Received: 8 July 2020 | Revised: 16 September 2020 | Accepted: 5 October 2020

DOI: 10.1111/zsc.12457

ORIGINAL ARTICLE

Zoologica Scripta | WILEY

Historical biogeography of Heteropterinae skippers via Beringian and post-Tethyan corridors

Emmanuel F. A. Toussaint[1,2] | **Emily A. Ellis**[2] | **Riley J. Gott**[2] | **Andrew D. Warren**[2] | **Kelly M. Dexter**[2] | **Caroline Storer**[2] | **David J. Lohman**[3,4,5] | **Akito Y. Kawahara**[2]

[1]Natural History Museum of Geneva, Geneva, Switzerland

[2]McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, University of Florida, Gainesville, FL, USA

[3]Biology Department, City College of New York, City University of New York, New York, NY, USA

[4]Ph.D. Program in Biology, Graduate Center, City University of New York, New York, NY, USA

[5]Entomology Section, National Museum of Natural History, Manila, Philippines

Correspondence
Emmanuel F. A. Toussaint, Natural History Museum of Geneva, CP 6434, CH 1211 Geneva 6, Switzerland.
Email: emmanuel.toussaint@ville-ge.ch

Funding information
Division of Environmental Biology, Grant/Award Number: DEB-1541500 and DEB-1541557

Abstract

Skippers are a species rich and widespread group of butterflies with evolutionary patterns and processes largely unstudied despite some recent efforts. Among Hesperiidae, the subfamily Heteropterinae is a moderately diverse clade comprising ca. 200 species distributed from North to South America and from Africa to the Palearctic region. While some regions are species rich, others are far less diverse. Using anchored phylogenomics, we infer a robust timetree and estimate ancestral ranges to understand the biogeographic history of these skippers. Inferences based on up to 383 exons recover a robust backbone for the subfamily along with the monophyly of all genera. Bayesian divergence time estimates suggest an origin of Heteropterinae in the late Eocene, ca. 40 million years ago. Maximum likelihood ancestral range estimates indicate an origin of the group in the New World. The eastern Palearctic was likely colonized via a Beringian route and a reverse colonization event resulted in two independent and extant American clades. We estimate a vicariant event between Central and South America that significantly predates estimates of the proto-Caribbean seaway closure, indicating active overwater dispersal in the Oligocene. The colonization of Africa from the east Palearctic is synchronous with the closure of the Tethys Ocean, while the colonization of Madagascar appears to be comparatively recent. Our results shed light on the systematics and biogeography of Heteropterinae skippers and unveil the evolutionary history of a new leaf in the skipper tree-of-life.

KEYWORDS

Anchored hybrid enrichment, Beringian route, butterfly biogeography, *Gomphotherium* land bridge, Hesperiidae, Lepidoptera evolution

1 | INTRODUCTION

Recent progress in phylogenomics and museomics (*i.e.* use of DNA from old museum specimens) have allowed a better understanding of Lepidoptera systematics, phylogenetics and evolution (Allio et al., 2020; Breinholt et al., 2018; Dowdy et al., 2020; Espeland et al., 2018; Hamilton et al., 2019; Homziak et al., 2019; Kawahara et al., 2019; Li et al., 2019; St Laurent et al., 2018; Toussaint et al., 2018). Among butterflies (Lepidoptera, Papilionoidea), increasingly well sampled and strongly supported trees-of-life are facilitating new evolutionary insights thanks to the application of new methods. A

striking example is the skippers (Lepidoptera, Papilionoidea, Hesperiidae) which have been the focus of many recent studies (Cong et al., 2019; Li et al., 2019; Toussaint et al., 2018, 2019; Zhang et al., 2019a, 2019b, 2020). This relatively diverse family (ca. 4,300 species, or >20% of butterfly species) is one of the most poorly studied butterfly families despite a wealth of unique ecologies and behaviours (Toussaint & Warren, 2019). The use of anchored phylogenomics and whole-genome sequencing has resolved the backbone of skippers and allowed discovery of new subfamilial-level lineages (Cong et al., 2019; Zhang et al., 2019a, 2020), illustrating the relative oblivion in which skippers had been left for decades. Yet, the fine-scale phylogenetics and evolution of most subfamilies remain poorly understood, and in-depth studies relying on extended taxon sampling are needed.

The skipper subfamily Heteropterinae is a derived lineage sister to the subfamily Hesperiinae (Li et al., 2019; Toussaint et al., 2018) with ca. 200 species distributed across the planet except for the Australasian region (see Warren, 2001 for a detailed introduction to the group). The highest diversity is found in the Neotropics, where the mostly Andean genera *Dalla* and *Ladda* are found (ca. 100 species) along with the monotypic genus *Freemaniana*. The genera *Argopteron* (3 species) and *Butleria* (7 species) are also found in South America but are located in the southernmost regions of the continent, in Argentina and Chile. Most species in the genus *Dardarina* (14 species) inhabit southern Brazil with a single species reaching as far north as Mexico. While about 20 species of *Dalla* occur in Central America, much of the diversity in this region and in the Nearctic is represented by the genus *Piruna* (ca. 20 species). One species in the genus *Carterocephalus* (19 species) is also found in the Nearctic region (*C. palaemon*), although its range extends to the entire Holarctic. The remaining *Carterocephalus* species are found in the Palearctic region where the two monotypic genera *Heteropterus* and *Leptalina* also occur. Finally, three genera are found in the Afrotropics, the genus *Hovala* (5 species) endemic to Madagascar and the genera *Metisella* (22 species) and *Willema* (2 species) distributed throughout Africa.

Although the subfamily Heteropterinae is distributed throughout much of the world, there has been limited foodplant and life-history information reported. *Apostictopterus fuliginosus* Leech [1893] has been bred from *Phyllostachys* sp. and other bamboos (Poaceae) in China (Igarashi & Harada, 2015), *Metisella syrinx* (Trimen, 1868) specializes on mountain bamboo, *Arundinaria tessellata* (Nees) Munro in South Africa (Dickson & Kroon, 1978) and adults of *Argopteron aureum* Peña 1968 have been observed fluttering around another bamboo, *Chusquea* sp., in Chile (Peña G. & Ugarte P., 1996). Additionally, Larsen (2005) suggested the entire genus *Metisella* (including *Willema*) feeds on Poaceae, and Cock and Congdon (2017) reported six partial life histories feeding on a mixture of grasses (Poaceae) in Tanzania and Kenya. The relatively few host plant records that exist for other species in this subfamily document individual Heteropterinae species feeding on a range of non-bamboo grasses (Poaceae) in multiple genera. The species are generally uni- or bivoltine and overwinter as larvae, but several *Metisella* species fly year-round in Africa (Kim & Ho, 2012; Williams, 2020).

Earlier phylogenetic studies of Warren et al. (2008, 2009) placed *Butleria* as sister to a clade comprising *Carterocephalus* and *Metisella* and another with *Piruna* and *Dardarina*. This phylogenetic hypothesis was complemented by the findings of Sahoo et al. (2016, 2017) which recovered *Butleria* as sister to the remainder of Heteropterinae, with *Heteropterus* as the following branch, in turn sister to the remainder of the subfamily except for *Butleria*. The remaining genera were placed in two clades, one comprising the South American *Dalla*, *Dardarina* and *Piruna*, and the other comprising the Old-World *Carterocephalus* and *Metisella*. This result was identical to the one recovered in Chazot et al. (2019) based on the same taxon sampling and data. In their phylogenomic study based on anchored enrichment (>380 exons), Toussaint et al. (2018) recovered *Carterocephalus* and *Piruna* as sister to *Heteropterus* and *Leptalina* with strong nodal support. Other recent phylogenetic hypotheses for the subfamily have placed the genera *Argopteron* and *Butleria* as sister to the remainder of the subfamily (Li et al., 2019). In their study, Cong et al. (2019) recovered *Heteropterus* and *Leptalina* as a clade sister to a larger clade comprising two subclades, one composed of *Carterocephalus*, *Hovala*, *Metisella* and *Willema*, and the second consisting of *Dalla*, *Dardarina*, *Ladda* and *Piruna*. The placement of the monotypic genus *Freemaniana* remains elusive.

Several genera were previously thought to belong in Heteropterinae but were placed in other clades based on recent phylogenomic evidence (Cong et al., 2019; Li et al., 2019; Toussaint et al., 2018; Zhang et al., 2019a). The genus *Tsitana* from South Africa appears to belong in Hesperiinae rather than Heteropterinae (Cong et al., 2019). The Indomalayan monotypic genera *Apostictopterus* and *Barca* are sister and closely related to Trapezitinae (Zhang et al., 2019a). Finally, the Afrotropical monotypic genus *Lepella* is nested within Hesperiinae in the tribe Aeromachini (Toussaint et al., 2018). Moreover, Cong et al. (2019) recently described two new genera, *Ladda* and *Willema*, but taxonomic delineation of these genera remains tentative considering the limited taxon sampling of their study.

Divergence times within Heteropterinae were estimated in Sahoo et al. (2017) based on a combination of secondary and fossil calibrations that recovered a crown age for the split between *Butleria* and the rest of sampled Heteropterinae genera at ca. 50 million years ago (Ma). In a more recent study investigating butterfly divergence times using a comprehensive fossil calibration set, Chazot et al. (2019) recovered the

age of Heteropterinae at ca. 34 Ma with the same taxon sampling. However, despite recovering plausible phylogenetic relationships within Heteropterinae, both of these studies recovered inter-subfamilial phylogenetic relationships among skippers that contradict the most recent phylogenomic studies on Hesperiidae (Cong et al., 2019; Li et al., 2019; Toussaint et al., 2018; Zhang et al., 2019). Hence, these estimates are likely to be biased by unsupported phylogenetic hypotheses. The most robust dated tree of butterflies presented in Espeland et al. (2018) recovers a crown age for the split of *Heteropterus* and *Piruna* at ca. 32 Ma but did not include sufficient taxon sampling to estimate the crown age of the subfamily. Therefore, a rigorous estimate of divergence times in Heteropterinae is needed.

In this study, we infer a new dated phylogenomic hypothesis for the Heteropterinae including all recognized genera and infer the historical biogeography of the group to understand its evolutionary history.

2 | MATERIALS AND METHODS

2.1 | Taxon sampling and molecular biology

Samples of 26 Heteropterinae species were collected in the field or sampled from museum collections, and sequences from 34 additional species from BOLD or previous studies were combined, for a total of 60 out of ca. 200 Heteropterinae described species covering all currently recognized genera (Appendix S1). Several specimens from the Florida Museum of Natural History, McGuire Center for Lepidoptera & Biodiversity (Gainesville, FL, USA) were dissected to allow a reliable species-level identification before DNA extraction (see below). We were able to sample few species from the large genera *Dalla* and *Ladda*, which together include ca. 100 species restricted to South America. We selected 32 outgroup skipper species from other subfamilies, four *Macrosoma* (Hedylidae) species and *Doxocopa agathina* (Nymphalidae) to allow use of secondary calibrations when estimating divergence times (see below).

DNA was extracted from butterfly abdomens or legs using an OmniPrep^{TM} DNA extraction kit (G-Biosciences). Quantified DNA extracts were submitted to RAPiD Genomics (Gainesville, FL, USA) for library preparation, hybridization enrichment and sequencing. Random mechanical shearing of DNA was conducted with an average size of 300 bp followed by an end-repair reaction and ligation of an adenine residue to the 3'-end of the blunt-end fragments to allow ligation of barcode adapters and PCR-amplification of the library. Following library construction, solution-based anchored hybrid enrichment (AHE) of Agilent SureSelect probes was conducted in a pool containing 16 libraries. These libraries were enriched with the SureSelect Target Enrichment System for Illumina Paired-End Multiplexed Sequencing. Paired-end 150-bp reads were sequenced from each library on an Illumina HiSeq.

We used two types of AHE probe sets to generate new genomic data using target exon capture methods. First, we used the BUTTERFLY2.0 probe set (Kawahara et al., 2018) to capture 13 gene regions including those most commonly used in studies on butterfly phylogenetics ("legacy genes"). These genes are as follows: acetyl-CoA (also called thiolase) (ACOA, 1,020 bp), carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase (CAD, 1854 bp), catalase (CAT, 1,290 bp), cytochrome oxidase c subunit 1 (CO1, 1,341 bp), dopa decarboxylase (DDC, 702 bp), elongation factor 1 alpha (EF1A, 1,059 bp), glyceraldehyde-3-phosphate dehydrogenase (GAPDH, 606 bp), hairy cell leukaemia protein 1 (HCL, 633 bp), isocitrate dehydrogenase (IDH, 708 bp), malate dehydrogenase (MDH, 681 bp), ribosomal protein S2 (RPS2, 471 bp), ribosomal protein S5 (RPS5, 555) and wingless (WGL, 240 bp). This approach has proven sufficient to generate robust genus-level phylogenetic hypotheses (Kawahara et al., 2018). Second, we used the BUTTERFLY1.1 probe set (Toussaint et al., 2018) to capture 383 gene regions including all loci captured with the BUTTERFLY2.0 probe set.

2.2 | AHE Data assembly and cleanup

Processing of raw AHE Illumina data followed Breinholt et al. (2018). Paired-end Illumina data were filtered for quality and adaptors were removed using Trim Galore! ver. 0.4.0 (www.bioinformatics.babraham.ac.uk), removing sequences less than a minimum read size of 30 bp and low-quality read ends with base Phred quality scores <20. Additionally, only reads with both forward and reverse reads that passed filtering were retained. The loci were assembled with iterative baited assembly (IBA.py, Breinholt et al., 2018) with a kmer coverage depth of 10 (-c), kmer length of 25 (-k) and paired read gap length of 200 (-g). The resulting assembled loci from the probe region were blasted against the reference genome of *Danaus plexippus* (Linnaeus, 1758), and BLAST results were used for single hit and orthology filtering (ortholog_filter.py, Breinholt et al., 2018). The loci were screened for orthology with a single hit threshold of 90% similarity and genome mapping following Breinholt et al. (2018), with the script ortholog_filter.py. This approach determines whether the probe hit and the reference sequence map to the same scaffold in the reference genome. Hits that do not map to the appropriate scaffold are unlikely to be orthologous and are removed from the data set. The identified orthologous sequences were screened for contamination by identifying and removing sequences that were nearly identical at the family and

genus level (Breinholt et al., 2018). The loci were aligned with MAFFT v7.294b (Katoh & Standley, 2013), and consensus sequences were generated for samples containing isoforms before being concatenated with FASconCAT-G 1.0.4 (Kück & Meusemann, 2010).

We searched the NCBI SRA database for previously published Heteropterinae genomic data. Genomic reads from the SRA database were trimmed using the same trimming scheme detailed above. The paired-end genomic reads were assembled using SPAdes v3.13 (Bankevich et al., 2012) with default parameter settings. The script genome_getprobe.py (Espeland et al., 2018) was used to extract probe regions from all de novo genome assemblies using *Danaus plexippus* as a reference. The program MAFFT v7.294b (Katoh & Standley, 2013) was then used to align and reverse-complement those alignments using—adjustdirectionaccurately. At this point, the genomic data set included AHE probe orthologs, but also paralogs. We removed these paralogs using a multi-layered filtering approach. First, a genome mapping approach was used to confirm orthology, again following Breinholt et al. (2018). The de novo genome assemblies produced low N50 scores (Ellis and Kawahara, in review), which increases the likelihood of erroneously assembled genes and may result in high levels of gene duplication (*e.g.* Denton et al., 2014). In order to determine orthology confidently in the face of these draft assemblies, our data set was further refined using a gene tree approach. We estimated gene trees in FastTree v2.1.7 (Price et al., 2009) with a GTR model and then used PhyloTreePruner (Kocot et al., 2013) to determine the maximally inclusive subtree for each locus. This approach selects the most likely ortholog and prunes all other paralogs, leaving one AHE ortholog per taxon. Additionally, available sequence data from BOLD was imported in Geneious R11 (Biomatters, USA), cleaned and aligned with individual loci using MUSCLE (Edgar, 2004).

We combined all data types into individual loci using FASconCAT-G 1.0.4 (Kück & Meusemann, 2010). Aliscore 2.2 (Kück et al., 2010) was used to check for saturation and sites that appeared to evolve randomly. Detailed information regarding taxon-specific genomic coverage and genomic matrix composition can be found in Appendices 1 and 2. The tree was rooted with *Doxocopa agathina* (Nymphalidae) (Breinholt et al., 2018; Espeland et al., 2018). The final matrix comprised up to 383 concatenated loci totalling 164,099 aligned nucleotides from 97 taxa including 60 out of ca. 200 described Heteropterinae species (Appendix S1). In total, 13 Heteropterinae taxa were sequenced using the BUTTERFLY1.1 probe set (383 loci), 13 were sequenced using the BUTTERFLY2.0 probe set (13 loci), 27 were only represented by a barcode sequence, and seven were obtained from whole-genome sequencing data mapping onto our probe sets.

2.3 | Data sets, model partitioning and phylogenetic analysis

We estimated the best partitioning scheme using PartitionFinder2 (Lanfear et al., 2017) with the *rcluster* algorithm applied to the 383 protein-coding loci (Appendix S2). We used the resulting scheme to estimate the best models of nucleotide substitution with ModelFinder (Kalyaanamoorthy et al., 2017) as implemented in IQ-TREE 2.0-rc1 (Nguyen et al., 2015). To find the most likely tree, 200 maximum likelihood (ML) searches were conducted in IQ-TREE, with two methods of nodal support: ultrafast bootstrap (UFBoot) and SH-aLRT tests. We generated 1,000 replicates for UFBoot ("–bb" command) (Hoang et al., 2018; Minh et al., 2013) and SH-aLRT ("-alrt" command) (Guindon et al., 2010). To reduce the risk of overestimating branch support with UFBoot due to severe model violations, we used hill-climbing nearest neighbour interchange (NNI) to optimize each bootstrap tree ("-bnni" command). All analyses were conducted on the University of Florida HiPerGator High Performance Computing Cluster (www.hpc.ufl.edu). Raw data, data sets and appendices from this study are submitted to a Dryad repository (https://doi.org/10.5061/dryad.zgmsbcc90).

2.4 | Divergence Time Estimation

We estimated divergence times in a Bayesian framework with BEAST 1.10.4 (Suchard et al., 2018). Because the full concatenated data set of 383 loci was too large to be analysed as a whole, we used the concatenated data set corresponding to the 13 loci of the BUTTERFLY2.0 probe set to optimize gene and taxon coverage and selected additional loci using "gene shopping" as implemented in SortaDate (Smith et al., 2018) using some underlying UNIX code and programmes implemented in phyx (Brown et al., 2017). Locus topologies were estimated in IQ-TREE, and the best-scoring ML tree from the concatenated data set was used as the reference species tree in SortaDate analyses. Loci were filtered following three criteria: (a) clock-likeness, (b) reasonable tree length and (c) least topological conflict with the species tree. We selected the 20 best loci based on this filtering and added them to the 13 loci from the BUTTERFLY2.0 probe set.

The best partitioning scheme and models of substitution were determined with PartitionFinder2 (Lanfear et al., 2017) using the *greedy* algorithm and the Bayesian information criterion corrected across all models included in BEAST. The data set was partitioned a priori by locus for a total of 33 initial partitions. We implemented clock partitioning by conducting analyses with (a) two clocks, one for the mitochondrial locus and one for all nuclear loci, and (b) one clock for each partition recovered in PartitionFinder (11 in total, see Results). We assigned a

Bayesian lognormal relaxed clock model to the different clock partitions. We also tested different tree models by using a Yule (pure birth) or a birth–death model in different analyses. Analyses consisted of 100 million generations with parameter and tree sampling every 5,000 generations. We estimated marginal likelihood estimates (MLE) for each analysis using path-sampling and stepping-stone sampling (Baele et al., 2012), with 1,000 path steps, and chains running for one million generation with a log-likelihood sampling every 1,000 cycles. The maximum clade credibility (MCC) trees of each analysis with median divergence age estimates were generated in TreeAnnotator 1.10.4 (Suchard et al., 2018) after removing the first 20 million generations as burn-in.

We used secondary calibrations derived from Espeland et al. (2018) to estimate divergence times because there are few skipper fossils. We constrained 10 nodes with soft lognormal prior densities encompassing the 95% credibility intervals estimated in Espeland et al. (2018) (with mean and standard deviation in real space in BEAUti): crown of Hedylidae + Hesperiidae (mean="106.01" stdev="18.62" offset="2.6"), crown of *Macrosoma hyacinthina* + *M. rubedinaria* (mean="30.71" stdev="7.525" offset="0.42"), crown of Hesperiidae (mean="76.13" stdev="13.09" offset="1.9"), stem of Euschemoninae (mean="67.12" stdev="11.44" offset="1.71"), stem of Eudaminae (mean="59.69" stdev="10.134" offset="1.41"), crown of Barcinae + Hesperiinae+Heteropterinae + Malazinae+Trapezitinae (corresponding to the crown of Hesperiinae + Heteropterinae+Trapezitinae in Espeland et al., 2018) (mean="53.01" stdev="9.034" offset="1.29"), crown of *Heteropterus* + *Piruna* (mean="34.12" stdev="7.412" offset="0.41"), stem of Hesperiinae (mean="48.48" stdev="8.32" offset="1.18") and crown of Hesperiinae (mean="43.9" stdev="7.447" offset="1.09").

2.5 | Ancestral Range Estimation

We used the R-package BioGeoBEARS 1.1.2 (Matzke, 2018) to estimate ancestral ranges in Heteropterinae. We relied on the BEAST MCC tree of the preferred analysis (see Results) without outgroups. Analyses were performed under the dispersal extinction cladogenesis (DEC) model (Ree & Smith, 2008) and a likelihood implementation of the dispersal–vicariance analysis (DIVA) model (Ronquist 1997) (*i.e.* DIVALIKE in BioGeoBEARS). We used the following areas: South America (S), Central America/ Nearctic (N), West Palearctic (W), East Palearctic (E), Africa (A) and Madagascar (M).

We took into account the dynamic geologic history of landmasses since ca. 40 Ma (Seton et al. 2012), by designating three time slices with different dispersal rate scalars; TS1 (40 – 27 Ma) corresponding to a period predating the closure of the Tethys Ocean in the mid-Miocene ca. 27 Ma (Pirouz et al., 2017), an event suggested to have been important in the historical biogeography of Old World skippers (Toussaint et al., 2019), TS2 (27 Ma – 15 Ma) corresponding to a period with a land bridge facilitating the connectivity between Africa and Asia, but also predating the current maximum estimate for the closure of the proto-Caribbean seaway (*i.e.* formation of the Isthmus of Panama) ca. 15 Ma (Bacon et al., 2015; Montes et al., 2015), and TS3 (15—present) corresponding to the a period with enhance connectivity between Central and South America.

The dispersal rate scalar values were selected according to terrain and water body positions throughout the evolutionary timeframe of the group (Appendix S3). The maximum number of areas per ancestral state was set to three. We compared the results of these stratified and designed analyses (M1) with null models (M0) that excluded the dispersal rate scalar and adjacency matrices, thereby relaxing all constraints derived from available geological data. Comparing parameterized and null models allowed us to test whether models taking into account geologic reconstructions are a better fit.

3 | RESULTS AND DISCUSSION

3.1 | Phylogenetic relationships among Heteropterinae

The topology from the best-scoring ML tree search (LnL = −2319884.252) is presented in Figure 1 (see Appendix S4). Nodal support across the topology is high except for branches with samples that had fewer, Sanger-sequenced loci (i.e., legacy loci). We recover Hedylidae as monophyletic (SH-aLRT = 100/UFBoot = 100) and sister to Hesperiidae (SH-aLRT = 100/UFBoot = 100), in line with most recent phylogenomic studies (Kawahara & Breinholt, 2014; Breinholt et al. 2018; Espeland et al., 2018; Toussaint et al., 2018; Kawahara et al., 2019; Li et al., 2019). We recover Heteropterinae as monophyletic (SH-aLRT = 100/UFBoot = 100) and as sister to a large clade comprising the subfamilies Barcinae, Hesperiinae, Malazinae and Trapezitinae (SH-aLRT = 100/UFBoot = 100). Only Zhang et al. (2020) tested the placement of Malazinae relative to other subfamilies and recovered it as sister to Heteropterinae. Their study was based on mitochondrial genomes (ca. 16 kb), whereas our study using 383 exons (ca. 160 kb) recovers Malazinae as sister to Barcinae and Trapezitinae with strong support (SH-aLRT = 97/UFBoot = 94), this latter clade being sister to Hesperiinae also with strong support (SH-aLRT = 100/

MALAZINAE
BARCINAE
TRAPEZITINAE
HESPERIINAE
Butleriini
I
Argopteron aureipennis BN000139
Argopteron puelmae SRR7174478
Butleria bisexguttatus BN005145
Butleria paniscoides BN000195
Butleria elwesi BN005146
Butleria flavomaculatus SRR7174380
II
Heteropterus morpheus RE07G368
Leptalina unicolor BN000462
Heteropterini
III
Hovala pardalina BN003761
Willema willemi BN005142
Metisella metis BN000484
Metisella kambove BN000483
Metisella quadrisignatus BN005141
Carterocephalus avanti HSPEC167_10
Cartero. flavomaculatus HSPED292_11
Carterocephalus micio HSPEC159_10
Carterocephalus dieckmanni BN005143
Cartero. houangty HSPEC169_10
Carterocephalus alcina HSPEC171_10
Carterocephalus palaemon BN000223
Carterocephalus silvicola BN000224
IV
Piruna aea SRR330376
Piruna roeveri BN005370
Piruna pirus LEP22847
Piruna penaea TAMIC071_10
Piruna ceracates BN005371
Piruna polingii ABLCU540_09
V
70/70
Freemaniana rawlinsi BN003731
Dardarina dardaris SRR7174550
Dardarina salta BN005144
Dardarina aspila BN004729
Ladda simplicis BN005137
Ladda carnis HSPEE063_12
Ladda caenides HSPEE051_12
Ladda ticidas HSPEE081_12
Ladda quadristriga SRR9330377
Ladda rubia HSPED830_12
Ladda crithote HSPED740_12
Ladda decca HSPED723_12
Ladda eburones BN007144
Ladda ibhara HSPEE095_12
VI
Dalla superior BN005139
Dalla costala SRR9330375
Dalla agathocles BN005140
Dalla riza HSPEE021_12
Dalla sepia HSPEE092_12
100/81
Dalla epiphaneus HSPEE019_12
Dalla caicus HSPEE033_12
Dalla dognini HSPEE048_12
Dalla dimidiatus BN005138*
Dalla miser HSPEE031_12
Dalla vinca HSPEE073_12
Dalla mesoxantha BN000303
Dalla scylla HSPEE053_12
Dalla oxaites HSPEE058_12
Dalla frater HSPEE054_12
Dalla merida HSPEE064_12
Dalla cyprius SRR9330372
Dalla cypselus HSPEE077_12
Dalla genes HSPEE056_12
0.03
AHE data (383 loci kit)
AHE data (13 loci kit)
Barcode Data
Whole Genome Data
SH-aLRT > 80 and UFBoot > 95
SH-aLRT > 80 or UFBoot > 95
SH-aLRT < 80 and UFBoot < 95

FIGURE 1 Maximum likelihood phylogenomic hypothesis for the subfamily Heteropterinae. Best-scoring maximum likelihood tree topology derived from the IQ-TREE analyses of the full nucleotide data set. Nodal support values expressed as ultrafast bootstrap (UFBoot) and SH-aLRT tests (SH-aLRT) are given following the inserted caption. Photographs of skippers*in natura*are presented:*Butleria flavomaculatus*(Credit: Roberto Herrera Pellizzari),*Heteropterus morpheus*(Credit: Darius Baužys),*Leptalina unicolor*(Credit: Zorac & Visar),*Metisella metis*(Credit: Andrew Massyn),*Carterocephalus palaemon*(Credit: Emmanuel Toussaint),*Piruna aea*(Credit: Alan Schmierer),*Dardarina dardaris*(Credit: Les Catchick),*Ladda ibhara*(Credit: Les Catchick) and*Dalla dimidiatus*(Credit: Les Catchick) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Results of the BEAST analyses using different calibration schemes, clock partitioning schemes and tree models

Analysis	Clocks	Tree Model	SS MLE	PS MLE	Crown Heteropterinae
A1	2	Yule	−423,928.905	−423,930.381	41.039 [34.990–46.962]
A2	2	Birth–death	−423,928.107	−423,929.875	40.361 [34.649–46.299]
A3	11	Yule	−424,175.575	−424,177.202	36.529 [32.493–40.847]
A4	11	Birth–death	−424,175.098	−424,177.159	36.650 [32.373–40.846]

Notes: SS, stepping-stone sampling marginal likelihood estimation; path-sampling marginal likelihood estimation; median post-burn-in divergence times in millions of years (95% credibility interval).

UFBoot = 100). We therefore believe that the placement reported in our study is likely more accurate than the one of Zhang et al. (2020).

Within Heteropterinae, we recover all genera as monophyletic with strong support and placed in six main clades (Figure 1), largely in line with the recent study of Cong et al. (2019). The genus *Argopteron* is sister to *Butleria* (SH-aLRT = 100/UFBoot = 100), forming clade I that is in turn recovered as sister to the remainder of Heteropterinae genera with strong support (SH-aLRT = 100/UFBoot = 100). In clade II, the two monotypic genera *Heteropterus* and *Leptalina* are recovered as sister (SH-aLRT = 100/UFBoot = 100). In clade III, the genus *Carterocephalus* (SH-aLRT = 100/UFBoot = 100) is sister to the African genera *Hovala*, *Metisella* and *Willema* (SH-aLRT = 100/UFBoot = 100). The genus *Piruna* (clade IV, SH-aLRT = 100/UFBoot = 100) is recovered as sister to clades V and VI with strong support (SH-aLRT = 100/UFBoot = 100). In clade V, the genus *Freemaniana* is recovered as sister to *Dardarina* (SH-aLRT = 100/UFBoot = 100) with reduced nodal support (SH-aLRT = 70/UFBoot = 70). Finally, in clade VI, the genera *Dalla* and *Ladda* are recovered as sister with strong support (SH-aLRT = 100/UFBoot = 96). Our results agree with the recent study of Cong et al. (2019), which described two new Heteropterinae genera. Our exon capture method is compatible with their whole-genome sequencing data set as exemplified by the inclusion of some of their specimens in this study. Our results further demonstrate the power of an integrative approach combining exon capture, whole-genome sequencing, and Sanger sequencing data.

3.2 | Evolutionary history of Heteropterinae skippers

The BEAST runs all converged with high ESS values and recovered similar age estimates with broadly overlapping credibility intervals for the crown of Heteropterinae regardless of the tree model (Yule or birth–death). Clock partitioning (two vs. 11 uncorrelated lognormal relaxed clocks) had the strongest impact on downstream estimates; more clocks inferred younger ages (Table 1). The best-fit analysis had two uncorrelated lognormal relaxed clocks and a birth–death model, and the resulting median ages of this analysis are presented in Figure 2.

Our estimates for the root and crown of Hesperiidae at, respectively, ca. 106 Ma (95% CI = 87–127 Ma) and ca. 77 Ma (95% CI = 67–88 Ma) unsurprisingly overlap with the secondary calibrations from Espeland et al. (2018) because these nodes were calibrated with informative priors derived from this study. Our study suggests new estimates for the crown age of Hedylidae at ca. 49 Ma (95% CI = 37–63 Ma), providing a first estimate for the divergence times of this family since *Macrosoma tipulata* (the species sister to the remainder of the family, see Kawahara et al., 2019) was included in the taxon sampling. We also estimate the crown of Heteropterinae at ca. 40 Ma (95% CI = 35–46 Ma). Our results indicate that most Heteropterinae genera diversified in the Oligocene to early Miocene, while intrageneric diversification mostly took place from the Miocene up to the Pleistocene (Figure 2).

Results of the BioGeoBEARS analyses are summarized in Table 2. Models M1 implementing paleogeographically

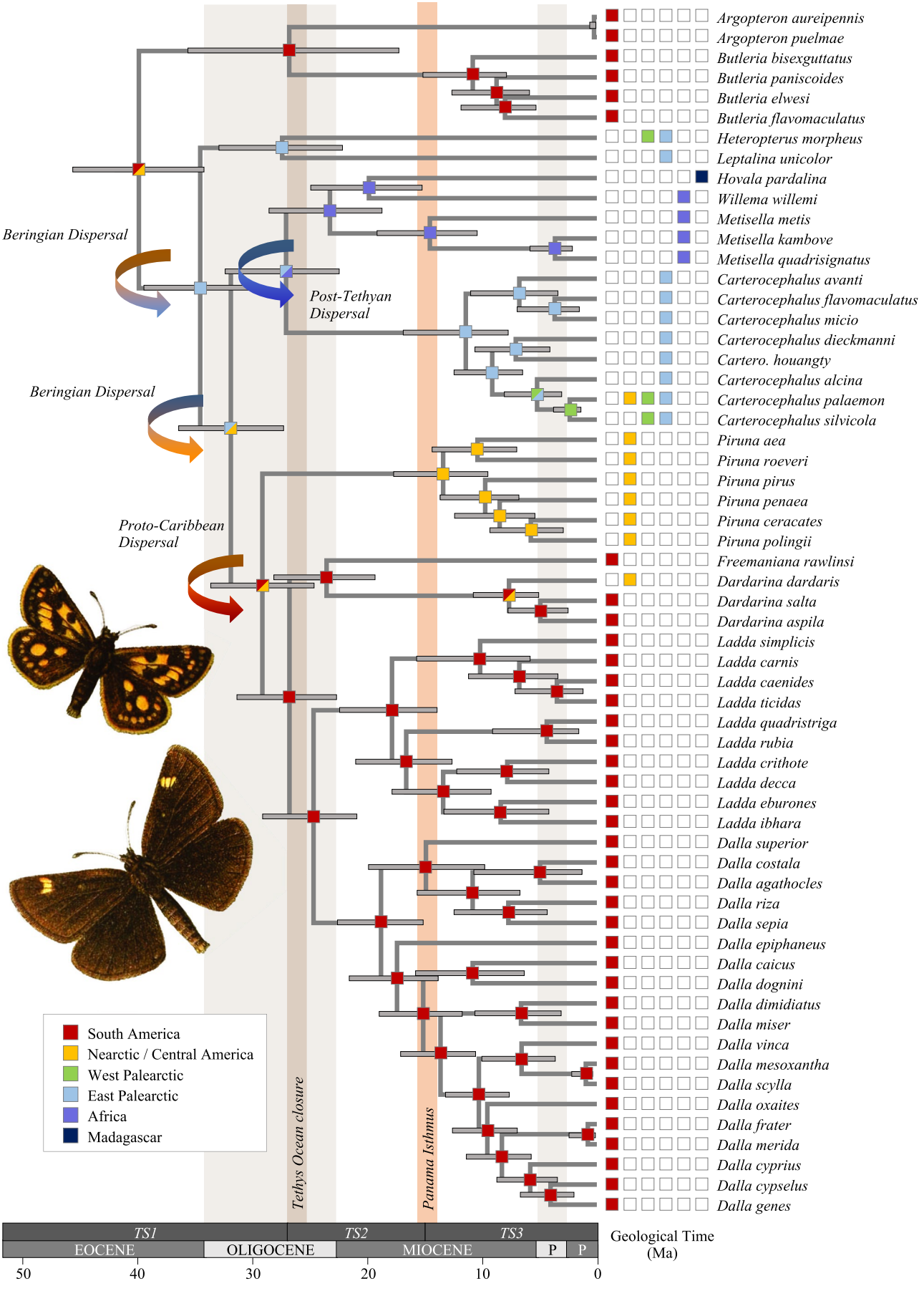
Argopteron aureipennis
Argopteron puelmae
Butleria bisexguttatus
Butleria paniscoides
Butleria elwesi
Butleria flavomaculatus
Heteropterus morpheus
Leptalina unicolor
Hovala pardalina
Willema willemi
Metisella metis
Metisella kambove
Metisella quadrisignatus
Carterocephalus avanti
Carterocephalus flavomaculatus
Carterocephalus micio
Carterocephalus dieckmanni
Cartero. houangty
Carterocephalus alcina
Carterocephalus palaemon
Carterocephalus silvicola
Piruna aea
Piruna roeveri
Piruna pirus
Piruna penaea
Piruna ceracates
Piruna polingii
Freemaniana rawlinsi
Dardarina dardaris
Dardarina salta
Dardarina aspila
Ladda simplicis
Ladda carnis
Ladda caenides
Ladda ticidas
Ladda quadristriga
Ladda rubia
Ladda crithote
Ladda decca
Ladda eburones
Ladda ibhara
Dalla superior
Dalla costala
Dalla agathocles
Dalla riza
Dalla sepia
Dalla epiphaneus
Dalla caicus
Dalla dognini
Dalla dimidiatus
Dalla miser
Dalla vinca
Dalla mesoxantha
Dalla scylla
Dalla oxaites
Dalla frater
Dalla merida
Dalla cyprius
Dalla cypselus
Dalla genes
Beringian Dispersal
Post-Tethyan Dispersal
Beringian Dispersal
Proto-Caribbean Dispersal
South America
Nearctic / Central America
West Palearctic
East Palearctic
Africa
Madagascar
Tethys Ocean closure
Panama Isthmus
TS1
TS2
TS3
EOCENE
OLIGOCENE
MIOCENE
P
P
50
40
30
20
10
0
Geological Time (Ma)

FIGURE 2 Historical biogeography of Heteropterinae skippers. Chronogram derived from the preferred BEAST analysis with outgroups removed. The most likely ancestral range estimated in BioGeoBEARS under the DIVALIKE model M1 is given at each node. Illustrations of*Heteropterus morpheus*and*Carterocephalus palaemon*are presented from*Die Schmetterlinge Deutschlands mit besonderer Berücksichtigung ihrer Biologie, Bd. 1–4,*by K. Eckstein (1913–1923) Stuttgart, Germany [Colour figure can be viewed at wileyonlinelibrary.com]

informed adjacency and rate scalar matrices in a time-stratified framework received better support than null models M0 under both DEC and DIVALIKE. The best-fit model according to AICc comparisons is the M1 DIVALIKE, and we present the results of this model in Figure 2. Under this model, the ancestral range of Heteropterinae skippers was in the New World (Nearctic + Central America + South America). From there, ancestors colonized the eastern Palearctic likely through a Beringian route ca. 35–40 Ma, before back-colonizing the Nearctic and Central American regions ca. 30–35 Ma. Similar patterns exist in other lineages such as snakes (Guo et al., 2012), and recent fossil discoveries from the Paleocene–Eocene boundary indicate that the Beringian route was likely important for insect dispersal in the Cenozoic (Garrouste & Nel, 2019). It appears that ancestors of Heteropterinae were highly mobile and took advantage of ephemeral connections between eastern Palearctic and Nearctic regions to disperse from one continent to the other in the Eocene and Oligocene (Figure 2). This pattern is in line with other studies suggesting the use of such routes in butterflies (*e.g*. Condamine et al., 2013).

We estimate that ancestors of *Dalla*, *Dardarina*, *Freemaniana*, *Ladda* and *Piruna* colonized South America ca. 30 Ma. This implies that these butterflies crossed the seaway separating North and South America at the time. The closure of the proto-Caribbean seaway is dated from ca. 15 Ma (Montes et al. 2015), but some archipelagic setting may have existed before, facilitating overwater dispersal via island "stepping-stones" (Toussaint et al., 2019). We surmise that a vicariant event separated *Dalla*, *Dardarina*, *Freemaniana* and *Ladda* (Clades V and VI) and the Central American genus *Piruna* (Clade IV), even though some species of *Dalla* and *Dardarina* are currently found in Central America. These latter lineages, while absent from our phylogeny, are likely recent colonizers that dispersed north via the Isthmus of Panama. The genus *Freemaniana* is placed for the first time in a phylogenetic framework, and its placement suggests that it separated from *Dardarina* ca. 25 Ma (Figure 2). Further diversification processes in this lineage may be obscured by a lack of sampling in key areas of Ecuador (Warren, 2001). The genera *Dalla* and *Ladda* are strongly associated with montane habitats (Warren, 2001). The role of Andean orogeny in the diversification of this large clade of two genera will be possible with more complete sampling and fine-scale understanding of species ranges.

Colonization of the Eastern Palearctic was followed by a split between *Heteropterus morpheus* and *Leptalina unicolor*, two monotypic genera that do not appear to have further diversified despite an initial split ca. 25 Ma. The long branches connecting these two lineages may reflect extinction processes during the Miocene and more recent periods of the Neogene. Since the Palearctic region is well known, it appears unlikely that additional undescribed species of these two genera could exist.

The Afrotropics were colonized once by range expansion from Eastern Palearctic by the ancestor of *Carterocephalus*, *Hovala*, *Metisella* and *Willema* ca. 25 Ma (Figure 2). This range expansion event coincides with the closure of the Tethys ocean and might reflect enhanced connectivity between Eastern Palearctic and the Arabian Peninsula through the formation of the *Gomphotherium* land bridge (Pirouz et al., 2017; Rögl, 1998). This geologic event was already inferred to impact ancient Asian-African dispersal in the skipper groups Baorini (Toussaint et al., 2019) and Coeliadinae (Toussaint et al. 2020) and is known to have shaped the biogeographic history of other butterfly lineages (Aduse-Poku et al., 2009, 2015; Kaliszewska et al., 2015). Colonization of Madagascar appears to be recent, although the lack of a crown age for *Hovala* precludes a better understanding of this timeline.

Within the predominantly East Palearctic genus *Carterocephalus*, we infer a colonization of West Palearctic in

TABLE 2 Results of the BioGeoBEARS analyses

Model	Adj.	Scal.	Pa.	d	e	LnL	AICc	ΔAICc	Node I	Node II	Node III
DEC M0	No	No	2	0.0017	0.001	−51.17	106.55	3.76	SNE	SNE	SNE
DIVALIKE M0	No	No	2	0.0023	0	−51.90	108.01	5.22	SE	E	NE
DEC M1	Yes	Yes	2	0.0066	0.001	−50.76	105.73	2.94	SN	NE	NE
DIVALIKE M1	Yes	Yes	2	0.0079	0	−49.29	102.79	-	SN	E	NE

Note: Adj., presence of absence of designed adjacency matrices; AIC, AICc, Akaike information criterion corrected (takes into account sample size); ΔAIC, difference in AIC between competing models; d, dispersal; e, extinction, LnL, log-likelihood; Pa., number of free parameters; Scal., presence or absence of designed rate scaler matrices.

the late Miocene to early Pliocene. This region only harbours two extant species, one of which, *C. palaemon* has extended its range to the northern Nearctic region likely during the Pleistocene. This species likely followed a De Geer or Thulean route to cross the seaways separating West Palearctic and Nearctic regions (Brikiatis, 2014; Condamine et al., 2013).

This study provides a robust dated phylogenomic framework for Heteropterinae and advances understanding of this widespread skipper butterfly lineage. Our results indicate that Heteropterinae skippers likely took advantage of geological reconfigurations over the past 40 million years to colonize most of the planet except for the Australian region where they are replaced by the grass skipper subfamily Trapezitinae, whose evolutionary history remains unknown. This study further demonstrates that combining exon capture and Sanger data will allow a deeper understanding of skipper evolution.

ACKNOWLEDGEMENTS

We are grateful to Ernst Brockmann for sharing unreleased DNA barcodes. We thank Marianne Espeland and Roger Vila for providing specimens, the photographers whose images were used in this study, and the members of the Kawahara Lab and curators of the McGuire Center for Lepidoptera and Biodiversity for their assistance. The University of Florida's HiPerGator cluster provided computational resources and technical support. This research was funded by NSF DEB-1541500 to AYK and NSF DEB-1541557 to DJL.

ORCID

Emmanuel F. A. Toussaint https://orcid.org/0000-0002-8439-1285
Emily A. Ellis https://orcid.org/0000-0003-3343-0923
Andrew D. Warren https://orcid.org/0000-002-6730-1938
Caroline Storer https://orcid.org/0000-0002-0349-0653
David J. Lohman https://orcid.org/0000-0002-0689-2906
Akito Y. Kawahara https://orcid.org/0000-0002-3724-4610

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. https://doi.org/10.1186/s12862-015-0449-3

Aduse-Poku, K., Vingerhoedt, E., & Wahlberg, N. (2009). Out-of-Africa again: A phylogenetic hypothesis of the genus *Charaxes* (Lepidoptera: Nymphalidae) based on five gene regions. *Molecular Phylogenetics and Evolution*, *53*, 463–478. https://doi.org/10.1016/j.ympev.2009.06.021

Allio, R., Scornavacca, C., Nabholz, B., Clamens, A. L., Sperling, F. A., & Condamine, F. L. (2020). Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail butterfly evolution. *Systematic Biology*, *69*(1), 38–60. https://doi.org/10.1093/sysbio/syz030

Bacon, C. D., Silvestro, D., Jaramillo, C., Smith, B. T., Chakrabarty, P., & Antonelli, A. (2015). Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences*, *112*(19), 6110–6115. https://doi.org/10.1073/pnas.1423853112

Baele, G., Li, W. L. S., Drummond, A. J., Suchard, M. A., & Lemey, P. (2012). Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, *30*(2), 239–243. https://doi.org/10.1093/molbev/mss243

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Breinholt, J. W., Earl, C., Lemmon, A. R., Lemmon, E. M., Xiao, L., & Kawahara, A. Y. (2018). Resolving relationships among the megadiverse butterflies and moths with a novel pipeline for anchored phylogenomics. *Systematic Biology*, *67*(1), 78–93. https://doi.org/10.1093/sysbio/syx048

Brikiatis, L. (2014). The De Geer, Thulean and Beringia routes: Key concepts for understanding early Cenozoic biogeography. *Journal of Biogeography*, *41*(6), 1036–1054. https://doi.org/10.1111/jbi.12310

Brown, J. W., Walker, J. F., & Smith, S. A. (2017). Phyx: Phylogenetic tools for unix. *Bioinformatics*, *33*(12), 1886–1888. https://doi.org/10.1093/bioinformatics/btx063

Chazot, N., Wahlberg, N., Freitas, A. V. L., Mitter, C., Labandeira, C., Sohn, J. C., Sahoo, R. K., Seraphim, N., de Jong, R., & Heikkilä, M. (2019). Priors and posteriors in Bayesian timing of divergence analyses: The age of butterflies revisited. *Systematic Biology*, *68*(5), 797–813. https://doi.org/10.1093/sysbio/syz002

Cock, M. J. W., & Congdon, T. C. E. (2017). Observations on the biology of Afrotropical Hesperiidae (Lepidoptera) with particular reference to Kenya. Part 11. Heteropterinae. *Zootaxa*, *4226*(4), 487–508. https://doi.org/10.11646/zootaxa.4226.4.3.

Condamine, F. L., Sperling, F. A., & Kergoat, G. J. (2013). Global biogeographical pattern of swallowtail diversification demonstrates alternative colonization routes in the Northern and Southern hemispheres. *Journal of Biogeography*, *40*(1), 9–23. https://doi.org/10.1111/j.1365-2699.2012.02787.x

Cong, Q., Zhang, J., Shen, J., & Grishin, N. V. (2019). Fifty new genera of Hesperiidae (Lepidoptera). *Insecta Mundi*, *0731*, 1–56.

Denton, J. F., Lugo-Martinez, J., Tucker, A. E., Schrider, D. R., Warren, W. C., & Hahn, M. W. (2014). Extensive error in the number of genes inferred from draft genome assemblies. *PLoS Computational Biology*, *10*(12), e1003998. https://doi.org/10.1371/journal.pcbi.1003998

Pennington, K.M., Dickson, C. G. C., & Kroon, D. M. (1978). *Pennington's Butterflies of Southern Africa*, Johannesburg: AD Donker.

Dowdy N. J., Keating S., Lemmon A.R., Lemmon E.M., Conner W.E., Scott Chialvo C.H., Weller S.J., Simmons R.B., Sisson M.S., Zaspel J.M. (2020). A deeper meaning for shallow-level phylogenomic studies: nested anchored hybrid enrichment offers great promise for resolving the tiger moth tree of life (Lepidoptera: Erebidae:

Arctiinae). *Systematic Entomology*, *45*(4), 874–893. http://dx.doi.org/10.1111/syen.12433.

Eckstein, K. (1913). *Die schmetterlinge Deutschlands mit besonderer berücksichtigung ihrer biologie*, Vol. *1*. KG Lutz'verlag.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Espeland, M., Breinholt, J., Willmott, K. R., Warren, A. D., Vila, R., Toussaint, E. F. A., Maunsell, S. C., Aduse-Poku, K., Talavera, G., Eastwood, R., Jarzyna, M. A., Guralnick, R., Lohman, D. J., Pierce, N. E., & Kawahara, A. Y. (2018). A comprehensive and dated phylogenomic analysis of butterflies. *Current Biology*, *28*(5), 770–778. https://doi.org/10.1016/j.cub.2018.01.061

Garrouste, R., & Nel, A. (2019). Alaskan Palaeogene insects: A challenge for a better knowledge of the Beringian 'route' (Odonata: Aeshnidae, Dysagrionidae). *Journal of Systematic Palaeontology*, *17*(22), 1939–1946. https://doi.org/10.1080/14772019.2019.1572235

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. https://doi.org/10.1093/sysbio/syq010.

Guo, P., Liu, Q., Xu, Y., Jiang, K., Hou, M., Ding, L., Pyron, R. A., & Burbrink, F. T. (2012). Out of Asia: Natricine snakes support the Cenozoic Beringian dispersal hypothesis. *Molecular Phylogenetics and Evolution*, *63*(3), 825–833. https://doi.org/10.1016/j.ympev.2012.02.021.

Hamilton, C. A., St Laurent, R. A., Dexter, K., Kitching, I. J., Breinholt, J. W., Zwick, A., Timmermans, M. J., Barber, J. R., & Kawahara, A. Y. (2019). Phylogenomics resolves major relationships and reveals significant diversification rate shifts in the evolution of silk moths and relatives. *BMC Evolutionary Biology*, *19*(1), 1–13. https://doi.org/10.1186/s12862-019-1505-1.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281.

Homziak, N. T., Breinholt, J. W., Branham, M. A., Storer, C. G., & Kawahara, A. Y. (2019). Anchored hybrid enrichment phylogenomics resolves the backbone of erebine moths. *Molecular Phylogenetics and Evolution*, *131*, 99–105. https://doi.org/10.1016/j.ympev.2018.10.038.

Igarashi, S., & Harada, M. (2015). *Sequel to "The Life Histories of Asian Butterflies, Vols. I & II"*, Tokyo: Roppon-Ashi Entomological Books.

Kaliszewska, Z. A., Lohman, D. J., Sommer, K., Adelson, G., Rand, D. B., Mathew, J., Talavera, G., & Pierce, N. E. (2015). When caterpillars attack: Biogeography and life history evolution of the Miletinae (Lepidoptera: Lycaenidae). *Evolution*, *69*, 571–588. https://doi.org/10.1111/evo.12599.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587. https://doi.org/10.1038/nmeth.4285.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010.

Kawahara, A. Y., & Breinholt, J. W. (2014). Phylogenomics provides strong evidence for relationships of butterflies and moths. *Proceedings of the Royal Society B*, *281*(1788), 20140970. https://doi.org/10.1098/rspb.2014.0970.

Kawahara, A. Y., Breinholt, J. W., Espeland, M., Storer, C., Plotkin, D., Dexter, K. M., Toussaint, E. F. A., St Laurent, R. A., Brehm, G., Vargas, S., Forero, D., Pierce, N. E., & Lohman, D. J. (2018). Phylogenetics of moth-like butterflies (Papilionoidea: Hedylidae) based on a new 13-locus target capture probe set. *Molecular Phylogenetics and Evolution*, *127*, 600–605. https://doi.org/10.1016/j.ympev.2018.06.002

Kawahara, A. Y., Plotkin, D., Espeland, M., Meusemann, K., Toussaint, E. F. A., Donath, A., Gimnich, F., Frandsen, P. B., Zwick, A., dos Reis, M., Barber, J. R., Peters, R. S., Liu, S., Zhou, X., Mayer, C., Podsiadlowski, L., Storer, C., Yack, J. E., Misof, B., & Breinholt, J. W. (2019). Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proceedings of the National Academy of Sciences*, *116*(45), 22657–22663. https://doi.org/10.1073/pnas.1907847116

Kim, S.-S., & Ho, Y.-H. (2012). *Life histories of Korean butterflies*. Seoul: Publishing Sagyejel.

Kocot, K. M., Citarella, M. R., Moroz, L. L., & Halanych, K. M. (2013). PhyloTreePruner: A phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. *Evolutionary Bioinformatics*, *9*. https://doi.org/10.4137/EBO.S12813

Kück, P., & Meusemann, K. (2010). FASconCAT: Convenient handling of data matrices. *Molecular Phylogenetics and Evolution*, *56*(3), 1115–1118. https://doi.org/10.1016/j.ympev.2010.04.024

Kück, P., Meusemann, K., Dambach, J., Thormann, B., von Reumont, B. M., Wägele, J. W., & Misof, B. (2010). Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Frontiers in Zoology*, *7*(1), 10. https://doi.org/10.1186/1742-9994-7-10

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*(3), 772–773. https://doi.org/10.1093/molbev/msw260

Larsen, T. B. (2005). *Butterflies of West Africa*, Stenstrup: Apollo Books.

Li, W., Cong, Q., Shen, J., Zhang, J., Hallwachs, W., Janzen, D. H., & Grishin, N. V. (2019). Genomes of skipper butterflies reveal extensive convergence of wing patterns. *Proceedings of the National Academy of Sciences*, *116*(13), 6232–6237. https://doi.org/10.1073/pnas.1821304116

Matzke, N. J. (2018). BioGeoBEARS: BioGeography with Bayesian (and likelihood) Evolutionary Analysis with R Scripts. version 1.1.1, published on GitHub on November 6, (2018). https://doi.org/10.5281/zenodo.1478250

Minh, B. Q., Nguyen, M. A. T., & von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Molecular Biology and Evolution*, *30*(5), 1188–1195. https://doi.org/10.1093/molbev/mst024

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J. C., Valencia, V., Ayala, C., Pérez-Angel, L. C., Rodriguez-Parra, L. A., Ramirez, V., & Niño, H. (2015). Middle Miocene closure of the Central American seaway. *Science*, *348*(6231), 226–229. https://doi.org/10.1126/science.aaa2815.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Peña G., L.E., & Ugarte P., A. J. (1996). *The Butterflies of Chile*, Santiago: Editorial Universitaria.

Pirouz, M., Avouac, J. P., Hassanzadeh, J., Kirschvink, J. L., & Bahroudi, A. (2017). Early Neogene foreland of the Zagros, implications for the initial closure of the Neo-Tethys and kinematics of crustal shortening. *Earth and Planetary Science Letters*, *477*, 168–182. https://doi.org/10.1016/j.epsl.2017.07.046

Price, M. N., Dehal, P. S., & Arkin, A. P. (2009). FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, *26*(7), 1641–1650. https://doi.org/10.1093/molbev/msp077

Ree, R. H., & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, *57*(1), 4–14. https://doi.org/10.1080/10635150701883881

Rögl, F. (1998). Palaeogeographic considerations for Mediterranean and *Paratethys ceaways* (Oligocene to Miocene). *Annals of the Natural History Museum in Vienna*, *99*, 279–310.

Sahoo, R. K., Warren, A. D., Collins, S. C., & Kodandaramaiah, U. (2017). Hostplant change and paleoclimatic events explain diversification shifts in skipper butterflies (Family: Hesperiidae). *BMC Evolutionary Biology*, *17*(1), 174. https://doi.org/10.1186/s12862-017-1016-x

Sahoo, R. K., Warren, A. D., Wahlberg, N., Brower, A. V., Lukhtanov, V. A., & Kodandaramaiah, U. (2016). Ten genes and two topologies: An exploration of higher relationships in skipper butterflies (Hesperiidae). *PeerJ*, *4*, e2653. https://doi.org/10.7717/peerj.2653

Seton, M., Müller, R. D., Zahirovic, S., Gaina, C., Torsvik, T., Shephard, G., Talsma, A., Gurnis, M., Turner, M., Maus, S., & Chandler, M. (2012). Global continental and ocean basin reconstructions since 200 Ma. *Earth-Science Reviews*, *113*(3–4), 212–270.

Smith, S. A., Brown, J. W., & Walker, J. F. (2018). So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *PLoS One*, *13*(5), e0197433. https://doi.org/10.1371/journal.pone.0197433

St Laurent, R. A., Hamilton, C. A., & Kawahara, A. Y. (2018). Museum specimens provide phylogenomic data to resolve relationships of sack-bearer moths (Lepidoptera, Mimallonoidea, Mimallonidae). *Systematic Entomology*, *43*(4), 729–761. https://doi.org/10.1111/syen.12301

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*(1), vey016. https://doi.org/10.1093/ve/vey016

Toussaint, E. F. A., Breinholt, J. W., Earl, C., Warren, A. D., Brower, A. V., Yago, M., Dexter, K. M., Espeland, M., Pierce, N. E., Lohman, D. J., & Kawahara, A. Y. (2018). Anchored phylogenomics illuminates the skipper butterfly tree of life. *BMC Evolutionary Biology*, *18*(1), 101. https://doi.org/10.1186/s12862-018-1216-z

Toussaint, E. F.A., Chiba H., Yago M., Dexter K. M., Warren A. D., Storer C., Lohman D. J., Kawahara A. Y. (2020). Afrotropics on the wing: Phylogenomics and historical biogeography of awl and policeman skippers. *Systematic Entomology*, http://dx.doi.org/10.1111/syen.12455

Toussaint, E. F. A., Vila, R., Yago, M., Chiba, H., Warren, A. D., Aduse-Poku, K., Storer, C., Dexter, K. M., Maruyama, K., Lohman, D. J., & Kawahara, A. Y. (2019). Out of the Orient: Post-Tethyan transoceanic and trans-Arabian routes fostered the spread of Baorini skippers in the Afrotropics. *Systematic Entomology*, *44*(4), 926–938. https://doi.org/10.1111/syen.12365

Toussaint, E. F. A., & Warren, A. D. (2019). A review of red-eye pigmentation and diel activity patterns in skippers (Lepidoptera, Papilionoidea, Hesperiidae). *Journal of Natural History*, *53*(35–36), 2165–2181. https://doi.org/10.1080/00222933.2019.1692090

Warren, A. D. (2001). A new genus and species of Cyclopidinae from Zamora, Ecuador (Lepidoptera: Hesperiidae). *Boletın Cientıfico, Museo De Historia Natural, Universidad De Caldas*, *5*, 138–153.

Warren, A. D., Ogawa, J. R., & Brower, A. V. (2008). Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperiidae (Lepidoptera: Hesperioidea). *Cladistics*, *24*(5), 642–676. https://doi.org/10.1111/j.1096-0031.2008.00218.x

Warren, A. D., Ogawa, J. R., & Brower, A. V. (2009). Revised classification of the family Hesperiidae (Lepidoptera: Hesperioidea) based on combined molecular and morphological data. *Systematic Entomology*, *34*(3), 467–523. https://doi.org/10.1111/j.1365-3113.2008.00463.x

Williams, M. C. (2020). Afrotropical Butterflies Website. http://www.metamorphosis.org.za/?p=articles&s=List&pt=166. Lepidopterists' Society of Africa. Accessed 30 June (2020).

Zhang, J., Cong, Q., Shen, J., Brockmann, E., & Grishin, N. V. (2019). Three new subfamilies of skipper butterflies (Lepidoptera, Hesperiidae). *ZooKeys*, *861*, 91. https://doi.org/10.3897/zookeys.861.34686

Zhang, J., Cong, Q., Shen, J., Brockmann, E., & Grishin, N. V. (2019). Genomes reveal drastic and recurrent phenotypic divergence in firetip skipper butterflies (Hesperiidae: Pyrrhopyginae). *Proceedings of the Royal Society B*, *286*(1903), p. (2019) 0609. https://doi.org/10.1098/rspb.2019.0609

Zhang, J., Lees, D. C., Shen, J., Cong, Q., Huertas, B., Martin, G., & Grishin, N. V. (2020). The mitogenome of a Malagasy butterfly *Malaza fastuosus* (Mabille, 1884) recovered from the holotype collected over 140 years ago adds support for a new subfamily of Hesperiidae (Lepidoptera). *Genome*, *63*(4), 195–202. https://doi.org/10.1139/gen-2019-0189

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Toussaint EFA, Ellis EA, Gott RJ, et al. Historical biogeography of Heteropterinae skippers via Beringian and post-Tethyan corridors. *Zool Scr*. 2021;50:100–111. https://doi.org/10.1111/zsc.12457