

Preliminary molecular phylogeny and biogeography of the monobasic subfamily Calinaginae (Lepidoptera, Nymphalidae)

Valentina Todisco¹, Vazrick Nazari², Donatella Cesaroni³, Valerio Sbordoni³

¹ University of Vienna, Department of Botany and Biodiversity, Vienna, Austria

² Agriculture and Agri-Food Canada, Eastern Cereal and Oilseed Research Centre, C.E.F., Ottawa, Ontario K1A 0C6, Canada

³ Department of Biology, University of Rome 'Tor Vergata', Via della Ricerca Scientifica, 00133 Rome, Italy

<http://zoobank.org/3AEA2B93-1B80-42A8-BAB6-7EEF0F130DFC>

Corresponding author: *Valentina Todisco* (valentina.todisco@gmail.com)

Abstract

Received 7 October 2016

Accepted 29 March 2017

Published 12 April 2017

Academic editor:

Michael Ohl

Key Words

Calinaga

Calinaginae

Nymphalidae

mtDNA

butterfly

Indochina

Oligocene

Calinaga (Moore 1857) is a rare and enigmatic Asian butterfly genus whose phylogenetic placement within Nymphalidae has only recently been established. The evolutionary history of *Calinaga* species however remains unknown. Here we explore the phylogeography of *Calinaga* using 1310 bp of sequence data from two molecular (mtDNA barcode and ribosomal protein S5 nuclear gene) and two morphological traits (genitalia and wing pattern). Within the proposed phylogenetic framework, we estimate the ages of divergence within the genus and reconstruct their historical biogeography. We found strong support for monophyly of *Calinaga* and support for the most recent accepted species in the genus. Our results indicate that the common ancestor of *Calinaga* first split in the Eocene (~43 million years ago) in southern China, probably as a consequence of geological and environmental impacts of the collision of the Indian and Asian subcontinents. In the Oligocene/Miocene, the extrusion of Indochina from the continent caused further dramatic orogenetic changes that promoted isolation and speciation events within the genus while Pleistocene climatic changes also influenced the distribution and further speciation. A dispersal–vicariance analysis suggests that vicariance events have played a far more important role than dispersal in the distribution of extant species.

Introduction

Calinaga (Moore, 1857) is an enigmatic butterfly genus mainly distributed in the Indochina region, but also extending to Taiwan and the Himalayas (Fig. 1A). It is a prime example of a pattern that is found in diverse Himalayan-Indochinese genera of butterflies (e.g. *Bhutanitis*, *Byasa*, *Callerebia*, *Loxerebia*, *Hemadara*, *Neope*, *Chrysozephyrus*) and in many other groups including vertebrates (e.g. birds - *Phylloscopus*: Martens et al. 2011; frogs - tribe Painsi: Che et al. 2010; plants - Roscoea: Zhao et al. 2015). Species of *Calinaga* inhabit forested areas at altitudes between about 1000–2000m where their larvae feed on various species of *Morus* (Lee 1958, Ashizawa and Muroya 1967, Chou 1994, Zhang and Hu 2012).

The systematics of the genus has long been a topic for taxonomic discussion. Moore (1857) described the genus

under “Strip III., with Chilopodiform or Scolopendriiform Larvae” alongside other nymphalids, with *C. buddha* as its type species. Kirby (1871) assigned *Calinaga* species to the Papilionidae (genus *Parnassius*), but he later (1877) placed the genus in the Nymphalidae (Hypolimnina group). Moore (1895) included *Calinaga* in the monobasic subfamily Calinaginae that is currently comprised of only this genus. Later Fruhstorfer (in Seitz 1927) recognized the genus under Nymphalidae based on certain morphological characters (genitalia, antennae, wing patterns), while Clark (1947) placed it in the “family” Danaidae and later gave it a separate family status, Calinagidae, between the “families” Argynnidae and Danaidae (Clark 1948). Ehrlich (1958a) recognized it as a distinct monotypic subfamily within the Nymphalidae, and, on the basis of characters of the adult integumental anatomy, suggested a close relationship to the subfamilies Satyrinae and Morph-

inae. This affinity was subsequently supported by studies on larval stages (Lee 1958) and adult morphology (Freitas and Brown 2004), as well as by recent molecular and phylogenetic studies that have firmly placed the subfamily as the sister to Satyrinae + Charaxinae (Wahlberg and Wheat 2008, Peña and Wahlberg 2008, Wahlberg et al. 2009, Heikkilä et al. 2012). At species level, Ehrlich (1958b) suggested that *Calinaga* was monotypic, comprised of a single species, *C. buddha* (Moore, 1857). The main problem of the traditional taxonomic species-level framework for *Calinaga*, from Fruhstofer (in Seitz 1927) to Ehrlich (1958b), has been the practice of placing the vast majority of available names as subspecies or synonymous under *C. buddha*. Presently, taxonomists consider the genus composed from 7 to 11 species (D'Abbrera 1985, 1993, Beccaloni et al. 2003, Lang 2012, Savela 2015, Sondhi et al. 2016, Vane-Wright, personal communication). A molecular study to shed light on systematic and evolutionary history of these species has never been attempted.

Located in the continental portion of Southeast Asia, lying to the southeast of the Himalayas and south of China, the Indochina bioregion has long been recognized as an area with globally important levels of biodiversity (Myers et al. 2000). Recent meta-analyses of geological, climatic, and biological datasets have demonstrated impressively that it is one of two major evolutionary hotspots for Southeast Asian biodiversity for a diverse range of plants and vertebrates (de Bruyn et al. 2014). It is a geologically and topographically complex region, with intricate current and historical climatic patterns that contribute to its rich biotic diversity (e.g. Fontaine and Workman 1997, Hall 1998, 2001, An 2000, Morley 2000, Sterling et al. 2006). Although several studies have examined the processes that have led to present patterns of species richness in plants and vertebrates (Bain and Hurley 2011, de Bruyn et al. 2014), little is known about the underlying causes for this remarkably high diversity, mechanisms of speciation, or origin of the Indochinese insect fauna. Until recently, some contributions addressing these topics were focused on species groups with high dispersal ability, such as butterflies (Monastyrskii and Holloway 2013).

For the first time in this study, we conduct an analysis of molecular data (mitochondrial COI-5' and ribosomal protein S5 nuclear gene (RpS5)) and we examine morphological characters (genitalia and wing pattern) for 51 specimens of *Calinaga* from widely distributed sites in the Indochina region (Fig. 1A, Table 1), for inferring preliminary information on the molecular phylogeny and the biogeographic processes which determined present-day distribution of the genus. An upcoming study by the first author (VT) will investigate the type material held at NHM London, with the aim of adding new information to the existing dataset.

Given the scarcity of information on the evolutionary history of *Calinaga*, the geographic and taxonomic coverage in our datasets will help to advance understanding of the diversification of this charismatic insect group in the Indochina bioregion and of other genera with similar distribution patterns.

Materials and Methods

We examined 51 *Calinaga* individuals from 29 localities across India, South China (Yunnan, Shaanxi, Qinghai, Sichuan), Laos, Vietnam, Myanmar and Thailand (Fig. 1A, Table 1). Photographs of all specimens used in the study as well as detailed collecting data are available in the Barcode of Life Data System (BOLD) (Ratnasingham and Hebert 2007) at <https://doi.org/10.5883/DS-CALI>. The specimens were initially identified as including representatives of six species: *C. aborica* (N= 2), *C. buddha* (N= 25), *C. buphonas* (N= 2), *C. davidis* (N= 6), *C. lhatso* (N= 2) and *C. sudassana* (N= 14). Samples were submitted for sequencing using the original identifications, but they were re-identified following collection of genetic and morphological data (see paragraph 2.1). We were unable to obtain sequences from, among others, the following taxa: *C. buddha buddha* (Northern India), *C. gautama* (Sikkim), *C. davidis davidis*, *C. fokienensis*, *C. kunagtungensis* and *C. pacifica* (China), *C. formosana* (Taiwan), and *C. distans* (Southern Vietnam).

Morphology

Morphological identifications in this study were made based on illustrations of type material in Moore (1857), Melvill (1893), Leech (1892), de Niceville (1897), Seitz (1909, 1927), Tytler (1915), Oberthür (1920), Lang (2012) and Sondhi et al. (2016). Male genitalia were dissected, prepared and photographed using traditional techniques (Supplementary Material S1: Sbordoni V. personal communication).

Amplification and sequencing analysis

DNA was extracted from one leg of each individual on Biomek FX liquid handling robot using a semi-automated DNA extraction protocol (Ivanova et al. 2006) on glass fiber plates (PALL Acroprep 96 with 3µm GF membrane over 0.2µm Bioinert membrane). DNA was eluted in 35-40µl of ddH₂O pre-warmed to 56°C and stored at -80°C. Polymerase chain reactions (PCR) and DNA sequencing were carried out following standard procedures for Lepidoptera (Hajibabaei et al. 2005, deWaard et al. 2008). A 658-bp fragment of cytochrome *c* oxidase subunit I (COI) was targeted for amplification using the primers LepF (5' ATTCAACCAATCATAAAGATATTGG3') and LepR (5' TAAACTTCTGGATGTCCAAAAAATCA3').

The entire nuclear RpS5 was initially amplified using the primers RpS5f/r (Wahlberg and Wheat 2008) and afterwards a fragment covering 483 bp was obtained by nested PCR using RpS5f and a novel specific primer (5' AC-MGTDCGYCGTCARGCTGTRGAYGTBTCWCC3'), developed from conserved regions. The PCR conditions were carried out as in Todisco et al. (2010).

Sequences were obtained by using an ABI 3730xl sequencer following the manufacturer's recommendations and they were edited and assembled using CodonCode Aligner 6.0.2. All sequences of *Calinaga* were submitted to GenBank (Accession Numbers in Table 1).

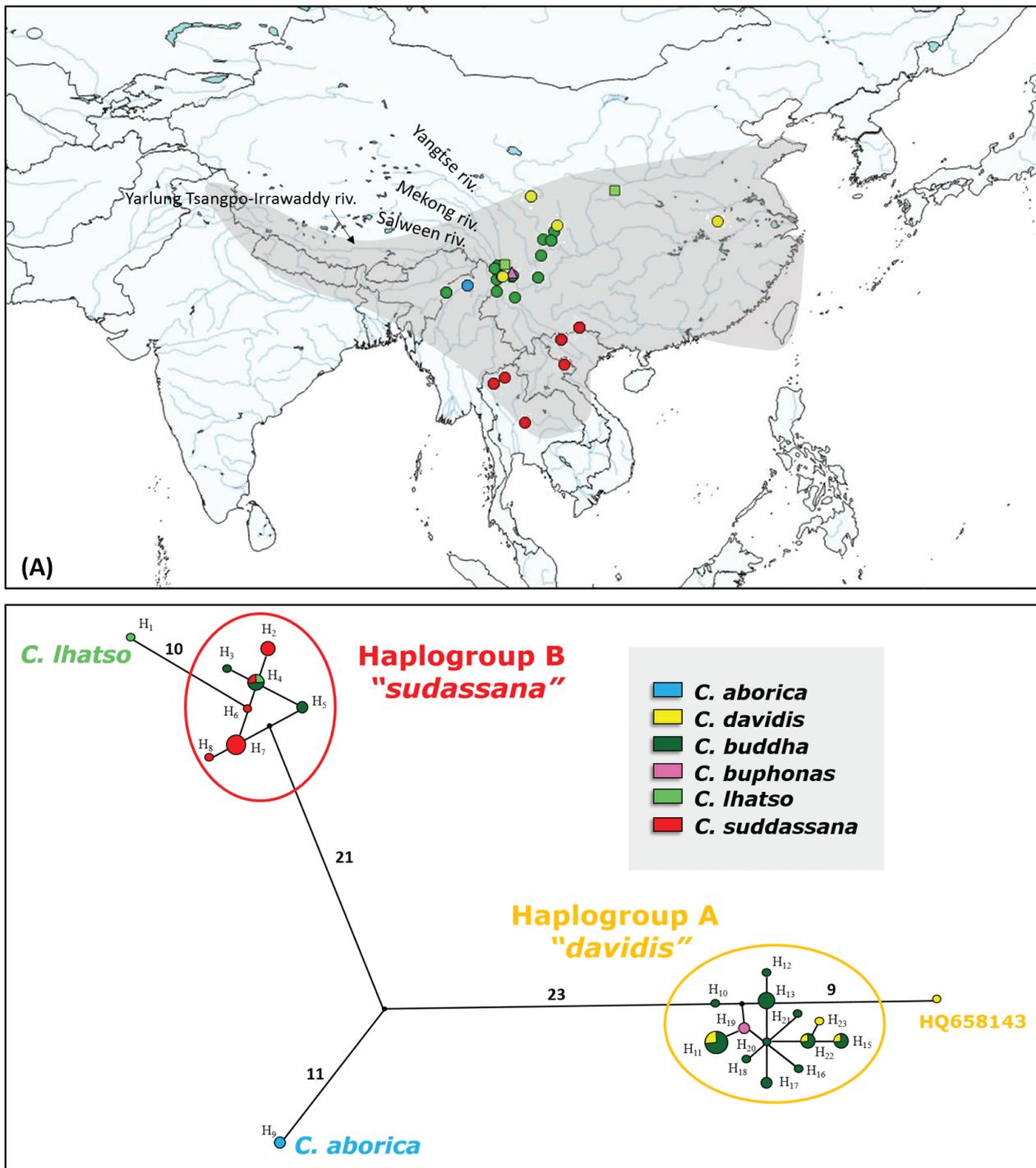


Figure 1. (A) Approximate geographic distributions (Shirôzu 1960, Lang 2012) and sampling localities (circles) for the species of *Calinaga* included in this study (with the exception of the sample CBUD-INDIN for which we do not have an exact locality). Species as initially identified are highlighted and shown in different colours. Note that many of these initially attributed names subsequently proved erroneous. The map was obtained using Quantum GIS 2.8.2 based on a map from Natural Earth (www.naturalearthdata.com). (B) Median-Joining Network of mtDNA. Circle size proportional to haplotype frequency; number of nucleotide substitutions indicated along connections, except for single or double substitutions. In both figures the species are highlighted and shown in different colours as initially identified.

Phylogenetic analyses and fossil calibration

Two sequences from *C. buddha* (COI: EF683684, AY090208; RpS5: EU141406) and one from *C. davidis* (COI: HQ658143), as well as COI and RpS5 sequences

for eight outgroups for species belonging to closely related lineages of Nymphalidae (Charaxinae and Satyrinae) from Wahlberg and Wheat (2008) were retrieved from GenBank and added to our dataset (see Table 1).

Table 1. Taxon= *Calinaga* species identified in this study. Sample numbers (N), collection locality, population code, mtDNA haplotype code and GenBank Accession numbers.

| taxon | N | Collection locality | Population code | Haplotype code | GenBank Accession | |
|--------------------------------|---|---|-----------------|-----------------------------------|----------------------------|-----------------|
| | | | | | COI KX233- | RpS5 KX283- |
| <i>aborica</i> | 2 | Burma: N Sagaing, Tailing Hka river | CABO-BURSA | H ₉ | 590/612 | 390/398 |
| <i>lhatso</i> | 1 | China: Shaanxi, Tsnling Mts, Foping nature reserve, 1600 m | CLHA-CHSTF | H ₄ | 570 | 379 |
| <i>davidis</i> | 2 | China: Qinghai, Ningyou Jiuzhi country | CDAV-CHNJQ | H ₁₅ , H ₂₃ | 576/581 | 382/384 |
| <i>davidis nubilosa</i> | 2 | China: Qinghai, Ningyou Jiuzhi country | CBUD-CHQNJ | H ₁₃ | 577/593 | 383/391 |
| <i>davidis</i> | 1 | China: Sichuan, Yingxiuwan, 1000 m | CDAV-CHSYI | H ₂₂ | 596 | 394 |
| <i>davidis nubilosa</i> | 1 | China: W Sichuan, Dayi district, 23Km north of Huashuiwan, 1500m | CBUD-CHSDH | H ₂₂ | 603 | - |
| <i>davidis nubilosa</i> | 1 | China: W Sichuan, Shimian-Mianning road, 2300 m | CBUD-CHSSM | H ₂₀ | 599 | - |
| <i>davidis nubilosa</i> | 1 | China: W Sichuan, 30km W of Tianquan, 1250 m | CBUD-CHSTI | H ₁₅ | 580 | - |
| <i>davidis nubilosa</i> | 3 | China: W Sichuan, Ya'an-ergaz Shan, 1400 m | CBUD-CHSYS | H ₁₂ , H ₁₃ | 601/614/609 | - |
| <i>davidis</i> | 2 | China: S Sichuan, SW of Dechang Miyi, 2100 m | CBUD-CHSDM | H ₁₈ , H ₂₁ | 572 615 | - |
| <i>davidis buphonas</i> | 3 | China: Yunnan, Wei-Xi 10km SE, 2000-2400 m | CDAV-CHYWE | H ₁₁ | 613/584/616 | 387/399 |
| <i>davidis buphonas</i> | 1 | China: Yunnan, Surroundings of Dali city, 2400 m | CBUD-CHYDA | H ₁₁ | 594 | 392 |
| <i>davidis</i> | 2 | China: Yunnan, Ningjing Shan Tse-Kou Mekong River, 2100 m | CBUD-CHYNM | H ₁₁ | 619/605 | 401/395 |
| <i>davidis genestieri</i> | 1 | China: Yunnan, Ningjing Shan Wei-Shi, 2200 m | CBUD-CHYNS | H ₁₁ | 583 | 386 |
| <i>davidis genestieri</i> | 1 | China: Yunnan, Wei-Xi, 2200 m | CBUD-CHYWX | H ₁₀ | 618 | - |
| <i>davidis</i> | 1 | China: Yunnan, 40km N of Xhinxrong, 1850 m | CBUD-CHYXH | H ₂₂ | 586 | - |
| <i>sudassana</i> | 2 | China: W Yunnan, Salween valley, 6 Km of Fugong, 1300 m | CBUD-CHYSA | H ₅ | 611/587 | 397 |
| <i>davidis</i> | 2 | China: W Yunnan, Salween valley, 16Km N of Gongshan, 1675 m | CBUD-CHYSG | H ₁₇ | 617/620 | 400/402 |
| <i>davidis cercyon</i> | 2 | China: N Yunnan, Sanba: Blotopel (Zhongdian), 2400 m | CBUP-CHYSB | H ₁₉ | 598/591 | - |
| <i>davidis buphonas</i> | 2 | China: N Yunnan, 75Km NW of Lijang, 2000 m | CBUD-CHYLI | H ₁₁ | 578/571 | - |
| <i>lhatso</i> | 1 | China: N Yunnan, 55km N of Zhongdian, 2450 m | CLHA-CHYZH | H ₁ | 592 | - |
| <i>sudassana</i> | 1 | NE India | CBUD-INDIN | H ₄ | 579 | - |
| <i>sudassana</i> | 2 | India: Nagaland Naga hills | CBUD-INDNNH | H ₃ , H ₄ | 608/597 | - |
| <i>sudassana</i> | 1 | Laos: Samneua | CSUD-LAOSAM | H ₄ | 607 | - |
| <i>sudassana</i> | 1 | Thailand | CSUD-TAITAI | H ₆ | 602 | - |
| <i>sudassana</i> | 1 | Thailand: Wiang Papao, Chiang Rai | CSUD-TAIRAI | H ₂ | 574 | - |
| <i>sudassana</i> | 2 | N Thailand: Samoeng | CSUD-TAISAM | H ₂ | 604/610 | - |
| <i>sudassana</i> | 3 | Vietnam: Dong Van, Meo Wac district, Ha Glang, 1700 m | CSUD-VNDVM | H ₇ | 589/573/575 | 389/380/381 |
| <i>sudassana</i> | 6 | Vietnam: Hoang Lien national Park Sapa, Cat Cat village and Muong Hoa river, 1350 m | CSUD-VNHLK | H ₇ , H ₈ | 595/606/588 600/585/582 | 393/396/388/385 |
| <i>davidis</i> | 1 | China | - | H ₁₅ | EF683684 | - |
| <i>davidis</i> | 1 | Stratford Butterfly Farm, UK | NW64-3 | H ₁₈ | AY090208 | EU141406 |
| <i>lactoris</i> | 1 | China: Tianma National Nature Reserve in Jinzhai County, Anhui | - | H ₁₄ | HQ658143 | - |
| <i>Euxanthe wakefieldi</i> | 1 | - | - | - | EU141357 | EU141390 |
| <i>Caligo telamonius</i> | 1 | - | - | - | AY090209 | EU141414 |
| <i>Anaea troglodyta</i> | 1 | - | - | - | DQ338573 | EU141428 |
| <i>Archaeoprepona demophon</i> | 1 | - | - | - | AY090220 | EU141424 |
| <i>Bicyclus anynana</i> | 1 | - | - | - | AY218238 | EU141374 |
| <i>Melanitis leda</i> | 1 | - | - | - | AY090207 | EU141408 |
| <i>Morpho peleides</i> | 1 | - | - | - | AY090210 | EU141407 |
| <i>Charaxes castor</i> | 1 | - | - | - | AY090219 | EU141422 |

We used NETWORK 5 (Bandelt et al. 1999) to calculate a Median Joining (MJ) network representing the genealogical relationships among mtDNA haplotypes (Fig. 1B). Genetic distances were calculated in MEGA 6 under the Kimura 2-parameter model of base substitution (Tamura et al. 2013). The program JModelTest 2.1.7 (Guindon and Gascuel 2003, Darriba et al. 2012) was employed to evaluate models of nucleotide substitutions. Only the best-fit models were subsequently used for estimations of tree topology.

We used the BEAST package 1.8.3 (Drummond et al. 2012) for Bayesian phylogeny reconstruction and divergence time estimation for concatenated genes. The analysis was allowed to run for 50 million generations and the first 5000 trees were discarded as burnin. A consensus tree was calculated with posterior probability limit of 0.5, maximum clade credibility tree and median heights. Monophyly was enforced on a) Satyrinae, b) Charaxinae, and c) Satyrinae + Charaxinae (Fig. 2). The tree prior was set to Yule Process Speciation model with a random starting tree. The analysis was performed using an uncorrelated relaxed clock (lognormal relaxed distribution not selected) for both genes. Trees were visualized using FIGTREE 1.4 (Rambaut 2009).

Recent phylogenetic studies calibrated with fossil data (Peña and Wahlberg 2008; Wahlberg et al. 2009) suggest that Nymphalidae began to diversify in the late Cretaceous around 90 Mya, satyrines (Calinaginae + Satyrinae + Charaxinae) around 75 Mya and Satyrinae + Charaxinae about 70 Mya. Thus, in our BEAST analyses the tree root was calibrated with normal distribution at 75 Mya (stdev 3), Satyrinae + Charaxinae at 70 Mya (stdev 3.5) and *Charaxes* + *Euxanthe* at 22 Mya (stdev 1) (see Fig. 2).

Tests of demographic equilibrium and population expansion in mtDNA haplogroup

Haplotype and nucleotide diversity were calculated using DnaSP 5.0 (Librado and Rozas 2009). Demographic equilibrium was tested in each mtDNA haplogroup by calculating F_s (Fu 1997) and R_2 (Ramos-Onsins and Rozas 2002) statistics, which have been shown to provide sensitive signals of population expansion (Ramos-Onsins and Rozas 2002). Arlequin 3.0 (Excoffier et al. 2005) and DnaSP 5.0 were used to compute F_s and R_2 , respectively, and to test their statistical significance by simulating random samples (10,000 replicates) under the null hypothesis of selective neutrality and constant population size using coalescent algorithms (both modified from Hudson 1990). P-values for the two statistics were obtained as the proportion of simulated values smaller than or equal to the observed values ($\alpha = 0.05$). Expected mismatch distributions and parameters of sudden expansion $\tau = 2 \mu T$ were calculated using Arlequin 3.0 by a generalized least-squares approach (Schneider and Excoffier 1999), under models of pure demographic expansion and spatial expansion (Ray et al. 2003, Excoffier 2004). The probability of the data under the given model was assessed by the goodness-of-fit test implemented in Arlequin 3.0.

Parameter confidence limits were calculated in Arlequin 3.0 through a parametric bootstrap (1000 simulated random samples).

Ancestral Area Reconstruction

Dispersal-vicariance optimization of ancestral areas implemented in DIVA 1.1 (Ronquist 1996) was used to infer the biogeographic history of the group. The distributions of populations were divided into five biogeographic units designated following Che et al. (2010) and Procheş and Ramdhani (2012), as follows: A) Southwestern China ecozone, B) Himalaya-Tibetan plateau region, C) Northern Sino-Himalaya, D) Southern Sino-Himalaya, E) Indochina (Fig. 2). The least number of dispersal events (12) was achieved when maxareas was set to 3.

Results

Morphology

The morphological study on the wing pattern of our specimens, based on illustrations in the most relevant literature, suggested the presence of four species: *C. aborica*, *C. davidis*, *C. lhatso* and *C. sudassana* (Table 1). Many of the samples procured as *C. buddha* proved to be misidentifications of either *C. davidis* or *C. sudassana*. In addition, the samples of *C. buddha* (COI: EF683684, AY090208; RpS5: EU141406) and *C. davidis* (COI: HQ658143) retrieved from GenBank were re-identified respectively as *C. davidis* and (possibly) *C. lactoris* (Table 1).

Male genitalia in *Calinaga* were found to be very uniform and thus genitalic characters were of little use in discrimination of species. Although Lang (2012) emphasized the shape and ‘pointiness’ of the dorsal apex of the valve as an important character, we did not find this trait useful or consistent within species. All the adult and genitalia images are available in Supplementary Material S1.

Phylogenetic analyses and network analysis

For COI (658 bp), 54 sequences (51 sequences analyzed in this study plus 3 sequences retrieved from GenBank) showed 23 different haplotypes (Fig. 1B; Table 1) with 71 variable sites. The 24 RpS5 sequences (483 bp) displayed 4 haplotypes and 6 variable sites.

Bayesian phylogeny analysis was used to reconstruct phylogenetic relationships (Fig. 2) for the concatenated genes. The resulting tree was based on an alignment of 1310 bp and it was rooted using outgroup sequences of members of two subfamilies: Charaxinae (*Charaxes castor*, *Anaea troglodyte*, *Archaeoprepona demophon*, *Euxanthe wakefieldi*) and Satyrinae (*Caligo telamonius*, *Morpho peleides*, *Bicyclus anynana*, *Melanitis leda*) (Table 1). The selected model of evolution for RpS5 nuclear gene was GTR+G+I (Rodriguez et al. 1990) with estimated base frequencies, 4 gamma categories, partitioned into positions (1+2),3 and for COI was HKY (Hasegawa et al. 1985) with estimated base frequencies, no heterogeneity model and not partitioned.

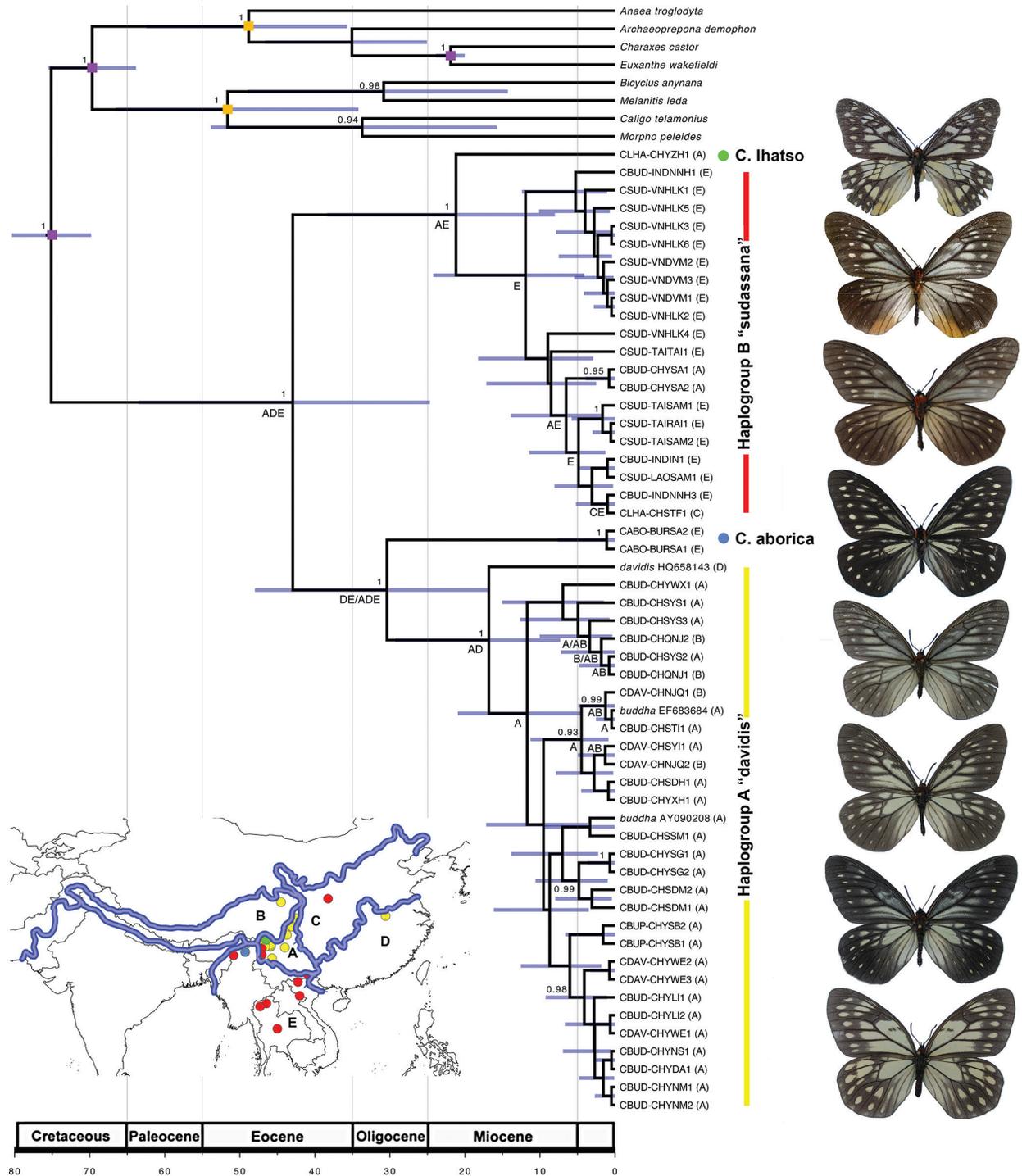


Figure 2. Bayesian phylogeny of *Calinaga* estimated in BEAST using concatenated data. Purple squares are calibration points (root: 75 ± 3 ; Satyrinae + Charaxinae 70 ± 3.5 , Charaxes + Euxanthe 22 ± 1). Monophyly was enforced on nodes marked with orange squares. The inset map shows the biogeographic regions used in DIVA analysis: **A)** Southwestern China ecozone, **B)** Himalaya-Tibetan plateau region, **C)** Northern Sino-Himalaya, **D)** Southern Sino-Himalaya, **E)** Indochina. Colored dots correspond to haplogroups on the tree.

The Bayesian tree for concatenated genes resulted in a well-supported *Calinaga* clade (Fig. 2). Beside distinct lineages for *C. lhatso* (N= 1) and *C. aborica* (N= 2), two major clades were observed. The first clade, consisting of specimens identified as *C. buddha* (N= 20), *C. davidis* (N= 6) and *C. buphonas* (N= 2), was called Haplogroup A “*davidis*” because all samples were re-identified as *C. davidis*

(Table 1). The latter also included two GenBank sequences for *C. buddha* (COI: EF683684, AY090208; RpS5: EU141406). Examination of the voucher images for these sequences showed that these also were indeed *C. davidis* (Min et al. 2008; Peña and Malm 2012). The second clade, consisted of specimens initially identified as *C. sudassana* (N= 14), *C. buddha* (N= 5) and *C. lhatso* (N= 1) and it was

called Haplogroup B “*sudassana*” because all samples were subsequently re-identified as *C. sudassana* (Table 1).

The MJ network analysis (Fig. 1.B) of the mtDNA haplotypes was fully consistent with our Bayesian phylogenetic analysis. Forty-five mutational steps joined Haplogroup A “*dauidis*” (including all *C. dauidis* sequences; Fig. 1B) and Haplogroup B “*sudassana*” (including all *C. sudassana* sequences with *C. lhatso* sequences CLHA-CHSFT1). The genetic differentiation between the two haplogroups, estimated as the mean distance (Kimura 2 parameter uncorrected p-distance) between groups, was 10.3%. The analysis also highlighted a star-like configuration of Haplogroup A “*dauidis*” with the H₁₅ haplotype shared among Yunnan populations localized in the headwater area for the three major rivers flowing from Tibet (Salween, Mekong and Yangtse; Fig. 1A, Fig. 2). In this geographical area *Callinga* displays its highest genetic variability, and Haplogroup A “*dauidis*” and B “*sudassana*” coexist (Fig. 2). Thirty-four and thirty-three mutational steps joined *C. aborica* with Haplogroup A and B respectively. In addition, in Haplogroup A “*dauidis*”, a GenBank sequence for *C. dauidis* (HQ658143) from Tianma National Nature Reserve in Jinzhai County, Anhui, China (Xia and Hao 2011), possibly representing the taxon *C. lactoris* Fruhstorfer 1908, was divergent (Fig. 1B). Finally, in the Haplogroup B “*sudassana*”, the single specimen CLHA-CHSFT1 from Tsinling Mts, Foping nature reserve in Shaanxi, initially identified as *C. lhatso* based on morphological similarity, showed identical COI and RpS5 sequences to *C. sudassana*. Extraction and amplification of this sample was conducted twice in two different laboratories and generated the same results. Eleven mutational steps (Fig. 1B) separated the DNA barcode sequence of this individual from the other *C. lhatso* from Yunnan (CLHA-CHYZH1) despite their similar wing patterns.

Tests of demographic equilibrium and estimation of divergence times

According to F_s and R₂ statistics, calculated for Haplogroup A “*dauidis*” (N= 20) and Haplogroup B “*sudassana*” (N= 30) mtDNA sequence sets (Table 2), the null hypothesis of constant population size could be rejected for these two phylogeographic units (bold in Table 2). Mismatch distribution of these groups was examined according to sudden expansion model (Table 2), and goodness of fit tests did not show significant deviations from expected distributions, so that parameter τ = 2μt could be

used to estimate the time (t) elapsed from population expansion: estimated values of τ and their 5% and 95% confidence limits are shown in Table 2. According to Brower (1994) proposed mutation rates in Lepidoptera (μ = 0.01 substitutions/site/lineage/Ma), demographic expansion of the Haplogroup B “*sudassana*” could be traced back to t = 260 kya (126-403 kya). A similar estimate was obtained for Haplogroup A “*dauidis*” t = 168 kya (70-362 kya).

The most recent common ancestor of *Calinaga* has been estimated to have split from Satyrinae + Charaxinae sometime between 67-84 Mya (Wahlberg et al. 2009). Results of our BEAST analysis reveal that the first major split within *Calinaga* occurred in the Eocene, around 43 Mya (63-25 Mya) (Fig. 2). The time of the most recent ancestor (tMRCA) of *C. aborica* + “*dauidis* group” was estimated to have occurred in the Oligocene ~30 Mya (47-17 Ma), whereas it was ~21 Mya (38-8 Ma) for *C. lhatso* + “*sudassana* group”. Finally, the tMRCA for both *dauidis* and *sudassana* groups was estimated at ~12 Mya (24-4 Ma).

Biogeographic ancestral reconstructions

Results of the ancestral area reconstructions showed that the common ancestor of *Calinaga* probably occurred before ~43 Mya in area ADE (Fig. 2), corresponding to Southwestern China ecozone, Southern Sino-Himalaya and Indochina. Before ~30 Mya this also represented the ancestral area for clade *C. aborica* + Haplogroup A “*dauidis*”. Subsequently in the Miocene before ~15 Mya, in a more restricted area ranging between the Southwestern China ecozone and Southern Sino-Himalaya (AD), the common ancestor of Haplogroup A “*dauidis*” was present with subsequent dispersal (in the late Miocene and Pleistocene) into the Himalaya-Tibetan Plateau region. The common ancestor of clade *C. lhatso* + Haplogroup B “*sudassana*” colonized the Southern Sino-Himalaya and Indochina (AE) before ~20 Mya. Finally, before ~10 Mya Indochina (E) represented the ancestral area of the Haplogroup B “*sudassana*”.

Discussion

Taxonomic implications

Despite six species initially identified (Fig. 1, Table 1), our molecular results identified only four main groups (*C. aborica*, *C. sudassana*, *C. dauidis* and *C. lhatso*). These were also consistent with subsequent careful mor-

Table 2. N, number of mtDNA sequences; H, number of unique mtDNA haplotypes; h, mtDNA haplotype diversity (±SD); π, mtDNA nucleotide diversity (±SD). Tests of demographic equilibrium and mismatch analysis in mtDNA phylogeographic groups: FS, Fu’s FS statistic; R₂, Rozas and Ramon-Onsins’ R₂ statistic; τ, sudden demographic expansion parameter, with 5% and 95% confidence limits; t, true time since population expansion, from τ = 2 μ T (where μ is the substitution rate per gene). Significantly small values of F_s and R₂ are in bold.

| Haplogroup | N | H | h | π (X 10 ²) | F _s | R ₂ | τ | τ (5%) | τ (95%) | t (ka) | t (5%) | t (95%) |
|------------|----|----|-------------|------------------------|----------------|----------------|-------|--------|---------|--------|--------|---------|
| A | 30 | 13 | 0.885±0.042 | 0.459±0.575 | -4.203 | 0.121 | 3.867 | 1.918 | 6.104 | 260 | 126 | 403 |
| B | 20 | 7 | 0.8±0.068 | 0.288±0.3 | -1.369 | 0.143 | 2.537 | 1.061 | 5.498 | 168 | 70 | 363 |

phological identifications (Fig. 2, Table 1). This is quite reflective of the persistent confusion about the taxonomic status and number of species within *Calinaga*. Historically, various authors have recognized different numbers and sets of valid *Calinaga* species, from one (Ehrlich 1958b) or three (Fruhstorfer 1914) to seven (D’Abrera 1985, 1993, Beccaloni et al. 2003), ten (Savela 2015) or eleven species (Sondhi et al. 2016). Many of the samples we procured as *C. buddha* proved to be misidentifications of either *C. davidis* or *C. sudassana*. A thorough examination of the type material of all the names under *Calinaga* is required to clarify whether some of these names should be synonymized. Within our examined specimens we found that *C. lhatso* and *C. aborica* were morphologically and genetically distinct entities. However, the identity of the specimen identified as *C. lhatso* (CLHA-CHSFT1), sharing its haplotype with *C. sudassana*, remains unclear. Many of the specimens initially identified to sub-species under *C. davidis* showed little or no genetic differentiation (e.g. *C. davidis nubilosa*, *C. davidis genestieri*, *C. davidis cercyon*, *C. davidis buphonas* etc.).

Historical biogeography

The question of where and when *Calinaga* originated and how its species are related have not been previously investigated. Although no fossils are known for the genus, previous phylogenetic studies on Nymphalidae that have used fossils for molecular calibration have dated the last common ancestor of *Calinaga* + (Satyrinae + Charaxinae) at 67–84 Mya (Peña and Wahlberg 2008; Wahlberg et al. 2009). Since these studies used only a single *Calinaga* species, the age of the crown group of *Calinaga* and divergences within the genus have remained unknown. Our results show that the first split in *Calinaga* occurred in Eocene at around 43 Mya (Fig. 2). The large gap between this date and the age of the most recent common ancestor of *Calinaga* and Satyrinae + Charaxinae (67–84 Mya) suggests diversification after the Cretaceous/Tertiary boundary with subsequent extinction of sister lineages (Ree and Smith 2008). The initial collision between India and Asia in the early Cenozoic (50–55 Ma) set in motion major orogenic associated environmental changes that continued through the Oligocene, Miocene and into modern times which were likely a major driving force in evolution of *Calinaga*, and numerous other plant (e.g. Liu et al. 2002, Mao et al. 2010) and animal lineages (e.g. Rüber et al. 2004, Zhang et al. 2006, Nazari et al. 2007, Che et al. 2010, Zhao et al. 2015) found in the Himalayan-Tibetan Plateau. In fact, the evolutionary history of organisms in this region helps elucidate parallel geological and climatic developments, and often aids the recognition of unrecorded geological events.

Our results suggest that the common ancestor of *Calinaga* probably occurred before ~43 Mya (63–25 Mya) in the region now corresponding to the Southwestern China ecozone, Southern Sino-Himalaya and Indochina (ADE: Fig. 2). Since the Indochinese peninsula did not exist before ~22 Mya (Leloup et al. 2001), the ancestral distribu-

tion of *Calinaga* was likely confined to the Southwestern China ecozone (D) and Southern Sino-Himalaya (A). Since its center of genetic diversity seems to lie on the south-eastern border of the Tibetan Plateau, in particular where the three major rivers sourced in Tibet (Salween, Mekong and Yangtze rivers; Fig. 1A, Fig. 2) are in proximity, *Calinaga* may well have originated in the Tibetan region. Geological records indicate that around the Oligocene/Miocene boundary (~23 Mya), the rapid extrusion of Indochina and the cumulative Cenozoic deformation of South-East Asia caused by underthrusting of the Indian Plate beneath the Eurasian Plate, resulted in intense orogeny in South-East Tibet (Zhang et al. 2013). In particular in the Early Miocene (~20 Mya), driven by rapid deformation of the eastern Himalayan syntaxis (Robinson et al. 2014), large modifications were induced in the river systems of this region (e.g. Salween, Mekong, Yangtze and Yarlung Tsangpo-Irrawaddy river: Zhang and Hu 2012). These geological upheavals have likely provided ample opportunities for allopatric isolation and subsequent speciation in *Calinaga*.

Our molecular dating suggests that *C. lhatso* and Haplogroup B “*sudassana*” split into distinct lineages in Early Miocene ~21 Mya (38–8 Mya), possibly isolating the last common ancestor of *C. lhatso* in the Southern Sino-Himalaya (A). Southern Sino-Himalaya and Indochina (AE) possibly represented the ancestral area for *C. lhatso* + Haplogroup B “*sudassana*”. Considering that before ~22 Mya the Indochinese peninsula did not exist, it is more likely that the ancestral area of this clade was mainly in Southern Sino-Himalaya (A). The common ancestor of Haplogroup B “*sudassana*” began diversifying around ~12 Mya (24–4 Mya) and it probably expanded its range from Southern Sino-Himalayan region (A) to Indochina peninsula (E). Conversely, the origin of *C. aborica* + Haplogroup A “*davidis*” must be traced back to Oligocene (~30 Mya), when the first uplift of the Himalayan–Tibetan Plateau occurred (29 – 65 Mya: Zhao et al., 2015). Afterwards, in Miocene ~12 Mya (21–4 Mya), Haplogroup A “*davidis*” diversified and expanded its range into Central-East China.

In addition, our analyses show that Haplogroup A “*davidis*” and Haplogroup B “*sudassana*” experienced demographic expansion in the Pleistocene, ~168 kya (70–363 kya) and ~260 ka (16–403 kya) respectively. Therefore, the genetic diversity and the present-day distributions of *Calinaga* should be attributed not only to the dramatic geological events in Oligocene/Miocene, but also to the impacts of the cycles of glacial advance/retreat that occurred in this region during the Pleistocene (Quan et al. 2014).

Finally, the taxon *C. formosana*, for which we examined several specimens but could not obtain any sequences (Supplementary Material S1), is morphologically allied to the Haplogroup A “*davidis*”. Since Taiwan formed ~4–5 Mya at a complex convergent boundary between the Philippine Sea Plate and the Eurasian Plate (“The Penglai Orogeny”: Teng 1990, Liu et al. 2000), the most plausible hypothesis is that *C. formosana* must have dispersed from mainland China to Taiwan recently in Pleistocene.

Conclusion

The genus *Calinaga* probably originated in the South-East Tibet in Eocene following the immense geological and environmental impact caused by the collision between Indian and Asian subcontinents. The extrusion of Indochina from the continent during the Oligocene/Miocene further prompted dramatic orogenetic changes that promoted isolation and speciation events in the genus. More recently, in the Pleistocene, climatic changes further modified the distribution of species and probably facilitated vicariant speciation events.

Since we did not sample or sequence specimens from all of the available names under *Calinaga*, we cannot make any definitive statements about the number of valid species warranted to be recognized as such, although the existence of many superfluous names is evident. From the names of the genus and the species coined by early British lepidopterists including F. Moore, it is apparent that they drew inspiration from Hindu mythological characters. In Sanskrit, Nāga refers to mythical reptilian creatures found in Indian religions (Hinduism, Buddhism and Jainism) who were often worshipped as deities. Among them, “Kaliya” (or Kalya, “Kalia-Naga”, *Calinaga*) was a particularly notorious and poisonous one living in Yamuna river in Vrindavan (Uttar Pradesh). After an encounter with Krishna, Kaliya surrendered and was sent to exile (Bhagavata Purana, 16:10). It seems that the modern taxonomy of *Calinaga* is in need of a Krishna to conquer these superfluous names and cleanse its taxonomy albeit after careful examination of the types and sequencing of additional material.

Acknowledgements

This study was supported in the framework of the SHIBU project (DNA barcoding, molecular systematics and evolution of Sino-Himalayan butterflies) funded by Tor Vergata University grants to Valerio Sbordoni and by grants to Paul D.N. Hebert from NSERC and from the Government of Canada through Genome Canada and the Ontario Genomics Institute in support of the International Barcode of Life Project. We are extremely grateful to Niklas Wahlberg for providing valuable insights in the early stages of this study, to Vadim Tshikolovets for providing identification materials, to Massimo Di Rao and Marco Lucarelli for the genitalia preparations and to Prof. R.I. Vane-Wright, Rodolphe Rougerie and an anonymous reviewer for their helpful comments.

References

- An Z (2000) The history and variability of the East Asian paleomonsoon climate. *Quaternary Science Reviews* 19: 171–187. [https://doi.org/10.1016/S0277-3791\(99\)00060-8](https://doi.org/10.1016/S0277-3791(99)00060-8)
- Ashizawa H, Muroya Y (1967) Notes on the early stages of *Calinaga buddha formosana* Fruhstorfer. *Special Bulletin of Lepidopterological Society of Japan* 1967: 79–89.
- Bain RH, Hurley MM (2011) A biogeographic synthesis of the amphibians and reptiles of Indochina. *Bulletin of the American Museum of Natural History* No. 360, 138 pp.
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036>
- Beccaloni G, Scoble M, Kitching I, Simonsen T, Robinson G, Pitkin B, Hine A, Lyal C [Eds] (2003) online. The Global Lepidoptera Names Index (LepIndex). World Wide Web electronic publication. <http://www.nhm.ac.uk/entomology/lepindex> [accessed 16th May 2016]
- Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato*, inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the U.S.A.* 91: 6491–6495. <https://doi.org/10.1073/pnas.91.14.6491>
- Che J, Zhou W, Hua JS, Yana F, Papenfuss TJ, Wake DB, Zhang YP (2010) Spiny frogs (Paini) illuminate the history of the Himalayan region and Southeast Asia. *Proceedings of the National Academy of Sciences USA* 107: 13765–13770. <https://doi.org/10.1073/pnas.1008415107>
- Chou L (1994) *Monographia Rhopalocerum Sinensium* Vols. 1–2. Henan Science and Technology Press, China, 854 pp.
- Clark AH (1947) The interrelationships of the several groups within the butterfly superfamily Nymphalioidea. *Proceedings of the Entomological Society of Washington* 49: 148–149.
- Clark AH (1948) Classification of the butterflies with the allocation of the genera occurring in North America, North of Mexico. *Proceedings of the Biological Society of Washington* 61: 77–81.
- D’Abbrera B (1985) *Butterflies of the Oriental Region. Part II. Nymphalidae, Satyridae and Amathusiidae*. Hill House, Melbourne.
- D’Abbrera B (1993) *Butterflies of the Holarctic Region. Part III. Nymphalidae (concl.), Libytheidae, Pieridae, Riodinidae, Lycaenidae*. Hill House, Black Rock, Victoria.
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods* 9(8): 772. <https://doi.org/10.1038/nmeth.2109>
- de Bruyn M, Stelbrink B, Morley RJ, Hall R, Carvalho GR, Cannon CH, Van Den Bergh G, Meijaard E, Metcalfe I, Boitani L, Maiorano L, Shoup R, Von Rintelen T (2014) Borneo and Indochina are major evolutionary hotspots for Southeast Asian biodiversity. *Systematic Biology* 63(6): 879–901. <https://doi.org/10.1093/sysbio/syu047>
- de Niceville L (1897) Little known butterflies from the Indo- and Austro-Malayan Regions. *Journal of the Asiatic Society of Bengal* 66: 543–577.
- deWaard JR, Ivanova NV, Hajibabaei M, Hebert PDN (2008) Assembling DNA barcodes: analytical protocols. In: Martin C (Ed.) *Methods in molecular biology: environmental genetics*. Humana Press, Totowa, NJ, 275–293. https://doi.org/10.1007/978-1-59745-548-0_15
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29: 1969–1973. <https://doi.org/10.1093/molbev/mss075>
- Ehrlich PR (1958a) A note on the systematic position of the butterfly genus *Calinaga* (Nymphalidae). *The Lepidopterists’ News* 12: 173.
- Ehrlich PR (1958b) The comparative morphology, phylogeny and higher classification of the butterflies (Lepidoptera: Papilionoidea). *The University of Kansas Science Bulletin* 39(8): 305–370.
- Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* 13: 853–864. <https://doi.org/10.1046/j.1365-294X.2003.02004.x>

- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47–50.
- Fontaine H, Workman DR (1997) In: Moores EM, Fairbridge RW (Eds) Chapman and Hill: London. *Encyclopedia of European and Asian Regional Geology*: 774–782. https://doi.org/10.1007/1-4020-4495-X_103
- Freitas AVL, Brown KSJ (2004) Phylogeny of the Nymphalidae (Lepidoptera: Papilionoidea). *Systematic Biology* 53: 363–383. <https://doi.org/10.1080/10635150490445670>
- Fruhstorfer H (1908) Eine neue *Calinaga* aus China. *Entomologische Zeitschrift*, 22: 147.
- Fruhstorfer H (1912–1914) Family: Nymphalidae. In: Seitz A (Ed) *Die Gross-Schmetterlinge der Erde*, Alfred Kernen, Stuttgart 9: 453–766, pls 107–137.
- Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.
- Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* 52: 696–704. <https://doi.org/10.1080/10635150390235520>
- Hall R (1998) The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. In: Hall R, Holloway JD (Ed) Leiden: Backhuys Publishers. *Biogeography and Geological Evolution of Southeast Asia*: 99–124.
- Hall R (2001) Cenozoic reconstruction of SE Asia and the SW Pacific: Changing patterns of land and sea. In: Metcalfe I, Smith JMB, Morwood M, Davidson I (Eds) Balkema AA. Lisse, Faunal and floral migrations and evolution in SE Asia–Australasia, 35–56.
- Hasegawa M, Kishino H, Yano T-A (1985) Dating of the human–ape splitting by molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160–174. <https://doi.org/10.1007/BF02101694>
- Hajibabaei M, deWaard JR, Ivanova NV, Ratnasingham S, Dooh RT, Kirk SL, Mackie, PM, Hebert PDN (2005) Critical factors for assembling a high volume of DNA barcodes. *Philosophical Transactions of the Royal Society B* 360: 1959–1967. <https://doi.org/10.1098/rstb.2005.1727>
- Heikkilä M, Kaila L, Mutanen M, Peña C, Wahlberg N (2012) Cretaceous origin and repeated tertiary diversification of the redefined butterflies. *Proceedings of the Royal Society B* 279: 1093–1099. <https://doi.org/10.1098/rspb.2011.1430>
- Hudson RR (1990) Gene genealogies and the coalescent process. *Oxford surveys in evolutionary biology*, Vol. 7. In: Futuyma DJ, Antonovics JD (Ed.) Oxford University Press, Oxford, 1–44.
- Ivanova NV, deWaard JR, Hebert PDN (2006) An inexpensive, automation friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes* 6: 998–1002. <https://doi.org/10.1111/j.1471-8286.2006.01428.x>
- Kirby WF (1871) *Synonymic Catalogue of Diurnal Lepidoptera*. John Van Voorst, London, 690 pp.
- Kirby WF (1877) *Synonymic Catalogue of Diurnal Lepidoptera*. Supplement. John Van Voorst, London, 883 pp. <https://doi.org/10.5962/bhl.title.11413>
- Lang SY (2012) *The Nymphalidae of China (Lepidoptera, Rhopalocera)*. Part 1: Libytheinae, Danainae, Calinaginae, Morphinae, Heliconiinae, Nymphalinae, Charaxinae, Apaturinae, Cyrestinae, Biblidinae, Limenitinae. Tshikolovets Publications, 456 pp.
- Lee CL (1958) *Butterflies*. Department of Entomology, Academia Sinica (Peking), Publication no. 4, 198 pp., 4 pls. 198 text figs. [in Chinese]
- Leech JH (1892) *Butterflies from China, Japan, and Corea*. Porter RH, London, 297 pp. <https://doi.org/10.5962/bhl.title.119071>
- Leleup PH, Arnaud N, Lacassin R, Kienast JR, Harrison TM, Trong TTP, Replumaz A, Taponnier P (2001) New constraints on the structure, thermochronology, and timing of the Ailao Shan-Red River shear zone, SE Asia. *Journal of Geophysical Research* 106: 6683. <https://doi.org/10.1029/2000JB900322>
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
- Liu TK, Chen YG, Chen WS, Jiang SH (2000) Rates of cooling and denudation of the Early Penglai Orogeny, Taiwan, as assessed by fission-track constraints. *Tectonophysics* 320: 69–82. [https://doi.org/10.1016/S0040-1951\(00\)00028-7](https://doi.org/10.1016/S0040-1951(00)00028-7)
- Liu J-Q, Gao T-G, Chen Z-D, Lu A-M (2002) Molecular phylogeny and biogeography of the Qinghai–Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Molecular Phylogenetics and Evolution* 23: 307–325. [https://doi.org/10.1016/S1055-7903\(02\)00039-8](https://doi.org/10.1016/S1055-7903(02)00039-8)
- Mao K, Hao G, Liu J, Adams R, Milne R (2010) Diversification and biogeography of *Juniperus* (Cupressaceae): variable diversification rates and multiple intercontinental dispersals. *New Phytologist* 188: 254–272. <https://doi.org/10.1111/j.1469-8137.2010.03351.x>
- Martens J, Tietze DT, Päckert M (2011) Phylogeny, biodiversity, and species limits of passerine birds in the Sino-Himalayan region—A critical review. *Ornithological Monographs* 70: 64–94. <https://doi.org/10.1525/om.2011.70.1.64>
- Melville JC (1893) Description of a new butterfly of the genus *Calinaga* from Siam. *Transactions of the Entomological Society of London* 1893(2): 121–122. <https://doi.org/10.1111/j.1365-2311.1893.tb02059.x>
- Min Z, Yang Z, Tianwen C, Yupeng G, Yuan Z, Ke J, Zhumei R, Rui Z, Yaping G, Enbo M (2008) Phylogenetic relationship and morphological evolution in the subfamily Limenitidinae (Lepidoptera: Nymphalidae). *Progress in Natural Science* 18: 1357–1364. <https://doi.org/10.1016/j.pnsc.2008.03.025>
- Monastyrskii AL, Holloway JD (2013) The biogeography of the butterfly fauna of Vietnam with a focus on the endemic species (Lepidoptera). *Current Progress in Biological Research*. <http://dx.doi.org/10.5772/55490>
- Moore F (1857) In Horsfield, T. and Moore, F., *A Catalogue of the Lepidopterous Insects in the Museum of the Hon. East-India Company*. Vol. I. London, WM. H. Allen and Co., 370 pp.
- Moore F (1895) *Lepidoptera Indica*. Vol. II: Nymphalidae (Elymniinae, Amathusiinae, Nymphalinae). L. Reeve and Co., London, 274 pp., 190 pls.
- Morley RJ (2000) *Origin and evolution of tropical rain forests*. Chichester: John Wiley, 362 pp. <https://doi.org/10.1038/35002501>
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Nazari V, Zakharov EV, Sperling FAH (2007) Phylogeny, historical biogeography, and taxonomic ranking of Parnassiinae (Lepidoptera, Papilionidae) based on morphology and seven genes. *Molecular Phylogenetics and Evolution* 42: 131–156. <https://doi.org/10.1016/j.ympev.2006.06.022>
- Oberthür C (1920) *Etudes de lépidoptérologie compare* 17. Plates DXII–DXVI.
- Peña C, Wahlberg N (2008) Prehistorical climate change increased diversification of a group of butterflies. *Biology Letters* 4: 274–278. <https://doi.org/10.1098/rsbl.2008.0062>

- Peña C, Malm T (2012) VoSeq: a Voucher and DNA sequence web application. *PLoS ONE* 7(6): e39071. <https://doi.org/10.1371/journal.pone.0039071>
- Procheş Ş, Ramdhani S (2012) The world's zoogeographical regions confirmed by cross-taxon analyses. *BioScience* 62: 260–270. <https://doi.org/10.1525/bio.2012.62.3.7>
- Quan Q-M, Chen L-L, Wang X, Li S, Yang X-L, Zhu Y-G et al (2014) Genetic diversity and distribution patterns of host insects of caterpillar fungus *Ophiocordyceps sinensis* in the Qinghai-Tibet Plateau. *PLoS ONE* 9(3): e92293. <https://doi.org/10.1371/journal.pone.0092293>
- Rambaut A (2009) Figtree v1. 2.2. Computer program and documentation distributed by the author. Available at <http://tree.bio.ed.ac.uk/software> [accessed 04 June 2013]
- Ramos-Onsins S, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* 19: 2092–2100. <https://doi.org/10.1093/oxfordjournals.molbev.a004034>
- Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes* 7: 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>
- Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution* 20: 76–86. <https://doi.org/10.1093/molbev/msg009>
- Ree RH, Smith SA (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57: 4–14. <https://doi.org/10.1080/10635150701883881>
- Robinson RAJ, Brezina CA, Parrish RR, Horstwood MSA, Nay Win O, Bird MI, Myint T, Walters A S, Oliver GJH, Khin Z (2014) Large rivers and orogens: the evolution of the Yarlung Tsangpo–Irrawaddy system and the eastern Himalayan syntaxis. *Gondwana Research* 26: 112–121. <https://doi.org/10.1016/j.gr.2013.07.002>
- Rodríguez F, Oliver JL, Marín A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* 142(4): 485–501. [https://doi.org/10.1016/S0022-5193\(05\)80104-3](https://doi.org/10.1016/S0022-5193(05)80104-3)
- Ronquist F (1996) DIVA version 1.1. Computer program and manual available by anonymous FTP from Uppsala University: <ftp.uu.se> or <ftp.systbot.uu.se>
- Rüber L, Britz R, Kullander SO, Zardoya R (2004) Evolutionary and biogeographic patterns of the Badidae (Teleostei: Perciformes) inferred from mitochondrial and nuclear DNA sequence data. *Molecular Phylogenetics and Evolution* 32: 1010–1022. <https://doi.org/10.1016/j.ympev.2004.04.020>
- Savelle M (2015) *Calinaga*. Tree of Life (online), http://ftp.funet.fi/index/Tree_of_life/insecta/lepidoptera/ditrysia/papilionoidea/nymphalidae/calinaginae/calinaga/ [accessed 10th October 2016]
- Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites application to human mitochondrial DNA. *Genetics* 152: 1079–1089.
- Seitz A (1909) Die Großschmetterlinge der Erde, Band 1: Abt. 1, Die Großschmetterlinge des palaearktischen Faunengebietes, Die palaearktischen Tagfalter. Seiten. Verlag Alfred Kernen, Stuttgart, 379 pp, 89 pl.
- Seitz A (1927) Die Großschmetterlinge der Erde, Band 9: Abt. 2, Die exotischen Großschmetterlinge, Die indo-australischen Tagfalter. Verlag Alfred Kernen, Stuttgart, 1197 pp, 177 pl.
- Shirôzu T (1960) Butterflies of Formosa in Color. Osaka, Hoikusha, 481 pp.
- Sondhi S, Karmakar T, Sondhi Y, Jhaveri R, Kunte K (2016) Re-discovery of *Calinaga aborica* Tytler, 1915 (Lepidoptera: Nymphalidae: Calinaginae) from Arunachal Pradesh, India. *Journal of Threatened Taxa* 8: 8618–8622. <https://doi.org/10.11609/jott.2354.8.3.8618-8622>
- Sterling EJ, Hurley MM, Le DM (2006) Vietnam: A Natural History. Yale University Press, New Haven, CT, 448 pp.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30: 2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Teng LS (1990) Geotectonic evolution of late Cenozoic arc–continent collision in Taiwan. *Tectonophysics* 183: 57–76. [https://doi.org/10.1016/0040-1951\(90\)90188-E](https://doi.org/10.1016/0040-1951(90)90188-E)
- Todisco V, Gratton P, Cesaroni D, Sbordoni V (2010) Phylogeography of *Parnassius apollo*: hints on taxonomy and conservation of a vulnerable glacial butterfly invader. *Biological Journal of the Linnean Society* 101: 169–183. <https://doi.org/10.1111/j.1095-8312.2010.01476.x>
- Tytler HC (1915) Notes on some new and interesting butterflies from Manipur and the Naga Hills. Part II. *Journal of Bombay Natural History Society* 23: 502–515.
- Wahlberg N, Wheat CW (2008) Genomic outposts serve the phylogenomic pioneers: Designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Systematic Biology* 57: 231–242. <https://doi.org/10.1080/10635150802033006>
- Wahlberg N, Leneveu J, Kodandaramaiah U, Peña C, Nylin S, Freitas AVL, Brower AVZ (2009) Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. *Proceedings of the Royal Society B* 276: 4295–4302. <https://doi.org/10.1098/rspb.2009.1303>
- Xia J, Hao JS (2011) Sequencing and analysis of the complete mitochondrial genome of *Calinaga davidis* Oberthur (Lepidoptera: Nymphalidae). *Acta Entomologica Sinica* 54: 555–565.
- Zhang P, Che Y-Q, Zhou H, Liu Y-F, Wang X-L, Papenfuss TJ, Wake DB, Qu L-H (2006) Phylogeny, evolution, and biogeography of Asiatic Salamanders (Hynobiidae). *Proceedings of the National Academy of Sciences of the United States of America* 103: 7360–7365. <https://doi.org/10.1073/pnas.0602325103>
- Zhang X, Hu S (2012) The immature stage of *Calinaga buphona* Oberthür, 1920 (Lepidoptera: Nymphalidae). *Far Eastern Entomologist* 239: 10–12.
- Zhang H, Clift PD, Wang P, Tada R, Jia J, He M, Jourdan F (2013) Pre-Miocene birth of the Yangtze River. *Proceedings of the National Academy of Sciences of the United States of America* 110: 7556–7561. <https://doi.org/10.1073/pnas.1216241110>
- Zhao JL, Xia YM, Cannon CH, Kress WJ, Li QJ (2015) Evolutionary diversification of alpine ginger reflects the early uplift of the Himalayan-Tibetan Plateau and rapid extrusion of Indochina. *Gondwana Research* 32: 232–241. <https://doi.org/10.1016/j.gr.2015.02.004>

Supplementary material 1

Adults and genitalia images of dissected specimens

Authors: Valentina Todisco, Vazrick Nazari, Donatella Cesaroni, Valerio Sbordoni

Data type: Images

Explanation note: Adult and genitalia images of dissected specimens examined in this study.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
