light traps at subzero temperatures (Danner et al. 1998), indicating their tolerance of a cold climate at high altitude. Since low temperatures during pupal development may lead to darker shades in the adult moths (see above), the species status of *H. nepalensis* is also questionable. *Hyles gallii* is distributed throughout the Holarctic, including localities in the mountains of northwest India (3100 – 4500 m) and northern Pakistan (2300 m), the

most westerly mountain range of the Himalaya-Tibetan Plateau upfolding. For comparison with our samples of *H. nepalensis* (which included the holotype and allotype), we included specimens from the Nearctic and the Ladakh area, specimens determined as *H. gallii* collected at low elevation in Nepal, and the holotype and two paratypes of *H. g. tibetanica* Eichler, 1971 (which Danner et al. 1998 had synonymized with nominotypical *H. gallii*) (Table 1).

### *Larval host plants*

With regard to their choice of larval host plants, the genus *Hyles* includes species that range from the extremely polyphagous (e.g., *H. lineata*) to those that are restricted to a single plant family (e.g., *H. hippophaes* on Elaeagnaceae, *H. zygophylli* on Zygophyllaceae and the species of the *H. euphorbiae*-complex on Euphorbiaceae). It has been suggested that polyphagy is an adaptation to living in arid or semi-arid habitats, where the available hosts are unpredictable, both in time and space, and in terms of their taxonomic (and thus chemical) composition. Consequently, the ability to feed on whatever host plant is present would be advantageous (Physiological Efficiency Hypothesis: Singer 2008 and references therein). In contrast, specialization onto a single host plant family, particularly if that family is well-protected by toxic secondary compounds, could also be evolutionarily beneficial, especially if those chemicals could be detoxified, then sequestered or otherwise co-opted for defence (Enemy-Free-Space Hypothesis: Singer 2008 and references therein). Larval host plant shifts and specialization in butterflies have also been linked with increased speciation rates (e.g., Ebel et al. 2015; Fordyce 2010; Wahlberg et al. 2013) but this has not yet been investigated in *Hyles*. The larval host plants of two of the three enigmas that are the focal species of this study are unknown, so by reconstructing the

evolution of larval host plant use in the genus, we aim to predict what the larval host plants of *H. salangensis* and *H. renneri* are.

# Methods

### *Material and laboratory techniques*

Of the 83 samples sequenced for this study (Table 1), 77 originated from dry museum collections made between the years 1930-2007 and were processed with ancient DNA (aDNA) techniques, due to assumed degradation of historic DNA (i.e., tissue that was not specifically preserved for later DNA analyses and is older than ~5 years). Type specimens were included from all the focal taxa studied (Table 1). Sequences of the mitochondrial genes COI and COII were obtained to merge the present dataset with the sequences analysed by Hundsdoerfer et al. (2005, 2009; Table 1). The 2284 bp targeted were amplified in three fragments of ~860 to ~990 bp for fresh or in 13 fragments of ~110 to ~280 for the historic specimens (details below). More recent tissue was available for the populations of *H. gallii* from Ladakh (Table 1) and Canada (both 2008).

For the recently collected tissue, DNA extracts were obtained using the protocols described by Hundsdoerfer et al. (2005, 2009) and PCR conditions were those used by Hundsdoerfer et al. (2009).

DNA from museum specimens was extracted using an ancient DNA (aDNA) methodology in a dedicated laboratory. Such tissue consisted of either the abdomen,

the contents of which were digested enzymatically following to the procedures described by Hundsdoerfer & Kitching (2010), or one or two legs, which were finely chopped with small scissors prior to addition of lysis buffer from a sbeadex forensic kit (LGC Genomics, Berlin). Manufacturers' instructions were followed except that the proteinase-K digestion was left overnight and the final elution was performed with 60 µL of elution buffer. The genes were amplified in up to 13 small, overlapping fragments with specifically designed *Hyles* primers (Table 3). The size of the amplified fragment depended on the degree of degradation of the DNA and ranged from longer fragments of 300–600 bp (fragments 1a, 1b, 2a, 2b, 3a, 3b; Table 3) to shorter ones of only 110–280 bp (fragments A–M). The fragments were amplified either individually or in a two-step multiplex approach to make efficient use of the limited amount of DNA available. To conduct the multiplex PCR, primer mixes were prepared with each primer at a concentration of 2 µM (2 pmol/µL). Primer mix 1 contained fragments A, C, E, G, I, K & M, and primer mix 2 targeted fragments B, D, F, H, J & L. Each multiplex-PCR was performed with 4 µL of DNA extract in a 25 µL volume containing 2.5 µL of primer mix, 12.5 µL of ready-made Type-it Microsatellite PCR Master Mix 2× (Type-it Microsatellite PCR Kit, Qiagen) and 6µL of ultra-pure $H_2O$. Cycler settings were initially 95 °C for 5min, then 40 cycles of 95 °C for 30 s, 57 °C for 1.5 min, 72 °C for 30 s and a final elongation at 60 °C for 30 min. Reamplification was performed in a 20 µL volume containing 1.5 µL of a 1:10 dilution of the multiplex PCR-product, 1 µL of each primer at 10 µM, 0.4 µL of dNTP-mix at 10 mM of each dNTP, 1 unit of Taq polymerase (Bioron DFS Taq, Ludwigshafen, Germany), 2 µL PCR buffer 10× incl. 25 mM $MgCl_2$ and 13.9 µL of ultra-pure $H_2O$. Cycler settings for reamplification were initially 95 °C for 5 min, then 40 cycles of 95 °C for 1 min, annealing temperature (see Table 3) for 1 min, 72 °C for

1 min and a final elongation at 72 °C for 10 min. Subsequent laboratory steps, such as purification of PCR products and cycle sequencing, followed Hundsdoerfer et al. (2009). The short sequences were then assembled by hand using BioEdit (Hall 1999).

### *Phylogenetic analyses*

To reconstruct phylogenetic relationships, we added the new sequences into a dataset comprising representative COI (1531 bp) and COII (682 bp) sequences from all *Hyles* species analysed by Hundsdoerfer et al. (2009), except for *H. robertsi*, as its samples formed an unresolved polytomy with those of *H. euphorbiae*. We omitted the intervening 71 bp t-RNA-Leu from the analyses. The final data set comprised 2213 bp for each of a total of 225 sequences.

Bayesian inference of phylogeny was performed using BEAST v.1.8.1 (Drummond et al., 2012) and run for 50,000,000 generations (trees sampled every 5,000 generations) under the uncorrelated lognormal clock model for all loci with the "auto-optimize" option activated and a birth-death process prior (with incomplete sampling assumed) applied to the tree. The log files were examined with Tracer v1.4.8 (Drummond & Rambaut 2007) to ensure effective sample sizes (ESS). Trees were summarized with TreeAnnotator v1.8.1 (posterior probability limit= 0.5) using a burn-in value of 3,000 (trees) and the median height annotated to each node.

For reconstruction of the input tree for ancestral range estimation we calculated a dated tree using only one specimen for each *Hyles* species (clade), resulting in 22 taxa. The data comprised COI, COII and EF1alpha (772 bp; from Hundsdoerfer et al. 2009) sequences with a total aligned length of 2985 bp (Table S3). EF1alpha

sequences were missing for *H. renneri* and *H. salangensis* (filled with Ns). The best partitioning scheme and evolutionary models were determined with PartitionFinder 1.1.1 (Lanfear 2012) using the AICc criterion, the 'beast' model set and the greedy search algorithm. The best scheme was determined to be a seven-partition scheme; see Table S1 for details.

Tree estimation was undertaken with BEAST 1.8.2 (Drummond et al. 2012) using a chain length of 11 million states with trees samples every 1000$^{th}$ state. The tree prior was set to the Birth-Death model (Stadler 2009). A maximum likelihood tree was calculated with RAxML (Stamatakis 2014) and applied as the starting tree to provide a valid initial state. BEAST log files were examined for convergence and sufficient effective sample sizes with Tracer v1.6. The first 1000 trees were subsequently removed as burn-in. The remaining 10000 trees were summarized using TreeAnnotator 1.8.2 and common ancestor heights were annotated on the maximum clade credibility tree. Additionally, we visualised the entire tree sample as a graphical cloud representation using Densitree v2.2.1 (Bouckaert & Heled 2014) (Fig S1). All topologies recovered in our analysis are illustrated including branch lengths. This plot shows the uncertainty with regard to topology and branch lengths (i.e., node age). Uncertainty observed at nodes with a posterior probability of one pertains exclusively to node age and not to topology.

For inference of divergence time estimates we calibrated our phylogeny using three calibration points (positions in the tree and time intervals are shown in Fig. 2). Because of a lack of reliable fossils we calibrated the root node (1) with a previous mean age estimate for *Hyles* from the fossil-biogeographic dating by Kawahara &

Barber (2015). We applied a second calibration (2) to the split of the Hawaiian *Hyles* clade from its sister crown clade corresponding to the volcanic age of the oldest islands of the Hawaiian archipelago Kauai and Niihau (5.1 Ma) (see dating approaches in Fleischer et al. 1998, Lerner et al. 2011). The two endemic species from Hawaii, *H. calida* and *H. perkinsi*, are restricted to the younger central island group of Oahu and Molokai (*H. calida* is also present on Kauai). We therefore applied a third calibration (3) to the split among the two Hawaiian endemics according to the younger volcanic age of 3.7 Ma for the respective islands (see Fleischer et al. 1998). As recommended by Drummond and Boukaert (2015) we validated our priors by performing BEAST runs while sampling from the prior only. We performed two test runs, one with node ages applied as hard maximum ages through uniform priors and a second run where we applied normal priors and set the upper 97.5% confidence interval to represent the desired boundary. From these runs it was evident that the latter approach best reflected our prior information and the use of uniform priors is not appropriate as they resulted in implausibly young age estimates. Calibration information was coded into priors as follows: root: normal prior (mean=7.5, sd=1); age of Hawaii archipelago: normal prior (mean=3.14, sd=1); age of younger Hawaiian islands: normal prior (mean=2.5, sd=0.612). The root prior was additionally nested within a uniform prior restricting the lower age to 5.1 Ma to avoid conflicts with the other priors.

### *Ancestral range estimation*

Ancestral distribution ranges of clades were estimated using the R package BIOGEOBEARS 0.2.1 (Matzke 2013, 2014a). The range of *Hyles* was divided into 10 regions (map in Fig. 3, Table 4) and the ranges of taxa in our tree coded in a

presence/absence matrix (Table S2, Supplement). We used BioGeoBEARS to test the fit of three commonly used biogeographic models on our data: Dispersal-Extinction-Cladogenesis (DEC) (Ree and Smith 2008); maximum likelihood versions of dispersal-vicariance analysis (DIVALIKE; Ronquist 1997); and Bayesian biogeographical inference (BAYAREALIKE; Landis et al. 2013). In addition, we tested whether models that allow founder event speciation (+j models; Matzke, 2014b) had a better fit on our data. We employed a stratified analysis, disallowing occurrence on Hawaii prior to 5.1 Ma. Model selection was carried out using the AICc criterion.

### *Ancestral state reconstruction*

We used BEAST v. 1.8.2 to reconstruct ancestral states of host plant use and the condition of the arolium concurrently with estimation of phylogeny. Character matrices (Table S4, Supplement) were coded by AKH and IJK and imported into BEAST as a discrete trait partition (if the host plant was unknown it was coded as missing and the state was reconstructed). Larval host plant data was extracted from an extensive literature database maintained by IJK that also includes personal communications from amateur breeders. Output data were extracted from the maximum clade credibility tree using custom R scripts, pie charts were plotted onto trees using the R package ggtree.

## Results

Newly generated sequences of 33 Palaearctic *H. gallii*, 14 Nearctic (formerly named) *H. g. intermedia*, 3 (formerly named) *H. g. tibetanica*, 7 (formerly named) *H. nepalensis*, 4 *H. livornica*, 14 *H. renneri* and 8 *H. salangensis* specimens were included in the analyses. Of all included sequences, the one with the highest proportion of missing data was # 9468 (the paratype of *H. gallii tibetanica*) with 73% of positions missing. We added a further 42 DNA barcode sequences downloaded from the public BOLD database (http://www.boldsystems.org/) or deposited in GenBank by BOLD for this study (11 Palaearctic *H. gallii*, 17 Nearctic (formerly named) *H. g. intermedia*, 1 *H. g. tibetanica*, 9 *H. livornica* and 4 *H. renneri*). The entire dataset of new sequences was deposited in BOLD (see Table 1 for GenBank accession numbers).

### *Phylogeny & Systematics*

*Hyles* is reconstructed as a robust monophylum (Fig. 1) aged 6.6 Ma B.P (5.1–8.0 Ma; Fig. 2). The species from the Nearctic, Neotropics, Madagascar, Australia and Hawaii branch off first and form the basal grade. The crown group includes all the Palaearctic species, including the migratory *H. gallii* (also occurs in the Nearctic) and *H. livornica* (also occurs in Africa and on Madagascar). Our study provides molecular data of a Madagascan *H. livornica* specimen for the first time, and confirms the synonymy with this species of the former *H. malgassica* (Denso, 1944) (*contra* Eitschberger 1999a and Eitschberger & Surholt 1999).

*Hyles salangensis* is a strongly supported lineage that forms the sister group of a clade comprising the Palaearctic species *H. nicaea* + the Holarctic *H. gallii/H. nepalensis*

clade (Figs. 1, 2). Together, this well supported clade (Bayesian posterior probability, pp = 0.96) forms the sister group of all other Palaearctic species.

*Hyles gallii* is a fully supported clade (Fig. 1) but the samples from different geographical areas are scattered among the subgroups, none of which has any support. The seven samples of the former *H. nepalensis* (# 1404–1407, 4366–4368 including the holotype and allotype), and five of the six samples from Nepal that were identified as *H. gallii* (# 1729, 6493, 6498, 6511, 6519), cluster into one poorly supported group (Fig. 1, group A, pp = 0.91), together with *H. gallii* samples from China, Mongolia, Russia, Kazakhstan and Europe. Ten of the 31 Nearctic samples of the former *H. g. intermedia* from Alaska, Canada and U.S.A. also cluster in this group. The majority of Nearctic samples form a second group (Fig. 1, group B; pp < 0.5%) together with nine *H. gallii* specimens from northern India (Kashmir, Ladakh) and Pakistan (Ladakh # 4774, 4777, 4779, 4783, 5384, 6499, Pakistan # 6518, 11260), the three types of the former *H. g. tibetanica* (# 9452, 9468, 9469) and another *H. gallii* from Tibet (# 6521). The sequences of all three *H. g. tibetanica* types obtained by us were identical. BOLD sequence SPUEB497-07 (BC-EMEM-1437) is also from the holotype of *H. g. tibetanica*. This specimen was sequenced independently by us as #9469 and apart from the Ns the two sequences do not differ. Groups A and B form a clade, but without support (pp < 0.5). A further specimen from Nepal (# 1730) that was identified as *H. gallii* appears as sister to groups A + B, and *H. gallii* as a whole forms a clade with full support (green branch in Fig. 1). Internal differentiation within the *H. gallii* (*sensu lato*) clade is negligible; none of the subclades within groups A and B has high pp support and the average pairwise sequence divergence is only 0.5% (p-distance; pairwise deletion). Values greater than 2% occur in only five comparisons

involving # 6511 from Nepal and over 3% only between # 6511 from Nepal and # 6521 from Tibet (data not shown, the two individuals are indicated with arrows in Fig. 1).

The samples of *H. renneri* and *H. livornica* form two well defined, but mixed clusters, each with very strong support (pp = 1 ; Fig. 1). Two moths from Nepal determined as *H. renneri* cluster with *H. livornica* (Fig. 1: # 4787, 4789). The ten other *H. renneri* moths (from Nepal and the neighbouring Tsou-La-Pass in Tibet; 2020–4500 m a.s.l.; Table 1), including the holotype, allotype and three paratypes, form the fully supported sister clade (Fig. 1). The two specimens of high elevation *H. livornica tatsienluica* (# 4757 and the syntype # 6524) group in the *H. renneri* clade.

### *Ancestral range analysis and dating*

Model selection with BioGeoBEARS showed the BAYAREALIKE+J model to have the best fit on our data (Table 2). Ancestral ranges at nodes towards the root of our *Hyles* phylogeny were reconstructed as highly equivocal (Fig. 3), but the Neotropics has the highest probability for the two oldest nodes. Indeed, the three terminal species of these two nodes all occur in South America, two of them exclusively. According to our dating approach the radiation of *Hyles* began in the late Miocene (Fig. 2) with the two New World clades as the earliest offshoots.

In the early Pliocene (mean: 4.67 Ma, 95% highest posterior density confidence interval, HPD: 3.49–6.06 Ma), an Old World clade of the Southern Hemisphere split from a large clade of species occurring exclusively in the Northern Hemisphere. The clade formed by the Madagascan and Australian endemics, *H. biguttata* and *H.*

*livornicoides*, is only poorly supported (Fig. 1; pp = 0.86). An alternative scenario, differing from that of the best tree (i.e., the maximum clade credibility tree, red in Fig. S1, Supplement) would be a successive and sequential divergence of these two taxa over a short period of time. This topology also occurs frequently in the total tree sample, as denoted by numerous fine blue lines (each representing single trees of the total sample) in the tree cloud representation condensing into two separate, stronger blue lines, the first forming a branch from the stem to *H. biguttata*, and the second a branch from the stem to *H. livornicoides* (Fig. S1, Supplement).

The separation of the Hawaiian lineage from its sister clade occurred in the middle to late Pliocene (mean 4.01 Ma; 95% HPD: 2.95–5.12 Ma). The biogeographic origin of the clade containing the Hawaiian species (*H. perkinsi* and *H. calida*) and the Palaearctic species is estimated to have been in the Palaearctic region. The split between the two island endemics occurred towards the end of the Pliocene or in the early Pleistocene (mean: 2.21 Ma; 95% HPD: 1.53–2.94 Ma). As expected, both of these divergences were dated as being more recent than the maximum age of the respective islands.

The sister clade of the Hawaiian lineage comprises a large group of 15 species representing a Palaearctic radiation that was dated to 3.53 Ma (95% HPD: 2.59–4.58 Ma). With the exception of the migratory *H. livornica* and *H. gallii*, the species of this clade have purely Palaearctic distributions. Both *H. livornica* and *H. gallii* show very recent (<~1 Ma) dispersal events into southern Africa/Madagascar and the Nearctic respectively. The ancestral area of this clade was confidently estimated to be in the Palaearctic (Fig. 3).

Two subclades can be identified within the Palaearctic clade, the smaller including only *H. salangensis*, *H. gallii* and *H. nicaea*, and the larger comprising the remaining twelve species. The estimated geographic origin of both these clades is the entire Palaearctic region, but within which the Himalaya shows the lowest likelihood. The larger clade has a higher likelihood of an ancestral range in the Eastern Palaearctic. Both Palaearctic subclades diversified around the beginning of the Pleistocene (taking into account the 95% HPD: small clade 3.75–1.83 Ma, large clade 3.98–2.15 Ma) and continued to diversify during the Pleistocene, to produce species including rather narrow range endemics in Central Asia and/or the Himalayas and wide-ranging migratory species.

***Ancestral State Reconstruction: Phylogeny to ecology & morphology***

Reconstruction of ancestral larval host plant use (Fig. 3) returned an unambiguous result for all but three nodes, in that the largest segment is > 50% (i.e., one host plant character state represents the majority). Larval host plant use at the origin of *Hyles* was reconstructed as "polyphagous", as were most of the deep nodes within *Hyles* with the exception of the three ambiguous nodes. Three shifts to other plant families occur on internal nodes: to Rubiaceae in the Hawaiian clade: to Asphodelaceae in *H. siehei* + *H. centralasiae*; and to Euphorbiaceae in the Palaearctic crown clade. All other shifts in host plant use occur on terminal branches, including a second, independent shift to Euphorbiaceae in *H. nicaea*. The larval host plant ranges of *H. salangensis* and *H. renneri* are both unambiguously reconstructed as most likely being polyphagous.

With regard to the arolium, ancestral state analysis produced an unambiguous reconstruction at all nodes (Fig. S2). Complete loss of the arolium occurred at the node comprising the species from *H. vespertilio* to *H. tithymali*, with a second, independent loss along the branch leading to *H. salangensis*. A fully developed arolium was developed in the Hawaiian clade. The vestigial state was reconstructed for all other nodes, including the root.

# Discussion

### *Phylogeny and biogeography of the three enigmatic taxa*

***Hyles salangensis*** The phylogenetic position of *H. salangensis* is perhaps surprising in view of the similarity of its forewing pattern to those of several other *Hyles* species that occur in the same general area, particularly *H. stroehlei*, *H. robertsi* and, to a lesser extent, *H. nervosa* (Kitching 2016). Some *H. stroehlei* resemble *H. salangensis* in having paler scales near the costa and a narrower postmedial band. Other *H. stroehlei* individuals appear extremely similar to *H. nervosa* (not analysed in this study), in that the scales near the costa are rather dark olive brown and the veins traversing the broad postmedial band are highlighted with pale scales (Kitching 2016). Other *H. salangensis* are very similar to *H. robertsi*, which occurs in eastern Afghanistan as subspecies *H. r. elisabethae* Ebert, 1996. *Hyles salangensis* has also been observed to 'resemble a dark hybrid between *H. hippophaes* and *H. euphorbiae*' (A.R. Pittaway, personal communication confirmed 27.09.2012). However, the phylogenetic positions of both *H. stroehlei* and *H. robertsi* are very far from *H. salangensis* (Hundsdoerfer et al. 2009, 2011; see also Fig. 1). We cannot comment on *H. nervosa*, as we have not yet been able to include any confirmed samples in an

analysis. Given the data currently available, there is no doubt that *H. salangensis* is a clearly delimited species that is closely related to *H. gallii* and *H. nicaea* (Figs. 1, 2).

The complete lack of an arolium on the pretarsus in *H. salangensis* suggested that the species might be closely related to a subclade of the larger Palaearctic clade, i.e., the the species from *H. vespertilio* to *H. tithymali* clade (Figs. 1–3, full support), members of which also lack this structure. In contrast, an arolium is present in all other *Hyles* (usually vestigial but fully developed in the Hawaiian clade). However, our ancestral state reconstruction indicated that the arolium had been independently lost in this Palaearctic subclade and in *H. salangensis*.

The ancestral biogeographic origin of *H. salangensis* was reconstructed as the Central Palaearctic. However, with so very few individuals collected, the full extent of its distributional range remains unknown. The only reliably identified specimens are the 15 males collected in the mid-1960s that constitute the type series. Although this rare species appears to be endemic to the mountain ranges of northeast Afghanistan, more extensive sampling might uncover further occurrences. The species' host plant is unknown, but it was reconstructed as "polyphagous" with around 65% likelihood.

***Hyles livornica/H. renneri*** The studied samples determined as *H. renneri* and *H. livornica* form two mixed clusters that are inconsistent with their *a priori* identifications based on a combination of adult external appearance and collection site data (in particular, altitude) (Fig. 1). The two moths that cluster with *H. livornica* but were initially determined as *H. renneri* (#4787 from 1250 m a.s.l. and #4789 from 3000 m a.s.l.) (Fig. 1, Table 1) do appear somewhat paler in wing coloration when

compared with the moths that form the *H. renneri* clade, including the types. This implies that these two moths had been misidentified as *H. renneri*. One of these paler moths (#4787) was collected only 50 km from the type locality of *H. renneri*, but the latter is 2750 m higher. The second (#4789) was captured at 3000 m a.s.l., the same altitude as *H. renneri* #4788, but the two sites are almost 200 km apart.

We were also able to verify the diagnostic features of the female genitalia of *H. renneri* (much larger overall, with a longer signum compared to *H. livornica*) given by Danner et al. (1998) for the moths from Bagarchap (#4375, 2200 m) and Tsou La Pass (#4373, 4500 m), thus confirming the identification of moths from the latter locality by U. Eitschberger and extending the range of the species into southern Xizang Zizhiqu (Tibet). As noted above, the Nepalese female from Kakani (2070 m) (NHMUK sphingid genitalia preparation #963) studied by Kitching & Cadiou (2000), and which in part led them to synonymize *H. renneri* with *H. livornica*, has small genitalia and a short signum. It is also a pale moth. They were thus correct in determining this specimen as *H. livornica*. However, they were unable to study first-hand the genitalia of a high elevation dark female and so missed the crucial distinction in size and signum length, leading them incorrectly to synonymize *H. renneri*. Nor, however, were Danner et al. (1998) entirely correct to separate the two species solely on geography for it is now clear that *H. renneri* and *H. livornica* both occur in central and east Nepal.

On the basis of the Nepalese and Tibetan samples studied here, *H. renneri* appears to be a higher elevation species occurring down to about 2000 m, whereas *H. livornica* is a more low elevation species, but ranging up to 3000 m. Thus the two species

potentially occur sympatrically between 2000 and 3000 m, although interestingly they have not yet been collected together in the same locality. It is also worth noting that Danner et al. (1998) listed three moths caught at Pokhara (1600 m) as paratypes of *H. renneri*, thus lowering the elevational range of *H. renneri* still further but we were unable to check the intensity of their wing coloration (see also discussion below). Thus, within Nepal at least, wing coloration does appear to provide the means to separate the darker *H. renneri* from the paler *H. livornica*, with female genital structure providing confirmation.

The two sequenced specimens (including a female syntype, #6524) of *H. livornica tatsienluica* also have dark wing coloration, and both clearly fall within the clade in which all other sequences are from moths determined as *H. renneri* (including the holotype and several paratypes). These two specimens were, however, collected 1800 km from the type locality of *H. renneri*, and thus they represent a major extension to the distribution range of this species from its previously known localities in Nepal and southern Xizang. We examined the genitalia of syntype #6524 (NHMUK specimen BMNH # 812462, NHMUK sphingid genitalia preparation #963) but they were too damaged to confirm their structure (the corpus bursae was missing). However, a third female in the NHMUK from Ta-tsien-lou (NHMUK specimen BMNH # 812888, NHMUK sphingid genitalia preparation #962) proved to have the large genitalia and long signum diagnostic of *H. renneri* (and the locality is also at high elevation, ~3000 m a.s.l.). Unfortunately, tissue of this moth was not available for sequencing. Nevertheless, we consider that the DNA evidence from syntype #6524, together with that of one other specimen from the type locality (Fig. 1; see above), together with the morphological evidence provided by the genitalia of BMNH # 812888, is sufficiently

strong to justify reinstating *Hyles tatsienluica* **stat. nov.** as a valid species and synonymizing *Hyles renneri* **syn. nov.** with it.

However, the situation regarding both forewing colour intensity and locality elevation is less clear cut when a broader perspective is taken. A female in the NHMUK from "Poo Bashahr state, Schipki-la, 4000m" (now Shipki La, a pass on the NW India-Tibet border) has a dark forewing pattern but the small genitalia and short signum of *H. livornica*. This shows that in this area, *H. livornica* can reach the same extremely high elevations as *H. tatsienluica* does in Nepal, but also suggests that when it does so, its general coloration darkens. A second dark female in the NHMUK from "Amur" also has the genitalia of *H. livornica*. Although the locality is very vague, there is very little land above 1000 m in Amurskaya and Khabarovsk Krai, and so the moth is almost certainly from a low elevation, though possibly from a relatively cool site. This then calls into question the identification of the low elevation (900 m) "fresh" male from Gangu County, Gansu, China by Eitschberger (1999) noted earlier. Although we now accept the differences in female genitalia between *H. livornica* and *H. tatsienluica*, we concur with Kitching & Cadiou (2000) that the purported diagnostic differences in the male genitalia do not hold and confirmation of identity cannot rest solely on wing colour.

In summary, *H. tatsienluica* is a valid species, separate from *H. livornica* that can be distinguished on female (but not male) genital morphology and COI–COII sequences. Where the two species occupy the same general area in Nepal (and maybe SW China), *H. tatsienluica* is darker and occurs at higher elevations, whereas *H. livornica* is paler and flies at lower elevations, but with a considerable overlap between about

2000 m and 3000 m. This may be due to seasonal movements of *H. livornica* to higher altitudes in summer or perhaps hill-topping behaviour in males (as is suspected to occur in other Sphingidae; Holloway 1987). The species is a well-known migrant in Africa and Western Europe and there is no reason to suspect such behaviour is absent in the Nepalese populations, though its ecology remains to be investigated. The type locality of *H. renneri* would seem to be near the far western end of the distribution of *H. tatsienluica* and is simply the place where the species is currently most accessible. The Tsou-La Pass locality in southern Xizang may be near the centre of the species' range, connecting the western end in Nepal with Ta-tsien-lou, which may represent the eastern end of its distribution. *Hyles tatsienluica* would thus appear to be a Tibetan Plateau species, but the exact limits of its range have yet to be fully determined.

With respect to the distribution range of *H. livornica*, contra Eitschberger (1999a) and Eitschberger & Surholt (1999) and in agreement with the revision by Kitching & Cadiou (2000), our data confirm the synonymy of *H. malgassica* with *H. livornica* and the occurrence of that species on Madagascar.

***Hyles gallii*** In contrast to the situation with *H. tatsienluica* and *H. livornica*, the seven individuals of the former *H. nepalensis* studied (all of which come from very high elevations of 3000–4000 m) show no distinct cluster formation. They group among *H. gallii* from the UK, Germany, Czech Republic, Finland and China (Shaanxi) (Fig. 1) with very little genetic differentiation. The type locality of *H. gallii* is Germany and as the present clade includes two samples from that country, as well

as the holotype of *H. nepalensis*, we consider it to represent the "true" *H. gallii* and hence agree with Kitching & Cadiou (2000) that the name *H. nepalensis* is a synonym of *H. gallii*.

Nearctic *H. gallii* can be separated morphologically from most Palaearctic *H. gallii* by the colour of the hindwing medial band and anal patch. Generally, in Palaearctic moths, the pink medial band distal to vein $CuA_2$ is largely white and the anal patch is large, circular and white. The hindwings of moths from Nepal and from Lanak-la in NW Xizang are likewise pale and so conform to the Palaearctic phenotype. In contrast, the entire medial area in Nearctic moths is more uniformly suffused with pink and the anal spot reduced in size (see illustrations in d'Abrera 1987: 184) and this was the main feature that has been used to justify subspecies status for the Nearctic populations as *H. g. intermedia* (Kirby, 1837). However, the hindwings of *H. gallii* specimens from the Ladakh region of NW India near the border to Pakistan, about 850 km from the type locality of *H. nepalensis*, also show the same pink-suffused hindwing medial band as Nearctic moths, and it was just such specimens that Kitching & Cadiou (2000) used as evidence to justify the synonymy of *H. g. intermedia*. Thus, we were most surprised to find that while some Canadian samples grouped with the phenotypically similar Ladakh moths in group B (Fig. 1), others were scattered among samples from lowland Palaearctic sites in group A (Fig. 1). DNA sequence data provide no support for *H. g. intermedia* as a separate Nearctic subspecies, suggesting an apparently very recent colonisation of the Nearctic, or even for a broader geographic group based on a pinker hindwing medial band. Indeed, the low level of differentiation within the *H. gallii*-complex as a whole reinforces the

rejection, and confirms the synonymies, of all subordinate taxa: *H. g. intermedia*, *H. nepalensis* and *H. g. tibetanica*.

***The origin of Hyles and possible early colonization routes***

The genus *Hyles* was first postulated to be of New World origin during the Oligocene/Eocene and from there two major radiations across the Southern and the Northern Hemisphere were suggested (Hundsdoerfer et al. 2005, 2009). However, the estimated ancestral ranges for the three basalmost nodes of our *Hyles* phylogeny were highly equivocal (Fig. 3), so our analyses did not confirm or disconfirm such a scenario, calling into question yet again the geographic origin of *Hyles* hawkmoths. In addition, the temporal scenario must also be reconsidered, because recent molecular dating approaches have suggested a much more recent origin and radiation of *Hyles* hawkmoths. Although examples of Oligocene to Early Miocene globally-distributed butterfly radiations exist (*Junonia* Hübner, 1819, Nymphalidae, Kodandaramaiah & Wahlberg 2007; *Mycalesina* Reuter, 1896, Nymphalidae, Aduse-Poku et al. 2015), the origin of *Hyles* was recently dated by Kawahara & Barber (2015) to the Middle to Late Miocene, when proto-*Hyles* separated from the Neotropical *Xylophanes* Hübner, [1819] and African *Chaerocina* Rothschild & Jordan, 1903 about 10 Ma. During this period several events of transcontinental faunal interchange between the Old World, particularly East Asia, and the Nearctic via Beringia have been suggested for some butterfly genera (*Papilio* Linnaeus, 1758, Papilionidae, Wu et al. 2015; *Oeneis* Hübner, 1819, Nymphalidae, Kleckova et al. 2015; *Polyommatus* (von Rottemburg, 1775), Lycaenidae, Vila et al. 2011).

Diversification within *Hyles* itself, however, has to be considered much more recent and was presumably unrelated to Miocene climatic oscillations and successive submergence or emergence of land bridges. According to our time calibrated phylogeny, the global *Hyles* hawkmoth radiation began in the Late Miocene.

Australia and Madagascar were colonised by the poorly supported clade comprising *H. livornicoides* and *H. biguttata*, which split from all other *Hyles* 4.7 Ma B.P. (3.5–6.1; Fig. 2). Colonisation routes have been previously been postulated via Antarctica or Fiji (Hundsdoerfer et al. 2005) under the assumption of the alternative tree topology in which the Australian and Madagascan species branch off successively (Fig. S1, Supplement). The route via Antarctica still remains a plausible scenario for both Australia and Madagascar with the topology recovered in this study, in which *H. livornicoides* and *H. biguttata* are sister species. Although ice sheets have been present on Antarctica over the past 40 million years, they have been extremely dynamic (Zachos et al. 2001). A major ice sheet had re-established on Antarctica by 10 Ma, but the early Pliocene is marked by a subtle warming trend between 6 and ~3.2 Ma (see Zachos et al. 2001 and references therein). Thus, Antarctica may have been able to serve as at least a stop-over for the ancestors of this clade. The alternative hypothesis of a westward dispersal via (proto-) Polynesian, Melanesian, Micronesian islands now seems less likely, considering the newly recovered sister group relationship between *H. biguttata* and *H. livornicoides*. However, this topology received poor support in BEAST analyses and the ancestral range was estimated to be highly equivocal, thus leaving this aspect of *Hyles* biogeography open and requiring further study.

In contrast to previous hypotheses, our results unequivocally show that the ancestors of the Hawaiian clade colonised Hawaii from the Palaearctic and not the Nearctic. However, in our opinion this pattern is an artefact of the ancestral range analysis caused by the lack of differential weighting of adjacent relative to more distant ranges, or simply to the imbalance in the number of Palaearctic and Nearctic taxa (15 vs. 3).

***Pliocene origin and Pleistocene diversification in the Palaearctic***

*Hyles* hawkmoth radiation culminated in a burst of diversification towards the end of the Pliocene (starting at ~4 Ma) resulting in no fewer than fifteen extant Palaearctic species, and from where two migratory species, *H. gallii* and *H. livornica*, more recently colonized the Nearctic and Africa/Madagascar respectively. At a first glance, our age estimate of ~6.5 Ma for *Hyles* appears quite young considering that during this short time interval *Hyles* hawkmoths colonized all continents. However, similar temporal scenarios have been reconstructed for a number of butterfly taxa. For example, from the early Pliocene onwards, three subgroups of cold-adapted butterflies of the genus *Oeneis* diversified and progressively dispersed from Central Asian ancestral ranges into the Northern and Western Palaearctic and the Nearctic (Kleckova et al. 2015). Similarly, at about 5 Ma several "Out-of-Asia" dispersals into Africa, Australia and the New World were hypothesized for butterflies of the genus *Junionia* (Kodandaramaiah & Wahlberg 2007). For these global butterfly radiations dispersal was considered a more important factor than vicariance (Kodandaramaiah & Wahlberg 2007) and climatic niche diversification was shown to promote speciation processes in cold-adapted taxa (Kleckova et al. 2015). These findings are paralleled by a number of Pliocene bursts of diversification in birds. At about 6 Ma *Turdus*

Linnaeus 1758 thrushes began diversifying in a global radiation that included repeated transatlantic dispersals (Voelker et al. 2010), and an early Pliocene onset of an "Out-of-Africa" radiation including several transcontinental dispersals gave rise to sixteen species of *Apus* Scopoli, 1777 swifts (Päckert et al. 2012; Tietze et al. 2015).

Furthermore, the effect of global climate cooling towards the end of the Pliocene is most striking in mountain ecosystems, where cold adapted species started occupying the newly emerging temperate and boreal niches, for example, in the Himalayas (global review in Fjeldså et al. 2012; birds: Päckert et al. 2012). Particularly in birds, there is evidence that the elevational aspect of the niche occupied by species evolves late during evolutionary history, with a steep increase in elevational disparity occurring from about 5 Ma to a peak at the beginning of the Pleistocene (Price et al. 2014). Elevational parapatry of *Hyles* hawkmoths in the Himalayas is complex and also dates back to a mid-Pliocene lineage separation. The elevational ranges of the sister species *H. livornica* (low elevations) and *H. tatsienluica* (moderate to high elevations up to alpine environments at 4000m and above) partly overlap. *H. gallii* and *H. nicaea lathyrus* (Walker, 1856) occur in local sympatry depending on the abundance of the larval host plant, *Euphorbia stracheyi* (Kitching & Cadiou 2000, Smetacek & Kitching 2012).

Furthermore, *Hyles* diversification during the Pleistocene epoch was limited exclusively to the Northern Hemisphere. Examples of Pleistocene diversification in other Lepidoptera mostly refer to single species or species pairs that (like *Hyles* hawkmoths) split into several genetic lineages across wide distribution ranges. One of the most widely distributed butterfly species, *Lampides boeticus* (Linnaeus, 1767) (Lycaenidae), started diversifying in the Early Pleistocene (2.3-1.5 Ma) and its

demographic history and extant distribution across the entire Old World were apparently shaped by the species' ability to undertake regional migrations (Lohman et al., 2008). In *Melitea cinxia* (Linnaeus, 1758) (Nymphalidae), seven phylogroups correspond to a characteristic circum-Mediterranean diversification that includes outliers in the East Palaearctic and western China (Wahlberg & Saccheri 2008). Other examples of more recent late Pleistocene lineage separations were found in three *Parnassius* Linnaeus 1758 (Papilionidae) species (Gratton et al. 2008, Todisco et al. 2012), the genus *Maniola* Linnaeus 1758 (Nymphalidae; Kreuzinger et al. 2015) and *Lopinga achine* (Scopoli 1763) (Nymphalidae; Kodandaramaiah et al. 2012). Even the enigmatic intraspecific structure found in *H. gallii* with two very young haplotype lineages distributed across the entire Holarctic is partly reflected in other butterfly groups, e.g. *Coenonympha tullia* (O. F. Müller, 1764) and allies (Kodandaramaiah & Wahlberg 2009) and four species (-pairs) within the Holarctic crown clade of Polyommatini (Lycaenidae, Vila et al. 2011). Alpine butterflies of the Holarctic *Parnassius phoebus* complex also diverged into a Eurasian-Beringian and a Nearctic lineage within the last 100,000 years (Todisco et al. 2012). However, *Hyles gallii* differs from these latter two examples in the lack of any clear phylogeographic structure within or among Eurasian and Nearctic haplogroups, which might be explained by a higher level of gene flow among *Hyles* hawkmoth refugia due to their greater dispersal ability. The absence of support for a pattern of vicariance in the *H. gallii* clade contrasts markedly with the support found for intraspecific East-West splits in other clades, e.g., the Western and Central Palaearctic lineages of *H. hippophaes* and *H. nicaea* (not shown in Fig. 1 due to collapse of the clades concerned). A Pleistocene origin of Palaearctic East-West disjunctions is frequently suggested in passerine birds (Haring et al. 2007, 2012; Päckert et al. 2012) but has

also been corroborated by divergence time estimates in butterfly species pairs/ species triplets from the Western, Central and/or Eastern Palaearctic (*Oeneis glacialis* Moll, 1763, *Oe. norna* (Thunberg, 1791) and *Oe. fulla* (Eversmann, 1851), Kleckova et al., 2015; *Coenonympha* Hübner, 1819, several terminal clades, Kodandaramaiah & Wahlberg 2009).

***Hyles ecology in the light of phylogeny***

Compared with the butterfly taxa discussed above, the burst of Pleistocene diversification in *Hyles* hawkmoths is exceptional. Phylogenetic comparison across butterfly genera suggested that increased speciation rates were frequently associated with host plant shifts in the respective clades (Ebel et al. 2015; Fordyce 2010). In *Hyles*, the Palaearctic stem species was likely polyphagous (Fig. 3). Within the clade reconstructed to have originated in the Palaearctic (including the Hawaiian species pair), specializations onto single plant families were detected at three internal nodes (in addition to four more that are restricted to single species: *H. nicaea*, *H. vespertilio*, *H. zygophylli* and *H. costata*). Utilizing Rubiaceae may have enabled the radiation into three species on Hawaii (*H. wilsoni* was not sampled) and specialization on Elaeagnaceae (and specifically the flowers and young fruits rather than the leaves) may have facilitated the radiation of three species in Central Asia (see Kitching 2016). Utilizing *Euphorbia* as host plant may have led to an increased speciation rate (and at least five species) within the crown group, the *Hyles euphorbiae*-complex (HEC). However, the phylogenetic relationships within the HEC are not supported, and much uncertainty is seen in the tree cloud (Fig. S1, Supplement), corroborating the ambiguity surrounding species boundaries in this group found by Mende et al. (2016). This observation may simply be due to the extremely young age of this group (1.1

Ma, 95% HPD 1.53–0.69 Ma; Fig. 2), one in which speciation processes appear to be still underway. Even so, the data do indicate that accelerated speciation rates could be associated with host plant shifts in the respective clades (as in Ebel et al., 2015; Fordyce, 2010), although this association is much under debate for Lepidoptera (e.g. Hardy & Otto 2014; Janz et al. 2016).

## Conclusions

The enigmas of the three high elevation Himalayan *Hyles* brain-teasers were elucidated in the present study. *Hyles salangensis* and *H. tatsienluica* were both found to be valid species, with *H. renneri* being a junior synonym of the latter. The known range of *H. tatsienluica* was confirmed as extending from the mountains of Southwestern China to southern Xizang/Tibet and Nepal. Identification of specimens from the Himalaya and surrounding areas as either *H. livornica* or *H. tatsienluica* cannot be made based on locality data alone, and even the overall darker appearance of the moths is not an infallible feature of *H. tatsienluica*. At present, only mtDNA sequencing of COI/COII and/or study of the female genitalia appear to allow unequivocal separation of the two species. In contrast, the status of the former *H. nepalensis*, *H. g. tibetanica* and *H. g. intermedia* as junior synonyms of the more widespread *H. gallii* is strongly supported by this study. The low genetic divergence within *H. gallii* indicates very recent trans-Holarctic dispersal and/or ongoing gene exchange across the immense Holarctic distribution range of this species. The vast distribution ranges and migratory behaviour of both *H. gallii* (Western Europe to Eastern North America) and *H. livornica* (Western Europe to Eastern Russia and

Africa, Madagascar) appear to be significant hindrances to ancestral biogeographic analyses in *Hyles*. Our analyses were neither able to confirm nor refute a New World origin of the genus, nor postulate a plausible scenario for the colonisation of both Australia and Madagascar by their respective endemic *Hyles* species. According to our time calibrated phylogeny, the global *Hyles* hawkmoth radiation must be considered much more recent than previously thought: it began in the Late Miocene and culminated in a Pleistocene burst of diversification in the Northern Hemisphere. Our data indicate accelerated speciation rates may be associated with host plant shifts in three clades, for an example utilizing *Euphorbia* as host plant in the crown group, the *H. euphorbiae*-complex.

## Acknowledgements

We thank the curators who allowed us to study the specimens in their collections, especially G. Martin (NHMUK London), Dr. S. Gaal-Hazsler (NHMV Vienna), T. J. Witt, Dr. W. Speidel & T. J. Greifenstein (MWM Munich), Dr. A. Hausmann & U. Buchsbaum (ZSM Munich), and Felix Sperling (UASM Edmonton). Thanks to Giovanny Fagua Gonzalez and Danny Shpeley (Edmonton) for insect leg sampling. Funding: This work was supported by the DFG [grant numbers HU 1561/1-1, 1-2 and PA 1818/3-1]; and the EU's Improving Human Potential program SYNTHESYS [grant numbers GB-TAF-3410, GB-TAF-5177]. Special thanks go to Elisabeth Hintelmann and the prize committee for the R.J.H. Hintelmann Prize (2007) for enabling analyses of the specimens from the museums in Munich (ZSM and MWM).

We thank Anja Rauh (Dresden) for additional lab work to include more samples from the Nearctic and Eastern Palaearctic. Thanks go to Ingo Michalak (Leipzig) for valuable comments and suggestions for an earlier version of the manuscript. We thank Rodolphe Rougerie, Jeremy de Waard and the BOLD consortium, as well as the respective providers, for the DNA barcodes included for comparison.

## References

Aduse-Poku, K., Brattström, O., Kodandaramaiah, U., Lees, D.C., Brakefield, P.M. & Wahlberg, N. (2015). Systematics and historical biogeography of the Old World butterfly subtribe Mycalesina (Lepidoptera: Nymphalidae: Satyrinae). *BMC Evolutionary Biology* 15, 167.

Bouckaert, R.R. & Heled, J. (2014). DensiTree 2: seeing trees through the forest. *bioRxiv*, 012401.

d'Abrera, B. (1986). *Sphingidae Mundi: hawk moths of the world*. Oxon: EW Classey Ltd.

Danner, F., Eitschberger, U. & Surholt, B. (1998). *Die Schwärmer der westlichen Palaearktis. Bausteine zu einer Revision (Lepidoptera: Sphingidae)*. Marktleuthen: Herbipoliana.

de Freina, J. J. (1994). Über Biologie, Morphologie und Taxonomie von *Hyles tithymali deserticola* (Bartel), mit vergleichenden Studien zu *Hyles euphorbiae mauretanica* (Staudinger) (Lepidoptera: Sphingidae). *Entomologische Zeitschrift* 104, 33–41.

Drummond, A.J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 1.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29, 1969–1973.

Drummond, A.J. & Bouckaert, R.R. (2015). *Bayesian evolutionary analysis with BEAST*. Cambridge: Cambridge University Press.

Ehrmann, W.U. & Mackensen, A. (1992). Sedimentological evidence for the formation of an East Antarctic ice sheet in Eocene/Oligocene time. *Palaeogeography, Palaeoclimate and Palaeoecology* 93, 85–112.

Ebel, E.R., DaCosta, J.M., Sorenson, M.D., Hill, R.I., Briscoe, A.D., Willmott, K.R. & Mullen, S.P. (2015). Rapid diversification associated with ecological specialization in Neotropical *Adelpha* butterflies. *Molecular Ecology* 24, 2392–2405.

Ebert, G. (1969). Afghanistan bombyces and sphinges (Sphingidae: Lepidoptera). Reichenbachia 12(5), 37-63.

Eitschberger, U. (1999a). Der Madagassische Linienschwärmer *Hyles* (*Danneria*) *malgassica* (Denso, 1944) - 4. Ergänzung zu "Die Schwärmer der westlichen Palaearktis" (Danner, Eitschberger & Surholt, 1998) (Lepidoptera, Sphingidae) – *Atalanta* 30, 169–173.

Eitschberger, U. (1999b). Zur Verbreitung einiger *Hyles*-Arten in China. 3. Ergänzung zu „Die Schwärmer der westlichen Palaearktis (Danner, Eitschberger & Surholt, 1998) (Lepidoptera, Sphingidae). *Atalanta* 30, 162-168.

Eitschberger, U. & Surholt, B. (1999). Die Arten des Subgenus *Danneria* Eitschberger & Zolotuhin, 1998 in der Gattung *Hyles* Hübner, [1819] 1.

Ergänzung zu „Die Schwärmer der westlichen Palaearktis“ (Danner, Eitschberger & Surholt, 1998) (Lepidoptera, Sphingidae). *Atalanta* 29, 13–24.

Fjeldså, J., Bowie, R.C.K. & Rahbek, C. (2012). The role of mountain ranges in the diversification of birds. *Annual Review of Ecology, Evolution and Systematics* 43, 249–265.

Fleischer, R. C., McIntosh, C. E., & Tarr, C. L. (1998). Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K–Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology*, 7(4), 533-545.

Fordyce, J.A. (2010). Host shifts and evolutionary radiations of butterflies. *Proceedings of the Royal Society of London B* 277, 3735–3743.

Gratton, P., Konopinski, M.K. & Sbordoni, V. (2008). Pleistocene evolutionary history of the Clouded Apollo (*Parnassius mnemosyne*): genetic signatures of climate cycles and a ‘time-dependent’ mitochondrial substitution rate. *Molecular Ecology* 17, 4248–4262.

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., & Hebert, P. D. (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Science* 103(4), 968-971.

Hall, T.A. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41, 95–98.

Harbich, H. (2009). Anmerkungen zur Wolfsmilchschwärmerpopulation (*Hyles euphorbiae* (Linnaeus, 1758) - Komplex) von Malta (Lepidoptera: Sphingidae). *Entomologische Zeitschrift* 119, 51–58.

Hardy, N.B. & Otto, S.P. (2014). Specialization and generalization in the diversification of phytophagous insects: tests of the musical chairs and oscillation hypotheses. *Proceedings of the Royal Society of London B* 281, 20132960.

Haring, E., Däubl, B., Pinsker, W., Kryukov, A. & Gamauf, A. (2012). Genetic divergences and intraspecific variation in corvids of the genus *Corvus* (Aves: Passeriformes: Corvidae) – a first survey based on museum specimens. *Journal of Zoological Systematics and Evolutionary Research* 50, 230–246.

Haring, E., Gamauf, A. & Kryukov, A. (2007). Phylogeographic patterns in widespread corvid birds. *Molecular Phylogenetics and Evolution* 45, 840–862.

Haruta, T. (1994). Moths of Nepal, Part 3. *Tinea* 14 (Suppl. 1), 154–158.

Haruta, T. (1995). Moths of Nepal, Part 4. *Tinea* 14 (Suppl. 2), 89.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrapes fulgerator*. *Proceedings of the National Academy of Science of the United States of America* 101: 14812–14817.

Holloway, J. D. (1987). *Moths of Borneo 3. Superfamily Bombycoidea: families Lasiocampidae, Eupterotidae, Bombycidae, Brahmaeidae, Saturniidae, Sphingidae*. Kuala Lumpur: Southdene Sdn. Bhd.

Hundsdoerfer, A.K. & Kitching, I. (2010). A method for improving DNA yield from century-plus old specimens of large Lepidoptera while minimizing damage to external and internal abdominal characters. *Arthropod Systematics and Phylogeny* 68, 151-155.

Hundsdoerfer, A.K., Mende, M., Kitching, I., & Cordellier, M. (2011). Taxonomy, phylogeography and climate relations of the Western Palaearctic spurge

hawkmoth (Lepidoptera, Sphingidae, Macroglossinae). *Zoologica Scripta* 40, 403-417.

Hundsdoerfer, A.K., Kitching, I.J. & Wink, M. (2005). A molecular phylogeny of the hawkmoth genus *Hyles* (Lepidoptera: Sphingidae, Macroglossinae). *Molecular Phylogenetics and Evolution* 35, 442–458.

Hundsdoerfer, A.K., Rubinoff, D., Attié, M., Wink, M. & Kitching, I.J. (2009). A revised molecular phylogeny of the globally distributed hawkmoth genus *Hyles* (Lepidoptera: Sphingidae), based on mitochondrial and nuclear DNA sequences. *Molecular Phylogenetics and Evolution* 52, 852–865.

Janz, N., Braga, M.P., Wahlberg, N. & Nylin, S. (2016). On oscillations and flutterings—A reply to Hamm and Fordyce. *Evolution* 70, 1150–1155.

Kawahara, A.Y. & Barber, J.R. (2015). Tempo and mode of antibat ultrasound production and sonar jamming in the diverse hawkmoth radiation. *Proceedings of the National Academy of Science* 112, 6407–6412.

Kitching, I.J. (2016, July). *Hyles*. Sphingidae Taxonomic Inventory. Available via http://sphingidae.myspecies.info/taxonomy/term/1276.

Kitching, I.J. & Cadiou, J.-M. (2000). *Hawkmoths of the World: an Annotated and Illustrated Revisionary Checklist (Lepidoptera: Sphingidae).* Ithaca and London: Cornell University Press and The Natural History Museum.

Kleckova, I., Cesanek, M., Fric, Z. & Pellisier, L. (2015). Diversification of the cold-adapted butterfly genus *Oeneis* related to Holarctic biogeography and climatic niche shifts. *Molecular Phylogenetics and Evolution* 92, 255–265.

Kodandaramaiah, U., Konvicka, M., Tammaru, T., Wahlberg, N. & Gotthard K. (2012). Phylogeography of the threatened butterfly, the woodland brown *Lopinga*

*achine* (Nymphalidae: Satyrinae): implications for conservation. *Journal of Insect Conservation* 16, 305–313.

Kodandaramaiah, U. & Wahlberg, N. (2007). Out-of-Africa origin and dispersal-mediated diversification of the butterfly genus *Junonia* (Nymphalidae: Nymphalinae). *Journal of Evolutionary Biology* 20, 2181–2191.

Kodandaramaiah, U. & Wahlberg, N. (2009). Phylogeny and biogeography of *Coenonympha* butterflies (Nymphalidae: Satyrinae) – patterns of colonization in the Holarctic. *Systematic Entomology* 34, 315–323.

Kreuzinger, A.J., Fiedler, K., Letsch, H. & Gill, A. (2015). Tracing the radiation of *Maniola* (Nymphalidae) butterflies: new insights from phylogeography hint at one single incompletely differentiated species complex. *Ecology and Evolution* 5, 46–58.

Landis, M.J., Matzke, N.J., Moore, B.R. & Huelsenbeck, J.P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology* 62, 789–904.

Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012). PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Molecular Biology and Evolution* 29, 1695–1701.

Lerner, H. R., Meyer, M., James, H. F., Hofreiter, M., & Fleischer, R. C. (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology*, 21(21), 1838-1844.

Lohman, D.J., Peggie, D., Pierce, N.E. & Meier, R. (2008). Phylogeography and genetic diversity of a widespread Old World butterfly, *Lampides boeticus* (Lepidoptera: Lycaenidae). *BMC Evolutionary Biology* 4, 301.

Matzke, N.J. (2013). Probabilistic historical biogeography: new models for founder event speciation, and imperfect detection, and fossils allow improved accuracy and model testing. *Frontiers in Biogeography* 5, 242–248.

Matzke, N.J. (2014a). BioGeoBEARS: an R package for model testing and ancestral state estimation in historical biogeography. Methods in Ecology and Evolution

Matzke, N.J. (2014b). Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Systematic Biology syu056.

Meerman, J.C. & Smid, G. (1988). Der *Hyles euphorbiae*-Komplex; die Wolfsmilchschwärmer von Kreta (Lepidoptera: Sphingidae). *Entomologische Zeitschrift* 98, 161–176.

Mende, M.B., Bartel, M. & Hundsdoerfer, A K. (2016). A comprehensive phylogeography of the *Hyles euphorbiae* complex (Lepidoptera: Sphingidae) indicates a 'glacial refuge belt'. *Scientific Reports* 6, 29527.

Päckert, M., Martens, J., Sun, Y.H., Severinghaus, L.L., Nazarenko, A.A., Ting, J., Töpfer, T. & Tietze, D.T. (2012). Horizontal and elevational phylogeographic patterns of Himalayan and Southeast Asian forest passerines (Aves: Passeriformes). *Journal of Biogeography* 39, 556–573.

Pittaway, A.R. (1993). *The hawkmoths of the Western Palaearctic*. London & Colchester: Harley Books.

Price, T. D., Hooper, D. M., Buchanan, C. D., Johansson, U. S., Tietze, D. T., Alström, P., ... & Martens, J. (2014). Niche filling slows the diversification of Himalayan songbirds. *Nature*, 509(7499), 222-225.

Ree, R.H. & Smith, S.A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57, 4–14.

Ronquist, F. (1997). Dispersal-Vicariance Analysis: a new approach to the quantification of historical biogeography. *Systematic Biology* 46, 195–203.

Rothschild, L.W. & Jordan, K. (1903). A revision of the lepidopterous family Sphingidae. Novitates Zoologicae 9 (supplement): cxxxv + 1-972.

Singer, M.S. (2008). *Evolutionary ecology of polyphagy. Specialization, speciation, and radiation. The evolutionary biology of herbivorous insects.* Berkeley and Los Angeles: Univ. California Press.

Smetacek, P. & Kitching, I.J. (2012). The hawkmoths of Ladakh, Jammu & Kashmir, India (Lepidoptera: Sphingidae). *Nachrichten des Entomologischen Vereins Apollo, N.F.* 32, 113-115.

Speidel, W. & Hassler, M. (1989). Die Schmetterlingsfauna der südlichen algerischen Sahara und ihrer Hochgebirge Hoggar und Tassili n'Ajjer (Lepidoptera). *Nachrichten des Entomologischen Vereins Apollo, N.F.* 8, 1–156.

Stadler, T. (2009). On incomplete sampling under birth-death models and connections to the sampling-based coalescent. Journal of Theoretical Biology 261, 58–66.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* btu033.

Tietze, D.T., Wink, M. & Päckert, M. (2015). Does evolution of plumage patterns and of migratory behaviour in Apodini swifts (Aves: Apodiformes) follow distributional range shifts? *PeerJ Preprints* e797v1

Todisco, V., Gratton, P., Zakharov, E.V., Wheat C.W., Sbordoni V. & Sperling F.A.H. (2012). Mitochondrial phylogeography of the Holarctic *Parnassius*

*phoebus* complex supports a recent refugial model for alpine butterflies. *Journal of Biogeography* 39, 1058–1072.

Vila, R., Bell, C.D., Mecniven, R., Goldman-Huertas, B., Ree, R.H., Marshall, C.R., Bálint, Z., Johnson, K., Benyamini, D. & Pierce, N.E. (2011). Phylogeny and palaeoecology of *Polyommatus* blue butterflies show Beringia was a climate-regulated gateway to the New World. *Proceedings of the Royal Society of London B* 278, 2737–2744.

Voelker, G., Rohwer, S., Outlaw, D.C. & Bowie, R.C.K. (2009). Repeated trans-Atlantic dispersal catalysed a global songbird radiation. *Global Ecology and Biogeography* 18, 41–49.

Wahlberg, N. & Saccheri, I. (2007). The effects of Pleistocene glaciations on the phylogeography of *Melitaea cinxia* (Lepidoptera: Nymphalidae). *European Journal of Entomology* 104, 675–684.

Wahlberg, N., Wheat, C.W. & Peña, C. (2013). Timing and patterns in the taxonomic diversification of Lepidoptera (butterflies and moths). *PLoS One* 8, e80875.

Wu, L.W., Yen, S.H., Lees, D.C., Lu, C.C., Yang, P.S. & Hsu, Y.F. (2015). Phylogeny and historical biogeography of Asian *Pterourus* Butterflies (Lepidoptera: Papilionidae): A case of intercontinental dispersal from North America to East Asia. *PLoS One* 10, e0140933.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292, 686–693.

# Figures

**Fig. 1.** Phylogenetic hypothesis of *Hyles* based on mitochondrial sequences (1531 bp COI & 682 bp COII) analysed using BEAST. Branches are colour coded according to Bayesian posterior probability (pp) support with a gradient from green (1) to red (<0.5) and pp values (above 0.8) are given to the right of nodes with more than three OTUs. Branches to OTUs of species that are not the focus of this work were collapsed for clarity and the outgroup species (see Table 1) were pruned. Maximal intraspecific distance of 3% within *H. gallii* is found between the two individuals marked with an arrow (#6511 and #6521).

Abbreviations: HT: holotype specimen, PT: Paratype specimen, ST: syntype specimen.

**Fig. 2.** Dated phylogenetic hypothesis of *Hyles* based on combined mitochondrial (1531 bp COI & 682 bp COII) and nuclear (772 bp EF1alpha) sequences of a single individual per species (22 taxon set) analysed using BEAST. Outgroups were omitted from the phylogeny and only the terminal *Hyles* clade is shown. Branch lengths correspond to the time Before Present (B.P.) corresponding to the scale on the axis below. Blue node bars show the 95% highest posterior density confidence intervals (HPD) of the time estimates. Branches are colour-coded according to Bayesian posterior probability (pp) support with a gradient from green (1) to red (<0.8) and pp values are given above nodes. The time intervals of the three calibration points coded for in the analyses are illustrated directly above the scale axis, their node positions are also indicated in the tree.

Abbreviation: CP: Central Palaearctic.

**Fig. 3.** Ancestral biogeography and larval host plant analyses of the twenty-two-taxon tree of *Hyles* as estimated with BEAST, depicting the results of ancestral range (above nodes) and host plant estimations (below nodes). The map and bar show respectively the colour coding of regions (Table 4) and larval host plants (Table S4). Asterisks indicate species with unknown host plant associations.

# Tables

### Table 1 - Specimens studied

List of specimens sequenced in the present study or included from BOLD (http://www.boldsystems.org/; *Hyles* abbreviated to *H.*). Accession numbers of COI+II (for BOLD, barcodes only) sequences are reported with the corresponding taxon and its author, internal number (institutional specimen voucher number), institution in which the tissue voucher is deposited, and collection data. The type status of samples is indicated by HT, AT, ST and PT, referring to tissue taken from the holotype, allotype, syntype and paratype respectively.

Abbreviations:

Acc. No.: GenBank accession number

Int. No.: Internal number (institutional specimen voucher number)

BIO: Biodiversity Institute of Ontario, Canada

EMEM: Entomologisches Museum Eitschberger, Marktleuthen, Germany

DAM: Alberta Museum, Edmonton, Canada

Haxaire: Research Collection of Jean Haxaire, France

IZCAS: Institute of Zoology, Chinese Academy of Sciences, China

LSNK: Landessammlung für Naturkunde, Karlsruhe, Germany

MTD: Museum für Tierkunde, Senckenberg Natural History Collections Dresden, Germany

MWM: Museum Witt München, Germany

NHMUK: Natural History Museum, London, U.K.

NHMV: Natural History Museum Vienna, Austria

RCRB: Research Collection of Ron Brechlin, Pasewalk, Germany

RBCM: Royal British Columbia Museum, Victoria, Canada

TLMF: Tiroler Landesmuseum Ferdinandeum, Innsbruck, Austria

UASM: University of Alberta, E.H. Strickland Entomological Museum, Edmonton, Canada

UHIM: University of Hawaii Insect Museum, Hawaii, U.S.A.

UM: Department of Entomology, University of Maryland, U.S.A.

UO: University of Oulu, Finland

ZFMK: Zoologisches Forschungsinstitut und Museum Alexander Koenig, Bonn, Germany

ZSM: Zoologische Staatssammlung München, Germany

† Sample was treated with aDNA methodology

* Mitochondrial marker of *Hyles livornica*

# Mitochondrial marker of *Hyles renneri*

**Table 2 - Model selection with BioGeoBEARS**

Results of model selection with BioGeoBEARS.

| Model | LnL | No. of Parameters | d | e | j | AICc | AIC weight |
|---|---|---|---|---|---|---|---|
| BAYAREALIKE | -85.9 | 2 | 0.015 | 0.29 | 0 | 176.4 | 0.77 |
| BAYAREALIKE+J | -85.76 | 3 | 0.014 | 0.25 | 0.0069 | 178.9 | 0.23 |
| DEC+J | -93.2 | 3 | 0.032 | 1.0e-12 | 0.032 | 193.7 | 0.0001 |
| DEC | -97.92 | 2 | 0.038 | 0.021 | 0 | 200.5 | 4.6e-06 |
| DIVALIKE+J | -97.43 | 3 | 0.037 | 1.0e-12 | 0.029 | 202.2 | 2.0e-06 |
| DIVALIKE | -100.5 | 2 | 0.043 | 2.0e-08 | 0 | 205.7 | 3.4e-07 |

**Table 3 - Primers**

Primers used to amplify the three mitochondrial marker genes, the first two subunits of cytochrome-c-oxidase (COI/II), and t-RNA-Leucine, in multiple short fragments using aDNA methodology. Fragments 1a, 1b, 2a, 2b, 3a & 3b yield 300-600 bp, A–M only 110–280 bp. Numbers in primer names refer to the approximate positions in the COI+t-RNA-Leu+COII alignment of Hundsdoerfer et al. (2005). Fragm. = fragment, CDS = coding sequence, Tm = annealing temperature, Pr. = primer.

| Fragm. code | Fragm. target CDS | Fragm. length | Tm | Pr. No. | Primer Name | Primer Sequence | Reference |
|---|---|---|---|---|---|---|---|
| 1a | COI | 300bp | 47° | 320 | LepF1 | ATTCAACCAATCATAAAGATATTGG | Hebert et. al. 2004 |
| | | | | 323 | MLepR1 | CCTGTTCCAGCTCCATTTTC | Hajibabaei et al. 2006 |
| 1b | COI | 400bp | 47° | 322 | MLepF1 | GCTTTCCCACGAATAAATAATA | Hajibabaei |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | et al. 2006 |
| | | | | 321 | LepR1 | TAAACTTCTGGATGTCCAAAAAATCA | Hebert et. al. 2004 |
| A | COI | 110bp | 47° | 320 | LepF1 | ATTCAACCAATCATAAAGATATTGG | Hebert et. al. 2004 |
| | | | | 352 | HylesCOIca130r | RTCATCTCCRATTAARGATC | This study |
| B | COI | 270bp | 47° | 351 | HylesCOIca100f | TAAGATTAYTAATTCGAGCAG | Mende et al. 2013 |
| | | | | 323 | MLepR1 | CCTGTTCCAGCTCCATTTTC | Hajibabaei et al. 2006 |
| C | COI | 240bp | 55° | 322 | LepF1 | ATTCAACCAATCATAAAGATATTGG | Hebert et. al. 2004a |
| | | | | 354 | HylesCOIca550r | AATGCTGTRATTYCIACAGCTC | This study |
| D | COI | 270bp | 47° | 353 | HylesCOIca480f | AGGRGCTGTAAATTTYTTAC | This study |
| | | | | 321 | LepR1 | TAAACTTCTGGATGTCCAAAAAATCA | Hebert et. al. 2004a |
| 2a | COI | 400bp | 47° | 345 | HylesMCOIf | CATCMTTTTTTSAYCCTGCTG | This study |
| | | | | 365 | HylesCOIca1070r | TGWARWARWRTAATRTCRATWGATG | This study |
| 2b | COI | 330bp | 49° | 366 | HylesCOIca1020f | CYTCYATTYTATGRAGATTAGG | This study |
| | | | | 370 | HylesCOIca1350r | GARATATANGAYCCYARTGATG | This study |
| E | COI | 200bp | 47° | 345 | HylesMCOIf | CATCMTTTTTTSAYCCTGCTG | This study |
| | | | | 356 | HylesCOIca880r | YGTATCRATRTCTATWCCTAC | This study |
| F | COI | 250bp | 47° | 355 | HylesCOIca830f | YTAYGCTATAATAGCAATTGG | This study |
| | | | | 365 | HylesCOIca1070r | TGWARWARWRTAATRTCRATWGATG | This study |
| G | COI | 150bp | 47° | 366 | HylesCOIca1020f | CYTCYATTYTATGRAGATTAGG | This study |
| | | | | 367 | HylesCOIca1170f | ATARWGGRTAYCARTGAATRAATCC | This study |
| H | COI | 180bp | 47° | 368 | HylesCOIca1110f | AYGAYACWTAYTAYGTTGTAGC | This study |
| | | | | 370 | HylesCOIca1350r | GARATATANGAYCCYARTGATG | This study |
| 3a | COI/II | 450bp | 47° | 369 | HylesCOIca1305f | GNTTRGCNATACCNCGACG | This study |
| | | | | 371 | HylesCOIIca1760r | TYTGNYCTTCTARTARRAATCG | This study |
| 3b | COII | 580bp | 47° | 359 | HylesCOIIca1710f | AATTATRATYACTATYTTAGTAGG | This study |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | 348 | HylesCOIIr | TTCAGTCATCTAATGAAG | This study |
| I | COI | 280bp | 47° | 369 | HylesCOIca1305f | GNTTRGCNATACCNCGACG | This study |
| | | | | 343 | HylesCOIr | TAAATTAGATCAAGTTGCCATTTC | This study |
| J | COII | 250bp | 47° | 347 | HylesCOIIf | YRCYGAACATTCATATAATG | This study |
| | | | | 371 | HylesCOIIca1760r | TYTGNYCTTCTARTARRAATCG | This study |
| K | COII | 200bp | 47° | 359 | HylesCOIIca1710f | AATTATRATYACTATYTTAGTAGG | This study |
| | | | | 349 | HylesMCOIIr | GAATATTCATATCTTCARTATC | This study |
| L | COII | 180bp | 47° | 350 | HylesMCOIIf | GATAYTGAAGATATGAATATTC | Mende et al. 2013 |
| | | | | 372 | HylesCOIIca2125r | TTGTTTGRTTTAAACGTCCAGG | Mende et al. 2013 |
| M | COII | 150bp | 47° | 358 | HylesCOIIca2130f | AYTAGGRGTAAAAGTAGATGC | This study |
| | | | | 348 | HylesCOIIr | TTCAGTCATCTAATGAAG | This study |

**Table 4 - Ancestral Biogeography analysis**

Species occurrences used for the BioGeoBears analysis.

Region codes as in Fig. 3: A, Nearctic; B, Neotropics; C, Hawaii; D, Madagascar; E, Australia; F, Africa (south of Sahara); G, Western Palaearctic (incl. N Africa); H, Central Palaearctic (excluding Himalaya); I, Eastern Palaearctic; J, Himalaya.

| **OTU** | **Occurrence codes** |
|---|---|
| *Hyles lineata* | A, B, C |
| *Hyles perkinsi* | C |
| *Hyles calida* | C |
| *Hyles biguttata* | D |
| *Hyles euphorbiarum* | B |

| | |
|---|---|
| *Hyles annei* | B |
| *Hyles livornicoides* | E |
| *Hyles livornica* clade | D, F, G, H, I, J |
| *Hyles renneri* clade | J |
| *Hyles vespertilio* | G |
| *Hyles gallii* clade | A, G, H, I, J |
| *Hyles nicaea* | G, H, J |
| *Hyles hippophaes* | G, H, I |
| *Hyles salangensis* | H |
| *Hyles stroehlei* | H |
| *Hyles centralasiae* | H, I |
| *Hyles siehei* | G, H |
| *Hyles costata* | I |
| *Hyles zygophylli* | G, H, I |
| *Hyles dahlii* | G |
| *Hyles euphorbiae* | G, H, I |
| *Hyles tithymali* | G |

# Supplement

**Figure S1 – "Tree cloud" representation of the phylogeny of *Hyles***

All 10000 trees in the BEAST tree sample were plotted (blue) with DENSITREE to illustrate the degree of phylogenetic uncertainty among them. The maximum clade credibility tree is superimposed in red with Bayesian posterior probabilities indicated at the nodes.

**Figure S2 - Reconstruction of the ancestral state of the arolium**

Probabilities of ancestral states obtained from reconstruction with BEAST are indicated at each node. A vestigial arolium is reconstructed as the plesiomorphic condition in *Hyles*; the arolium was then lost twice independently. It is fully developed only the two species from Hawaii.

**Table S1**

The best partitioning scheme and evolutionary models determined with PartitionFinder.

| Subset | Best Model | Subset Partitions | Subset Sites |
|---|---|---|---|
| 1 | TrN+I+G | COIposition1, COIIposition1 | 1-1531\3, 1532-2212\3 |
| 2 | HKY | COIposition2 | 2-1531\3 |
| 3 | GTR+I+G | COIposition3, COIIposition3 | 3-1531\3, 1534-2212\3 |
| 4 | HKY | COIIposition2 | 1533-2212\3 |
| 5 | HKY+I | Ef1aposition1 | 2213-2985\3 |
| 6 | HKY+I | Ef1aposition2 | 2214-2985\3 |
| 7 | HKY+G | Ef1aposition3 | 2215-2985\3 |

**Table S2**

Input matrix for ancestral range estimation with BioGeoBEARS.

| Taxon | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| *H. annei* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *H. biguttata* | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| *H. calida* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *H. centralasiae* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
| *H. costata* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| *H. dahlii* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *H. euphorbiae* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| *H. euphorbiarum* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *H. gallii* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| *H. hippophaes* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| *H. lineata* | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *H. livornica* | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| *H. livornicoides* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| *H. nicaea* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| *H. perkinsi* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| *H. renneri* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| *H. salangensis* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *H. siehei* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| *H. stroehlei* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| *H. tithymali* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *H. vespertilio* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| *H. zygophylli* | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |

## Table S3

Overview of sequences used for Fig. 2. Unless indicated, the first (or the first two) number(s) is (are) for COI & COII and the last number for EF1alpha. H.: *Hyles*, HT: Holotype.

| **Accession Numbers** | **Species** | **Internal Number** |
|---|---|---|
| AJ749414, FN392890 | *Basiothia medea* | 23272 |
| FN386541, FN392891 | *Chaerocina dohertyi* | 23253 |
| AJ749419, FN392901 | *Theretra alecto* | 23266 |
| FN386543, FN392899 | *Theretra japonica* | 23269 |
| AJ749418, FN392900 | *Theretra silhetensis* | 23260 |
| AJ749415, AJ749416, FN392902 | *Xylophanes chiron* | 23250 |
| FN386544, FN392903 | *Xylophanes falco* | 23256 |
| AJ749417, FN392904 | *Xylophanes porcus* | 23271 |
| AJ749420, FN392894 | *Deilephila elpenor* | 695903 |
| FN386542, FN392892 | *Deilephila porcellus* | 23265 |
| AJ749423, FN392896 | *Hippotion celerio* | 695972 |
| AJ749425, FN392898 | *Hippotion echeclus* | 23261 |
| AJ749421, FN392895 | *Hippotion eson* | 695958 |
| AJ749429, FN392908 | *H. annei* | 16162 |
| FN386546, FN392909 | *H. biguttata* | 507 |
| AJ749426, FN392912 | *H. calida* | 0105 |
| FN386551 (COI & COII) | *H. centralasiae* | 1323 |
| FN392917 (EF1alpha) | *H. centralasiae* | 3915 |
| AJ749470, AJ749471, FN392918 | *H. costata* | 23245 |
| AJ749455, FN392921 | *H. dahlii* | 0030 |
| AJ749469, FN392925 | *H. euphorbiae* | 23280 |
| AJ749428 (COI & COII) | *H. euphorbiarum* | 23274 |
| FN392935 (EF1alpha) | *H. euphorbiarum* | 23273 |

| | | |
|---|---|---|
| AJ749434 (COI & COII) | *H. gallii* | 4233 / 695850 |
| FN392939 (EF1alpha) | *H. gallii* | 489 |
| AJ749452 (COI & COII) | *H. hippophaes* | 695817 |
| FN392942 (EF1alpha) | *H. hippophaes* | 1317 |
| AJ749436 (COI & COII) | *H. lineata* | 23259 |
| FN392905 (EF1alpha) | *H. lineata* | 23258 |
| FN386572 (COI & COII) | *H. livornica* | 2537 |
| FN392947 (EF1alpha) | *H. livornica* | 950 |
| AJ749442 (COI & COII) | *H. livornicoides* | 23297 |
| FN392952 (EF1alpha) | *H. livornicoides* | 23296 |
| FN386582, FN392956 | *H. nicaea* | 625 |
| FN386583, FN392958 | *H. perkinsi* | HW19hi |
| #########, - | *H. renneri* HT | 1428 |
| #########, - | *H. salangensis* HT | 1414 |
| AJ749453 (COI & COII) | *H. siehei* | 16137 |
| FN392971 (EF1alpha) | *H. siehei* | 862 |
| FN386594 (COI & COII) | *H. stroehlei* | 584 |
| FN392973 (EF1alpha) | *H. stroehlei* | 588 |
| AJ749486, FN392985 | *H. tithymali* | 084a |
| AJ749445, FN392991 | *H. vespertilio* | 0007 |
| FN386608, FN392995 | *H. zygophylli* | 3445 |

**Table S4**

Character matrix used for reconstruction of ancestral host plant use and the ancestral state of the arolium.

| Taxon | Host plant | Arolium |
|---|---|---|
| *H. annei* | polyphagous | vestigial |
| *H. biguttata* | Ericaceae | vestigial |
| *H. calida* | Rubiaceae | fully developed |
| *H. centralasiae* | Asphodelaceae | absent |
| *H. costata* | Polygonaceae | absent |
| *H. dahlii* | Euphorbiaceae | absent |
| *H. euphorbiae* | Euphorbiaceae | absent |
| *H. euphorbiarum* | polyphagous | vestigial |
| *H. gallii* | polyphagous | vestigial |
| *H. hippophaes* | Elaeagnaceae | absent |
| *H. lineata* | polyphagous | vestigial |
| *H. livornica* | polyphagous | vestigial |
| *H. livornicoides* | Nyctaginaceae-Zygophyllaceae | vestigial |
| *H. nicaea* | Euphorbiaceae | vestigial |
| *H. perkinsi* | Rubiaceae | fully developed |
| *H. renneri* | ? | vestigial |
| *H. salangensis* | ? | absent |
| *H. siehei* | Asphodelaceae | absent |
| *H. stroehlei* | ? | absent |
| *H. tithymali* | Euphorbiaceae | absent |
| *H. vespertilio* | Onagraceae | absent |
| *H. zygophylli* | Zygophyllaceae | absent |