Around the World in 26 Million Years: Diversification and Biogeography of Pantropical Grass-Yellow Eurema Butterflies (Pieridae: Coliadinae)

Article *in* Journal of Biogeography · February 2025

DOI: 10.1111/jbi.15107

22 authors, including:

Jing Vir Leong
Czech Academy of Sciences, Biology Centre
15 PUBLICATIONS **183** CITATIONS
SEE PROFILE

Pável Matos-Maraví
Czech Academy of Sciences, Biology Centre
87 PUBLICATIONS **1,492** CITATIONS
SEE PROFILE

Rayner Núñez
Research Museum Alexander Koenig
62 PUBLICATIONS **428** CITATIONS
SEE PROFILE

Renato Nunes
The Graduate Center, CUNY
10 PUBLICATIONS **57** CITATIONS
SEE PROFILE

Journal of Biogeography

WILEY

Journal of Biogeography

RESEARCH ARTICLE OPEN ACCESS

Around the World in 26 Million Years: Diversification and Biogeography of Pantropical Grass-Yellow *Eurema* Butterflies (Pieridae: Coliadinae)

Jing V. Leong[1,2,3] | Pável Matos-Maraví[2] | Rayner Núñez[4,5] | Renato Nunes[1,6] | Weijun Liang[1,7] | Michael F. Braby[8,9] | Tenzing Doleck[1,6] | Kwaku Aduse-Poku[1,10] | Yutaka Inayoshi[11] | Yu-Feng Hsu[12] | Niklas Wahlberg[13] | Djunijanti Peggie[14] | Alma B. Mohagan[15,16] | Dave P. Mohagan[15,16] | Julio A. Genaro[17] | Antonio R. Perez-Asso[18] | Krushnamegh Kunte[19] | Dino J. Martins[20,21] | Szabolcs Sáfián[22,23] | Akito Y. Kawahara[24,25] | Naomi E. Pierce[26] | David J. Lohman[1,6,27]

[1]Department of Biology, City College of New York, New York City, New York, USA | [2]Biology Centre of the Czech Academy of Sciences, Ceske Budejovice, Czech Republic | [3]Department of Zoology, Faculty of Science, University of South Bohemia, Ceske Budejovice, Czech Republic | [4]SNSB–Zoologische Staatssammlung München, Lepidoptera Section, Munich, Germany | [5]Lepidoptera and Trichoptera Section, Leibniz Institute for the Analysis of Biodiversity Change – Museum Koenig, Bonn, Germany | [6]Ph.D. Program in Biology, Graduate Center, City University of New York, New York City, New York, USA | [7]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut, USA | [8]Division of Ecology and Evolution, Research School of Biology, The Australian National University, Acton, Australian Capital Territory, Australia | [9]The Australian National Insect Collection, National Research Collections Australia, Canberra, Australian Capital Territory, Australia | [10]Department of Biology, Howard University, Washington, DC, USA | [11]Sritana Condominium 2, Chiang Mai, Thailand | [12]Department of Life Science, National Taiwan Normal University, Taipei, Taiwan | [13]Department of Biology, Lund University, Lund, Sweden | [14]Museum Zoologicum Bogoriense, Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Bogor, Indonesia | [15]Center for Biodiversity Research and Extension in Mindanao, Central Mindanao University, Maramag, Philippines | [16]College of Arts and Sciences, Institute of Biological Sciences, Central Mindanao University, Maramag, Philippines | [17]Division of Plant Industry, Florida State Collection of Arthropods, Gainesville, Florida, USA | [18]National Museum of Natural History, Santo Domingo, Dominican Republic | [19]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru, India | [20]Turkana Basin Institute, Stony Brook University, Stony Brook, New York, USA | [21]Insect Committee of Nature Kenya, the East Africa Natural History Society, Nairobi, Kenya | [22]Institute of Silviculture and Forest Protection, University of Sopron, Sopron, Hungary | [23]African Butterfly Research Institute, Nairobi, Kenya | [24]McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, University of Florida, Gainesville, Florida, USA | [25]Entomology and Nematology Department, University of Florida, Gainesville, Florida, USA | [26]Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts, USA | [27]Entomology Section, National Museum of Natural History, Manila, Philippines

Correspondence: Jing V. Leong (jing.leong@entu.cas.cz) | David J. Lohman (lohman-dlohman@ccny.cuny.edu; djlohman@gmail.com)

Received: 27 March 2024 | **Revised:** 30 January 2025 | **Accepted:** 31 January 2025

Funding: This work was supported by National Geographic grants 9285-13 and WW-227R-17 to D.J.L.; a Fulbright ASEAN Research Award to D.J.L.; Czech Science Foundation GAČR grant 22-35084J to P.M.-M.; National Science Foundation grants DEB-1541500 to A.Y.K., DEB-1541557 to D.J.L. and DEB-1541560 to N.E.P.; Alexander von Humboldt Foundation grant 1162549-CUB-GFHERMES-E to Rayner Núñez, and Ramanujan Fellowship from the Department of Science and Technology, Government of India, and a research grant from NCBS to K.K., which funded fieldwork and sequencing in India.

Keywords: Beringia | biogeography | diversification | Papilionoidea | phylogenomics | systematics

ABSTRACT

Aim: Grass-yellow butterflies (*Eurema*) are a group of pantropical Pieridae distributed throughout Asia, Australasia, Africa and the New World. However, little is known about their diversification, including the biogeographic mechanism(s) explaining their circumglobal distribution. We present the first densely sampled, time-calibrated phylogeny and biogeographic reconstruction of grass-yellows to confirm the monophyly of the genera, re-evaluate their taxonomy and infer the biogeographic events contributing to their worldwide distribution.

Location: Global tropics and subtropics.
Taxon: The butterfly tribe Euremini (Pieridae: Coliadinae).
Methods: We analysed up to 391 genetic loci from 126 samples of 66 ingroup species. Divergence dating was accomplished in a Bayesian phylogenetic framework using secondary calibration points, and maximum likelihood models of various biogeographic models were fitted to the data using the R package BioGeoBEARS. We used the best fitting model to estimate relative dispersal events with biogeographical stochastic mapping. Finally, we estimated branch-specific speciation and extinction rates to assess the diversification history of the group.
Results: Different phylogenomic analyses converged on similar topologies with robust support. Grass-yellows emerged *ca*. 26 Mya in the New World, and a single extant lineage dispersed to Asia in the early Miocene, where they diversified and dispersed to Africa and Australasia. The fastest rates of diversification occurred in the Old World tropics during the late Miocene. Many of the grass-yellow genera were either paraphyletic or polyphyletic as traditionally circumscribed. To maintain nomenclatural stability, we place all grass-yellows in *Eurema sensu lato* and recognise two subgenera: *Eurema* (*Abaeis*) and *Eurema* (*Eurema*).
Main Conclusions: Grass-yellow butterflies originated in the Americas and attained their global distributional patterns via dispersal. The Indo-Australian and Caribbean archipelagoes seem to have accelerated the diversification of the group, and movement in and out of these island regions was frequent. Although the traditional view of founder-event speciation envisions migrants from large landmasses ('mainland') colonising smaller landmasses ('islands'), we find that island to mainland dispersal and differentiation were equally or more common.

1 | Introduction

Cosmopolitan and pantropical taxa attain their circumglobal distributions by ancient vicariance, recent dispersal, or some combination of the two (Lin et al. 2024). The relative importance of vicariance versus dispersal in shaping biodiversity distribution patterns remains contentious (Heads et al. 2023; Zink et al. 2000), and there are many proposed hypotheses to explain intercontinental disjunctions. Continental drift and other dynamic geologic factors explain the biogeographic history and diversification of some widespread taxa across the Old World and New World. For instance, ancient vicariance caused by the fragmentation of Gondwana has strongly influenced the distributions of many animal and plant groups in the southern hemisphere, including geckos (Gamble et al. 2007), marsupials (Mitchell et al. 2014) and gymnosperms (Wang and Ran 2014).

In contrast, dispersal—movement that potentially leads to gene flow—is a common response to environmental change, including habitat loss, habitat fragmentation and climate change (Chen et al. 2011), and can lead to speciation (Mayr 1954). Dispersal through Beringia contributed to widespread distributions in both the Old and New Worlds for many taxa, including the dispersal of modern humans around 15,000–30,000 years ago, a period when *Homo sapiens* walked from Asia to the Americas (Goebel et al. 2008). From the Late Cretaceous to the Neogene, Beringia was an important biogeographic route and dispersal filter for many terrestrial taxa that are unlikely to disperse over water, including butterflies in the taxa Riodinidae, Limenitidinae and Polyommatini (Espeland et al. 2015; Tseng et al. 2022; Vila et al. 2011).

Species-rich taxa with cosmopolitan or pantropical distributions enable testing hypotheses related to dispersal, vicariance and diversification. Grass-yellow butterflies, 71 species that we place in the genus *Eurema* (Pieridae: Coliadinae: Euremini), can serve as a model for understanding biogeographic processes. These butterflies are common and ubiquitous inhabitants of degraded forests and other open habitats in nearly every tropical and subtropical areas on Earth (Yata and Morishita 1985). Some species are widespread across open habitats in the North American, South American, African, Asian, or Australian following tropics/subtropics, whereas others are island endemics in the Caribbean archipelago and Indo-Australian Archipelago (IAA). The monophyly of sulphur butterflies (Pieridae: Coliadinae), which includes grass-yellows, is well established (Braby et al. 2006; Carvalho et al. 2024; Espeland et al. 2018; Wahlberg et al. 2014). The subfamily includes ~300 described species with larvae that feed predominantly on legumes (Fabaceae), but species-level systematics and taxonomy within individual genera have been problematic because some genera are para- or polyphyletic (Braby 2005; Braby et al. 2006; Carvalho et al. 2024).

The taxonomy of *Eurema* and taxa historically allied to it, *Abaeis*, *Leucidia*, *Pyrisitia*, *Terias*, and *Teriocolias*, have been unstable because of phenotypic similarity among species, because of few synapomorphies and because the group lacks a densely sampled phylogeny. *Terias*, for example, was originally described as a genus by Swainson, then regarded as a subgenus by Yata (1989), before being reinstated as a genus by Zhang et al. (2023) with no reappraisal of diagnostic features. We review the turbulent nomenclatural history of these taxa and revise the taxonomy in light of our results.

Currently, no study has attempted to resolve species-level relationships or infer the evolution and biogeographic history of *Eurema* on a global scale, which requires a well-supported phylogenetic framework. Yata (1989) hypothesised that the ancestor of *Eurema* lived in the Old World tropics, then dispersed and rapidly radiated in the New World. Yata's logic was based on the then-current knowledge of the distribution of these butterflies and Darlington's (1957) observation that more animal taxa have dispersed from the Old World to the New World than *vice versa*. Factors affecting the evolutionary success and pantropical distribution of this species-rich butterfly group are still poorly understood.

We present the first globally sampled, time-calibrated phylogenetic and biogeographic reconstruction of pantropical

grass-yellow butterflies—a group that has included the genera *Abaeis*, *Eurema*, *Leucidia*, *Pyrisitia*, *Terias* and *Teriocolias*. We aim to assess the monophyly of these taxa and of the widespread species *E. hecabe* while testing Yata's (1989) biogeographic prediction that grass-yellow butterflies originated in the Old World and diversified via vicariance. Specifically, we address the following questions: (1) Did grass-yellow butterflies originate in the New World or the Old World? (2) What roles have vicariance and dispersal played in shaping their diversification and pantropical distribution? and (3) Does their taxonomy need to be revised in light of our results so that genera and species are monophyletic?

2 | Methods

2.1 | Taxon Sampling and Dataset Assembly

Specimens were identified by the authors of this paper, often with reference to relevant morphological keys (Klots 1928, 1929; Yata 1989, 1991, 1992, 1994, 1995) and verified using DNA barcodes. We generally followed Lamas' (2015) taxonomy at the outset of this study and consider the following former genera to be grass-yellow butterflies (the number of sampled and described species prior to our taxonomic changes are provided in parentheses): *Abaeis* (1/2), *Eurema* (50/62), *Leucidia* (2/2), *Pyrisitia* (11/11) and *Teriocolias* (1/1). On the basis of prior taxonomic assignments using morphology, seven of the unsampled species are from clade A (Figure 1), one unsampled species is from clade B, one unsampled species is from clade C and five unsampled species are from clade D. As none of these rare species were morphologically unusual enough to merit a supraspecific taxon that we did not sample, it is unlikely that their inclusion would substantially change the topology of our tree or represent a major unsampled clade. Dataset assembly used a hierarchical approach, and most of the samples were sequenced via anchored hybrid enrichment (AHE), a technique that enables DNA to be sequenced reliably from dried museum specimens (Nunes et al. 2022). The BUTTERFLY 1.1 probe set, which captures one mitochondrial (cytochrome c oxidase subunit I, COI) and 390 nuclear loci (Toussaint et al. 2019), was used to capture sequences from 21 samples prior to sequencing on the Illumina platform. Then, 12 nuclear loci and COI were sequenced via AHE from another 83 samples using the BUTTERFLY 2.0 probe set (Kawahara et al. 2018). All 13 loci captured by the BUTTERFLY 2.0 probe set are found in the BUTTERFLY 1.1 probe set and include many loci amplified by the primers designed by Wahlberg and Wheat (2008), which have been widely used in butterfly molecular systematics. Up to eight of these Wahlberg and Wheat loci (CAD, COI, EF1a, GAPDH, IDH, MDH, RpS2 and RpS5) were PCR amplified and Sanger sequenced from 30 additional samples.

Prior to AHE, DNA was extracted with an OmniPrep DNA Extraction Kit (gbiosciences.com) from leg, thorax or abdominal tissue of specimens that were ethanol preserved, dried/papered, or spread/pinned. Extracts to be sequenced via AHE were submitted to RAPiD Genomics (rapid-genomics.com) for library preparation, hybrid enrichment and Illumina sequencing. To prepare each library, gDNA was first sheared into 300 bp fragments, then an adenine residue was ligated onto the blunt 3′ end of each fragment to allow ligation of barcode adapters and the library was amplified via PCR. Following library preparation, target enrichment and Illumina sequencing, raw reads were cleaned and assembled using the iterated bait assembly software pipeline for AHE phylogenomics (Breinholt et al. 2018).

Preliminary analyses only using samples with 391 or 13 loci captured via AHE resulted in trees with some important branches that were weakly supported or with radically different topologies in different analyses. To rectify this problem, we added additional samples with 391 loci to clades that lacked them by using publicly available short-read genome sequence data from NCBI. Extracting our protein-coding loci from short-read sequence data was accomplished with SEquence CApture PRocessor (SECAPR) (Andermann et al. 2018; Ribeiro et al. 2021). This program was used to extract the 391 loci captured by the BUTTERFLY 1.1 probe set from nine short-read archives on NCBI (ncbi.nlm.nih.gov/sra) following the recommended protocol https://github.com/AntonelliLab/seqcap_processor. This bioinformatic pipeline downloads raw FASTQ reads from NCBI using the SRA Toolkit (https://github.com/ncbi/sra-tools/wiki/01.-Downloading-SRA-Toolkit). FastQC checks for sequence quality (Andrews 2010) and acceptable reads are trimmed with Trimmomatic (Bolger et al. 2014) before *de novo* assembly with ABySS (Simpson et al. 2009) using a k-mer value of 35 (Compeau et al. 2011). The alignment program LASTZ (Harris 2007) was then used to extract our 391 target sequences from the assembly.

The final dataset had 126 ingroup samples (66 species) and 17 outgroup samples (17 genera). The number of loci per sample ranged between 1 and 376 (Table S1); simulation has demonstrated that large amounts of missing data do not adversely affect tree inference (Talavera et al. 2022). We then made a reduced dataset with a single individual per species except for the widespread species *E. hecabe*, which we suspected might include cryptic species (Bickford et al. 2007). Many species were represented by more than one sample in the full dataset, and we selected the specimen of each species with the most sequence data for inclusion in a reduced dataset, which had 69 ingroup and 17 outgroup samples (Table S1). We recognise the recent publication of Irungbam et al. (2023), which separated the *E. brigitta* complex into two species: *E. brigitta sensu stricto* from Africa and Madagascar, and *E. drona* from Asia and Australasia (the latter region includes Australia and Melanesia).

2.2 | Alignment, Model Selection and Phylogenetic Analyses

Sequences for each locus were aligned using MUSCLE (Edgar 2004) implemented in AliView 1.26 (Larsson 2014), which was also used to determine the codon positions of each locus by minimising stop codons. SequenceMatrix 1.9 (Vaidya et al. 2011) was used to concatenate aligned loci while retaining codon information. ModelFinder Pro (Kalyaanamoorthy et al. 2017) implemented in IQ-TREE 2.2.0 (Minh et al. 2020) was used to find the optimal partitioning scheme and DNA substitution models. Partitioning analyses had different settings depending on the downstream analysis. The full dataset was analysed using the Bayesian information criterion (BIC) with the MrBayes model subset for subsequent ExaBayes analysis. To partition the reduced dataset for BEAST, model selection used

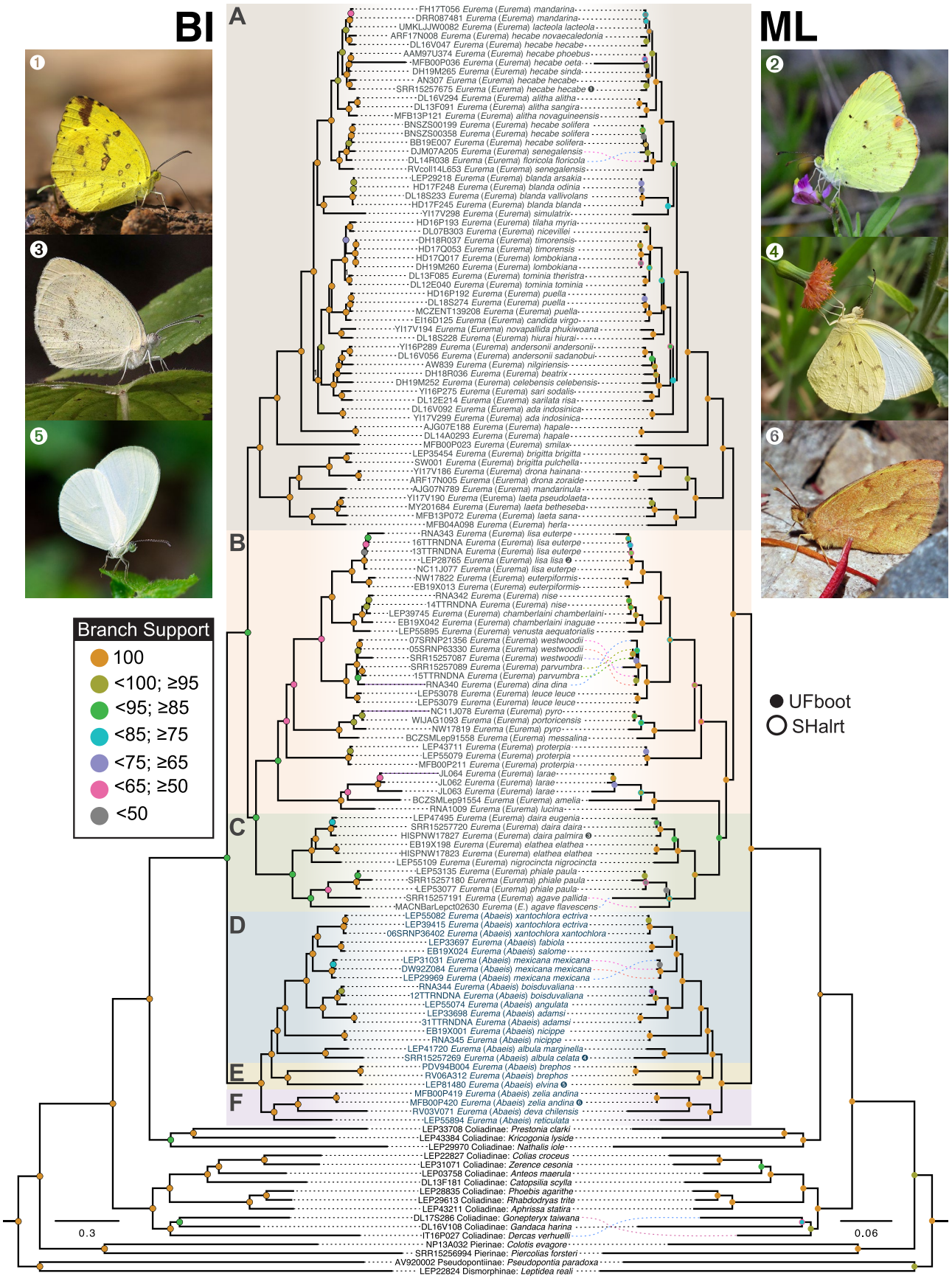

FIGURE 1 | Legend on next page.

FIGURE 1 | Consensus ExaBayes Bayesian inference (BI, left) and IQ-TREE ML (right) trees of *Eurema* grass-yellow butterflies and relatives. The inferred relationships are well supported as indicated by the nodal posterior probability support values in the BI tree and UFBoots with SH-aLRT tests in the ML tree. The sample code, genus, subgenus and species are indicated for each specimen, along with subspecies if applicable. The two subgenera of *Eurema* are indicated with differently coloured tip labels, and exemplar photos of select taxa are identified with a number corresponding to their name among the tip labels. The six clades of grass-yellows are indicated with colour blocks and labelled A–F. Photo credits: (1) Antonio Giudici, used with permission; other photos from inaturalist.com; (2) Edward Perry IV (CC BY-NC); (3) upupamartin (CC BY-NC-ND); (4) alessandradalia (CC-BY-SA); (5) Luciano Bernardes (CC-BY-NC); (6) Carlos Schmidtutz (CC BY-NC).

the BIC and only included models that can be implemented in that programme. The full dataset was analysed with the corrected Akaike information criterion (AICc) using the full model set for subsequent analysis in IQ-TREE.

Bayesian inference (BI) and maximum likelihood (ML) optimality criteria were used to infer the evolutionary history of grass-yellow butterflies using ExaBayes 1.5.1 (Aberer et al. 2014) and IQ-TREE 2.2.0, respectively. Four Markov chains, one cold and three heated, were run simultaneously for 10 million generations in ExaBayes, sampling every 1000 generations until the average standard deviation of split frequencies reached 5.05% and did not improve. The first 25% of trees were discarded as burn-in before a majority-rule consensus tree was generated. Analyses were run on the CIPRES Science Gateway (Miller 2019). To assess branch support in the ML tree, 1000 ultrafast bootstrap (UFBoot) replicates (Hoang et al. 2017) were conducted along with 1000 replicates of the Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) to assess branch support (Anisimova and Gascuel 2006). We regard branch support values as high if SH-aLRT and UFBoot values were ≥ 95 in ML analyses and posterior probabilities were ≥ 0.95 in BI analyses (Simon 2022).

2.3 | Time Calibration and Divergence Time Estimates

Divergence times were estimated in a Bayesian framework with BEAST 1.10.4 (Drummond et al. 2012) using the dataset with one sample per species partitioned as described above, which resulted in two partitions of mtDNA and 14 partitions of nDNA data from our 390 nuclear loci. We first ran four BEAST path sampling/stepping-stone analyses using a Yule or Birth-Death tree prior with either one or two clocks with an MCMC chain length of 500,000 while sampling every 5000 steps, followed by 500 path steps (Baele et al. 2012). These analyses identified a Yule tree prior with two lognormal relaxed clocks (one for mtDNA and another for nDNA) as the best fitting parameters.

The fossil record of butterflies and other Lepidoptera is unusually poor, probably because these soft-bodied insects with non-wetting scales were rarely subjected to conditions leading to fossilisation (de Jong 2017). To examine the effect of calibration points on our inferred ages, we ran two separate dating analyses in BEAST. The first analysis used secondary calibrations from a study of lepidopteran evolution that used 2098 loci and was calibrated with 16 fossils (Kawahara et al. 2019). The first calibration point was a clade of Coliadinae with the genera *Anteos*, *Aphrissa*, *Catopsilia*, *Colias*, *Phoebis*, *Rhabdodryas* and *Zerene*, which has a mean age of 29.01 Mya and a 95% HPD of 18.07–40.05 Mya (StDev = 4.6). The second calibration point used the crown age of the Coliadinae+Pierinae at 51.66 Mya with a 95% HPD of 40.42–63.87 Mya (StDev = 4.5).

The second divergence dating analysis used three calibration points from a molecular phylogeny of butterfly genera inferred with 391 loci and calibrated with up to 11 fossils (Kawahara et al. 2023): the crown age of Coliadinae at 40.12 Mya with a 95% HPD of 38.22–42.94 Mya (StDev = 0.9), the crown age of *Eurema* grass-yellows *sensu lato* at 24.55 Mya with a 95% HPD of 22.46–28.72 Mya (StDev = 1) and the clade including *E.* (*Abaeis*) *mexicana* and *E.* (*Abaeis*) *nicippe* at 16.63 Mya with a 95% HPD of 13.97–21.09 (StDev = 1.3).

For each BEAST analysis, we ran three independent analyses of 500 million generations sampled every 30,000 generations. We confirmed that effective sampling size (ESS) values were above 200 using Tracer v.1.7.1 (Rambaut et al. 2018). LogCombiner 1.10.4 was used to discard the first 25% of trees from each analysis as burn-in. The three analyses were combined, resulting in a set of trees sampled from 1.125 billion MCMC generations. TreeAnnotator 1.10.4 then summarised the pooled trees as two maximum clade credibility trees. BEAST analyses were performed on the Excelsior high-performance computing cluster at the City College of New York (sbgrid.org; Morin et al. 2013).

2.4 | Biogeographic Estimation

The R package BioGeoBEARS v1.1.2 (BioGeography with Bayesian [and likelihood] Evolutionary Analysis in R Scripts) (Matzke 2013) was used to reconstruct ancestral ranges of grass-yellow butterflies. BioGeoBEARS selects among historical biogeographic models including DIspersal Vicariance Analysis (DIVA-like), Dispersal-Extinction-Cladogenesis (DEC) and BAYAREA-like (Landis et al. 2013; Ree and Smith 2008). We used the R package *ape* (Paradis and Schliep 2019) to prune most of the outgroups from the time-calibrated Bayesian tree that included only one sample of each ingroup species (with an exception for the widespread species *E. hecabe*, which had multiple samples; Table S1). To infer the geographic origin of the grass-yellows more accurately, we retained their sister lineage in the analysis. This well-supported clade including *Prestonia clarki*, *Kricogonia lyside* and *Nathalis iole* was recovered as sister to the grass-yellows in comprehensively sampled phylogenies of Pieridae (Carvalho et al. 2024) and of butterfly genera (Kawahara et al. 2023), making us confident that this New World clade is the true, extant sister to our ingroup. We defined two sets of biogeographic regions to examine complementary hypotheses without overparameterising our models. The first analysis used four regions to investigate transcontinental dispersal (Table S2): Africa

(A), Asia (B), Australasia including Australia and Melanesia (C) and the New World including North America, South America and the Caribbean (D). The second parameterisation used six geographic areas to investigate the role of islands in diversification (Table S3): North and South America (α); the Caribbean (β); Africa (γ); continental Asia and Japan (δ); Sundaland, the Lesser Sundas, Wallacea, and the Philippines (ε); and Australasia (ζ). Individuals of *E. hecabe* were coded as occurring in the region where the specimen was caught. Geographic distributions of grass-yellow butterflies and related groups were characterised using numerous sources (butterfliesofamerica.com; gbif.com; Scott 1986; Williams 2023; Yata 1989).

Three sets of parameters were run under three biogeographic models (DEC, DIVALIKE, BAYEAREALIKE): (1) a time-stratified analysis with dispersal rate scalars reflecting changes in geography due to changes in sea levels and continental drift over geologic time; (2) another time-stratified analysis with binary dispersal rates input as 0 or 1 and (3) an unconstrained analysis with no time stratification nor different dispersal scalars between regions (Table S4). We allowed all possible ancestral states for all biogeographic analyses. The evolutionary history of the group was divided into four time periods: 0–5, 5–10, 10–20 and 20–35 Mya. We considered these to be biogeographically relevant intervals during which sea levels changed and continents/crustal fragments drifted—particularly in the IAA—which likely affected dispersal probabilities between geographic areas (Lohman et al. 2011). Dispersal probabilities between geographic areas in the past were estimated by examining distances between landmasses in geological reconstructions performed with GPlates (Müller et al. 2018) using the default dataset and literature describing major geological events including sea level change (Table S4; Lohman et al. 2011; Toussaint et al. 2019). Although there has been criticism of the founder-event speciation "*j* parameter" in BioGeoBEARS (Ree and Sanmartín 2018), the use of this parameter has been subsequently justified (Matzke 2022), and the likelihood of founder-event speciation seems likely in the Caribbean archipelago and IAA where many grass-yellow species are found. We therefore performed analyses with and without this parameter to ascertain the magnitude of its effect.

Following ancestral range reconstruction, we performed biogeographical stochastic mapping (BSM) in BioGeoBEARS using the best-fitting model to estimate the relative number of dispersal events for the four or six biogeographic regions we specified (Dupin et al. 2017; Ree 2005; Ree, Moore et al. 2005). Using 100 pseudoreplicated BSM biogeographic histories, we calculated relative dispersal events for the four or six regions in each time slice following Matos-Maraví et al. (2021). The results were visualised with the R package *qgraph* (Epskamp et al. 2012). We used different minimum dispersal values and cutoff values in *qgraph* to visualise dispersal networks for overall dispersal (min = 1, cut = 4) and for each time slice (min = 0.25, cut = 1).

2.5 | Diversification Analyses

Using molecular phylogenies to estimate shifts in diversification rates through time can be controversial (Barido-Sottani and Morlon 2023; Helmstetter et al. 2022; Louca and Pennell 2020; Meyer and Wiens 2018; Quental and Marshall 2011). We therefore estimated diversification rate shifts using two methods and compared the results. In the first analysis, we inferred changes in branch-specific diversification (BSD) rates with RevBayes (Höhna et al. 2019, 2016) on the tree with secondary calibrations from Kawahara et al. (2019). RevBayes uses a birth-death model to infer diversification rates of different lineages within the tree and determines whether rate shifts occurred on a phylogeny. The configuration script to run the BSD estimation with RevBayes was adapted from https://revbayes.github.io. RevBayes can approximate the continuous base distributions for the diversification-rate parameters by using a discrete rate category (Höhna et al. 2019), a method similar to that of Yang (1994) and Drummond et al. (2006). We set the discrete rate category to 20, as this likely approximated the continuous diversification rate parameter distribution. We used clade-specific sampling fractions in the analysis to account for known species that were not sampled. The net diversification rates were estimated over 5000 MCMC generations with two runs and a tuning interval of 200. All other priors and parameters were left as default values.

In the second diversification rate shift analysis, we used the program ClaDS2 in conjunction with the tree calibrated with divergence times from Kawahara et al. (2023). Branch-specific speciation rates were estimated in the R package RPANDA using the ClaDS2 model (Maliet et al. 2019; Morlon et al. 2016). This Cladogenetic Diversification rate Shift (ClaDS) model estimates speciation (λ) and extinction (μ) rates across lineages and assumes a constant turnover rate (ε) (Maliet et al. 2019). This analysis was run using a sampling fraction of 83%, with three independent MCMC chains, sampling every 100 iterations for 1 million generations.

3 | Results

3.1 | Phylogenetic Relationships and Taxonomy

Bayesian and ML analyses of the full dataset recovered highly congruent topologies (Figure 1). The ingroup taxon is split into two major lineages and several smaller clades. Some of these clades have previously been considered separate genera or subgenera, but we hereby place all ingroup species into the single genus *Eurema* with two subgenera: *Eurema* (*Abaeis*) and *Eurema* (*Eurema*). The rationale and justification for these changes are discussed below.

Most nodes near the root and the nodes defining each main clade had high branch support (PP > 95, UFBoot > 95). Some nodes near the tips were weakly supported (Figure 1). The most significant difference between the trees was the position of a clade with all three Cuban-endemic *Eurema* species, which were previously placed in the genus *Pyrisitia*: *E. amelia*, *E. larae* and *E. lucina*. In the ExaBayes BI analysis, this strongly supported clade was sister to all other species formerly placed in *Pyrisitia* (PP = 0.91; clade B, Figure 1), but the trio of Cuban species was sister to a different clade (*E. agave*, *E. daira*, *E. elathea*, *E. nigrocincta* and *E. phiale*; clade C, Figure 1) in the ML tree (UFboot = 88; SH-aLRT = 37; Figure 1). The topology among our three *E. mexicana* samples was different between analyses, as were the relationships of our two *E. agave* samples (Figure 1). The latter species was recovered as paraphyletic with respect to

E. phiale; however, one sample has 319 loci and the other has only a DNA barcode—possibly explaining the topological instability. Samples of *E. westwoodii*, *E. parvumbra* and *E. dina* formed a clade in both trees, but the topology of the six samples within it differed between analyses. The two samples of *E. senegalensis* were not recovered as sister taxa. *Eurema hecabe*, undoubtedly the most common and widespread Old World species, was recovered as polyphyletic in three separate lineages with relatively strong support (Figure 1). Samples from southern Vietnam and New Caledonia formed a clade that was sister to *E. mandarina* and *E. lacteola*. This group was sister to all other Asian/Australasian *E. hecabe*, with samples from Australia, India, New Guinea, Sulawesi and Thailand. Other geographically widespread species represented by multiple samples—including *E. alitha*, *E. blanda*, *E. brigitta*, *E. laeta*, *E. daira* and *E. mexicana*—were recovered as monophyletic (Figure 1).

There were slight differences between these well-sampled trees (Figure 1) and the BEAST BI timetree analyses (Figure 2 and Figures S1 and S2), which were based on a reduced dataset. Importantly, the composition of the major clades and relationships among them were unchanged, but the trio of Cuban species was recovered in yet another position (PP = 0.67; Figure 2 and Figures S1 and S2). The positions of *E. alitha* and *E. mandarina* differed, and three clades within the New World *Eurema* (*Eurema*) subgenus differed in their relationships to each other. The two BEAST analyses of the reduced dataset differed only in the secondary calibrations used to estimate divergence times. However, the topologies of the two major clade compatibility trees differed: *Eurema proterpia* was recovered in a slightly different location (Figure S1). The posterior probability of this divergence was low (PP = 0.43; Figure S2).

3.2 | Divergence Times and Biogeography

In all analyses of the four- and six-area datasets, the DEC + *j* model had the highest log-likelihood value, and time-stratified analyses with binary dispersal scalars were the best fit to the data: LnL = 57.26 for four areas and LnL = −111.17 for six areas (Table 1 and Table S5). Log-likelihood differences among the various models in the time-stratified analyses (with binary and non-binary scalars) were minimal. However, log-likelihoods in the unconstrained analysis were considerably worse. Using binary scalars substantially improved the fit of each biogeographic model. We present results from the ancestral area reconstruction using the best fitting time-stratified DEC + *j* model with binary scalars for four areas (Figure 2), and dispersal histories under the time-stratified DEC + *j* model with binary scalars for four and six areas (Figure 3).

Most estimated divergence times were approximately 1 Myr older in the tree with secondary calibrations from Kawahara et al. (2019) (Figure 2 and Figure S2) than the tree with calibrations from Kawahara et al. (2023) (Figure S1). We elected to use the divergence estimates of the former in Figure 2 because this study uses many more markers and fossil calibrations than the latter, and because we viewed the older ages to be more conservative. The ingroup clade of grass-yellow butterflies began diversifying ~26.35 Mya (Figure 2 and Figure S2; crown age; 95% HPD 21.17–31.82 Mya). The New World was the most likely inferred ancestral area occupied by the MRCA of the ingroup in the BioGeoBEARS analysis with four areas. The ancestral range of *Eurema* (*Eurema*), which emerged 22.91 Mya (crown age; 95% HPD 18.22–27.73 Mya), was inferred to be the New World or the New World and Asia. Within this subgenus, the Old World clade emerged 19.74 Mya (crown age; 95% HPD 15.68–24.12 Mya) in Sundaland. The subgenus *Eurema* (*Abaeis*) began diversifying *ca*. 19.25 Mya (95% HPD 14.98–23.79) in the New World. Figure S2 shows the full, time-calibrated tree with all outgroups, posterior probability values and 95% HPD bars.

The earliest dispersal event in the four-area analysis was from the New World to Asia, whereas the six-area analysis inferred dispersal between North and South America and the Caribbean. In the six-area analysis, the dispersal event from the New World to Asia, which seems to have involved a single event, was below the cut-off threshold for inclusion in Figure 3. Dispersal from Asia to Africa appears to have been predominantly unidirectional, and the six-area analysis suggests that this was predominantly from the IAA rather than the Asian mainland. *Eurema* species in Australasia resulted from lineages that dispersed from Asia, including the IAA and the mainland. It is notable in the six-area analysis that dispersal out of the IAA to all other Old World regions (Africa, continental Asia, and Australasia) was more common than dispersal into the IAA (Figure 3), suggesting that the islands of this region are a source of many extant lineages.

3.3 | Diversification of Grass-Yellow Butterflies

Variation in the BSD rate λ calculated with RevBayes differed over a narrow range across the tree (Figure 2 and Figure S1). The highest values were in the Old World clade, particularly between *ca*. 2.5–9.5 Mya. The clade with the highest diversification near the tips included *E. tilaha*, *E. nicevillei*, *E. lombokiana* and *E. timorensis*, all of which are endemic to one or more of the Sunda islands in the IAA. Among the New World taxa, the clade of three Cuban-endemic species had the highest diversification rate. Branch-specific speciation rates calculated with ClaDS2 were similar to BSD rates in that the most derived clade (corresponding in part to clade A labelled in Figure 1) had the highest rates (Figure S1). However, ClaDS2 inferred that differences in diversification between this clade and the other branches were less pronounced than estimated by RevBayes (Figure S1).

4 | Discussion

4.1 | Phylogenetics of Grass-Yellow Butterflies

Our molecular phylogenies confirm the previous results of Braby et al. (2006) and Wahlberg et al. (2014) that *Eurema* is not monophyletic as traditionally circumscribed. With only 14 grass-yellow ingroup samples, Zhang et al. (2021) split *Terias* as a separate genus from *Eurema* and redefined the genus *Teriocolias* to include *E. deva* and *E. reticulata*. Subsequent revisionary work by Zhang et al. (2023) made sweeping taxonomic changes with a larger but still incomplete taxon sample. Neither the scope of their genetic data nor the analytical methods were discussed, and there was scant discussion of prior morphological work. Approximately half of the described *Eurema* species occur

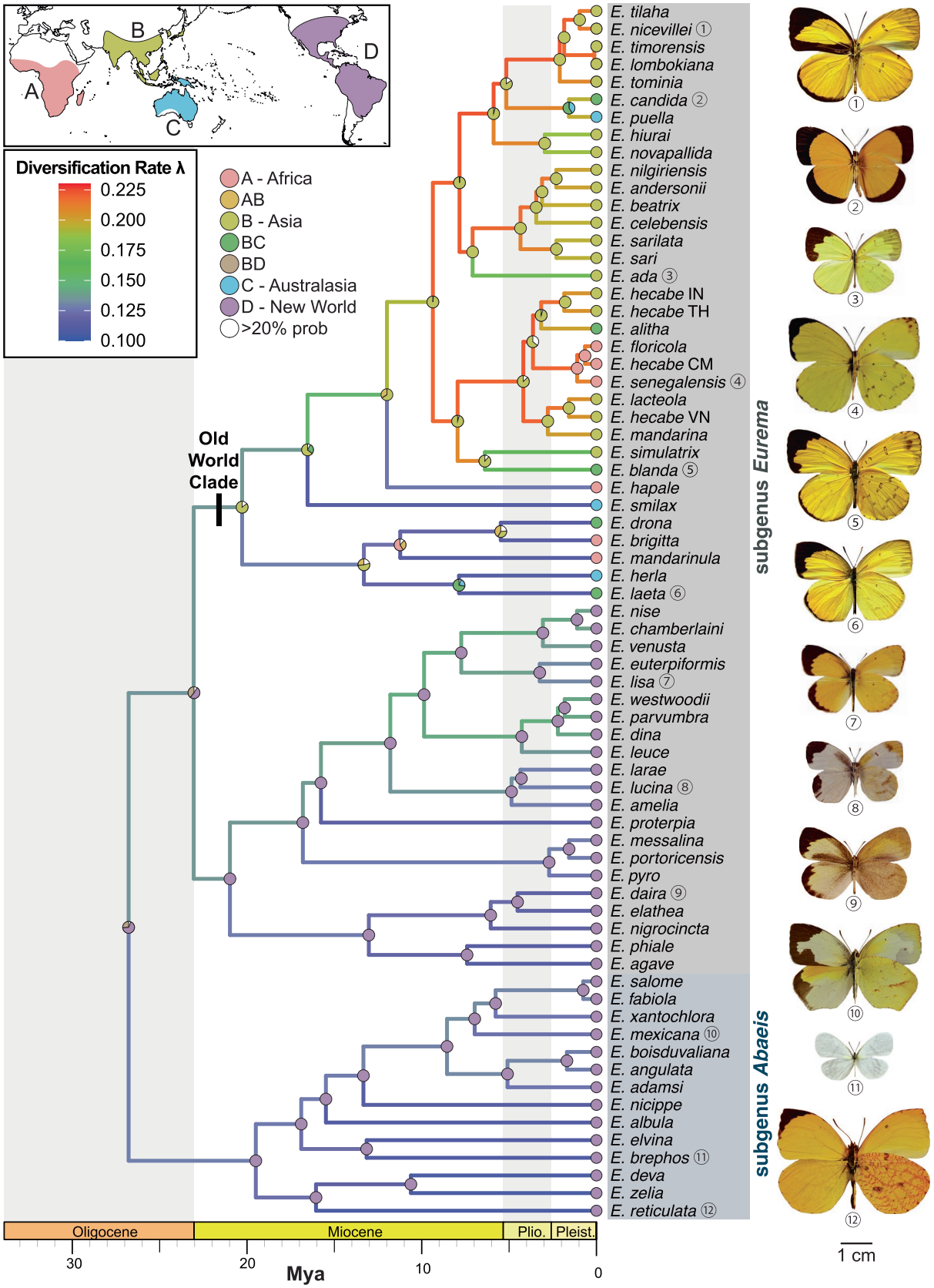

FIGURE 2 | Legend on next page.

in the Old World. Relegating all these taxa to the genus *Terias* would create much taxonomic confusion, as there are innumerable field guides, websites and other resources that use the genus name *Eurema*. Therefore, to maintain nomenclatural stability, and in agreement with Yata (1989), we apply the name *Eurema* in the broadest sense and recognise only a single genus with two

FIGURE 2 | Ultrametric chronogram of grass-yellow butterflies inferred with BEAST using secondary calibrations from Kawahara et al. (2019) with a geological time scale along the bottom. Specimen codes of the samples are indicated in Table S1; ISO country codes of collection are provided for *Eurema hecabe* samples. Coloured blocks under the species names indicate their subgenus. Branch colours correspond to the *diversification rate* λ scale and indicate diversification rate shifts inferred by a BSD analysis. Coloured circles at the tips indicate the species' range with reference to the map and legend. Geographic areas and combinations of areas are indicated with colour in the inset map and in pie charts at nodes, which indicate relative probabilities of an ancestor residing in one or more areas as inferred with BioGeoBEARS. All areas or combinations of areas with $<20\%$ probability are indicated with white. The BioGeoBEARS analysis was conducted with its most closely related outgroup, a single New World outgroup clade with *Prestonia clarki*, *Kricogonia lyside* and *Nathalis iole*, which were removed for presentation. The Old World Clade is indicated with a tick mark; all other taxa live in the New World. Selected species are shown to scale. The left half is a dorsal image, and the right half is a ventral image. The best side was selected for presentation before manipulation in Adobe Photoshop (adobe.com). Image credits: 1, 3, 5, 6 Y. Inayoshi; (10) Field Museum of Natural History, FMNHINS-0000-124-056. All other photos Yale Peabody Museum: (2) YPM-ENT-736151, (4) YPM-ENT-736034, (7) YPM-ENT-735775, (8) YPM-ENT-737671, (9) YPM-ENT-736140, (11) YPM-ENT-756643 and (12) YPM-ENT-735823.

subgenera for this large group of butterflies (see Taxonomy of *Eurema sensu lato* below).

Osamu Yata revised the taxonomy of Old World *Eurema* in over 300 pages spanning five publications (Yata 1989, 1991, 1992, 1994, 1995). In the first of these papers, Yata (1989) presented hypothesised phylogenies based on morphological characters for his conception of subgenus *Terias* (Figure S3). Our results are somewhat congruent with his hypotheses for relationships among the species that he included in that group. However, we recovered *E. brigitta*, *E. drona*, *E. mandarinula* (which Yata considered a subspecies of *E. desjardinsii*—not sampled in this study), *E. laeta* and *E. herla* as a clade within the subgenus *Terias* that is sister to the species that Yata placed in that group. He placed these taxa in the subgenus *Eurema* (*Eurema*). In addition, Yata placed the species *E. lacteola* in a clade with *E. novapallida* and *E. hiurai*, but the single *E. lacteola* DNA barcode we obtained from GenBank was recovered as sister to *E. mandarina* in the *hecabe* group. Additionally, our results suggest that *Eurema ada* is sister to Yata's *sari* group rather than an early diverging lineage (Figure S3).

Eurema hecabe seems to be polyphyletic as currently circumscribed. We included nine *E. hecabe* samples from eight countries in our full dataset and four samples from four countries in the reduced dataset. They were recovered in three distinct lineages in all analyses (Figures 1 and 2 and Figures S1 and S2). Samples from Thailand, India, Sulawesi, New Guinea and Australia formed a clade sister to *E. mandarina*, *E. lacteola* and *E. hecabe* from Vietnam and New Caledonia. The African subspecies *E. hecabe solifera* was always separate from, and not sister to, Asian/Australasian *E. hecabe*. It was consistently recovered in a clade including the African species *E. floricola* and *E. senegalensis*. Although our sampling of *E. hecabe* does not cover its full geographic range, these results strongly suggest that the African taxon is a distinct species. We recommend further geographic sampling and an investigation of species limits within *E. hecabe*. The type locality of *E. hecabe* is Guangzhou [Canton], China (Corbet 1949).

4.2 | Origins, Biogeographic History and Diversification of Grass-Yellow Butterflies

Yata (1989) hypothesised that *Eurema* originated in the Old World and dispersed into the New World and that Old World *Eurema* diversification was mainly driven by vicariance. Our results suggest the opposite. The closest outgroup lineage and the first two diverging *Eurema* lineages are all confined to the New World, suggesting the taxon originated there before a single dispersal event to the Old World (Figures 2 and 3), most likely to Asia via Beringia. The stem age of the Old World Clade dates to the early Miocene at 22.91 Mya (95% HPD 21.18–31.82 Mya), with a crown age of 19.74 Mya (95% HPD 18.21–27.73 Mya). The Beringian land bridge between North America and Eurasia was severed relatively recently in the late Miocene or early Pliocene (Marincovich and Gladenkov 1999). Prior to this, pulses of plant and animal dispersal were strongly affected by climatic fluctuations (Alfimov and Berman 2001; Wolfe 1994), but the peak of trans-Beringian arthropod dispersals 15–30 Mya (Jiang et al. 2019) coincides with our estimate of the period during which the ancestors of the Old World Clade made the journey.

Most species in the first three diverging lineages of the Old World Clade are currently distributed in Australasia and Africa, with only a few species also found in continental Asia: only *E. laeta* and *E. drona* occur in continental Asia and the IAA in addition to Australasia. No species in this grade are confined to continental Asia, which would have been the first area occupied by butterflies traversing Beringia. Thus, it seems that ancestral *Eurema* dispersed from the Americas to Asia via Beringia and then dispersed further south without leaving immediate descendants in continental Asia.

This situation is similar to other butterfly taxa, including blue butterflies in the *Polyommatus* section (Lycaenidae) and metalmark butterflies in the subtribe Corrachiina (Riodinidae). Each of these groups has South American species descended from an Asian lineage that crossed Beringia and left no extant relatives in North America, which is where the Asian migrants would have first lived after their eastward journey (Espeland et al. 2015; Vila et al. 2011). We regard trans-oceanic dispersal or stepping-stone dispersal of their ancestor from the New World to the Old World to be improbable because distances between the Pacific islands are considerable. Similarly, although their diversity in Australasia and South America might suggest a Gondwanan origin with subsequent northbound dispersal, the inferred divergence time of the Old World Clade is too young (estimated to be *ca*. 20 Mya) for vicariance to have played a role in its evolutionary history, as this requires an estimated date of divergence of at least *ca*. 34 Mya for the final separation of South America and Australia from Antarctica (Kawahara et al. 2023).

Diversification in both hemispheres was likely fuelled by archipelagoes. In the New World, dispersal to and from the islands

TABLE 1 | Parameter values and log-likelihoods for time-stratified biogeographic models with four or six areas implemented in BioGeoBEARS with binary scalars describing the ancestral ranges of grass-yellow butterflies (plus *Prestonia clarki*, *Kricogonia lyside* and *Nathalis iole*) given their phylogeny.

Model	# areas	*d*	*e*	*j*	AICc	ΔAICc	AIC_w
DEC	4	0.0122	1.00E-12	0	110.1	1.8	0.238
DEC + *j*	4	**0.0082**	**1.00E-12**	**0.0184**	**108.3**	**0**	**0.591**
DIVALIKE	4	0.016	1.00E-12	0	112.5	4.2	0.072
DIVALIKE+*j*	4	0.0112	1.00E-12	0.0176	111.9	3.6	0.0959
BAYAREALIKE	4	0.0138	0.0146	0	147.8	39.5	1.50E-09
BAYAREALIKE+*j*	4	0.0057	6.78E-04	0.0347	118.73	10.4	3.20E-03
DEC	6	0.0244	1.00E-12	0	237.0528	8.354251	0.015
DEC + *j*	6	**0.0195**	**1.00E-12**	**0.0494**	**228.6985**	**0**	**0.95**
DIVALIKE	6	0.0308	1.00E-12	0	244.2714	15.572913	0.0004
DIVALIKE+*j*	6	0.0231	1.00E-12	0.0494	235.2493	6.550777	0.036
BAYAREALIKE	6	0.0262	0.048	0	2.99E+02	70.301498	5.10E-16
BAYAREALIKE+*j*	6	0.0136	1.00E-07	0.0889	244.8557	16.157229	0.0003

Note: The bold value indicates the best fit model with the highest likelihood.
Abbreviations: *d* = dispersal, *e* = extinction, *j* = jump dispersal.

of the Caribbean was common. In the Old World, there were many dispersal events out of the IAA to the Asian mainland, Australasia, and to Africa (Figure 3). The strong improvement of model likelihoods with the *j* parameter for jump dispersal (Table 1 and Table S5) further suggests that founder event speciation was common. However, rather than dispersal from a 'mainland' area to an island as originally hypothesised by Mayr's (1954) conception of founder-event speciation, the reverse seems to be equally likely in the New World and more likely in the IAA. Evidence for dispersal to and from islands, rather than vicariance, comes from geographic reconstructions of land and sea during the Cenozoic, suggesting that dispersal over land to islands currently inhabited by *Eurema* would have been improbable (Crews and Esposito 2020; Hall 2017; Lohman et al. 2011; Toussaint et al. 2019). Thus, Yata's assumption of vicariance as the predominant mode of diversification is not supported.

Branch-specific diversification analysis reinforces our conclusion that dispersal to and from islands played a pivotal role in grass-yellow evolution (Figures 2 and 3). The highest diversification rates that we detected occurred within the Old World Clade during the late Miocene, including ancestors of many extant island endemics. Similarly, but less pronounced, the fastest diversifying clade in the New World includes many species found in the Caribbean (Figure 2 and Figure S1). This clade had moderately elevated diversification rates during the late Miocene but took a longer time—compared to the IAA—to accumulate species. In contrast, clades found primarily in South America had low diversification rates.

4.3 | Taxonomy of *Eurema sensu lato*

Numerous attempts over the past century to revise this genus using external features, including genitalia, male sex brands, wing colour and wing pattern, have not solidified the nomenclature. Identification of species in this group is difficult, and morphological similarity between non-sister taxa is high. For instance, wing markings in the widespread species *E. hecabe* vary markedly with season (Kato 2000; Yata 1989), and prior molecular phylogenetic studies have demonstrated that *Eurema* is not monophyletic with respect to taxa previously recognised to be in the genera *Abaeis*, *Pyrisitia*, *Teriocolias* and *Leucidia* (Braby et al. 2006; Wahlberg et al. 2014), indicating that these genera are not well defined. The last detailed taxonomic revision of New World *Eurema* and related grass-yellow genera by Klots (1931 [1933]) noted the difficulty of distinguishing *Leucidia* from *Eurema*. Genitalia in *Leucidia* are quite similar to some *Eurema* species, which caused Braby (2005) to suggest that *Leucidia* may be a sister taxon to *Eurema*. Although the type species of the genus, *Eurema daira* (Godart, 1819), was described from the United States, the alpha taxonomy of *Eurema* has been best studied in the Old World by Yata (1989, 1991, 1992, 1994, 1995). Braby (2005) proposed two solutions to resolve the taxonomy of *Eurema*: (1) to raise the Old World subgenus *Terias* to generic status and recognise multiple genera or (2) to recognise a single genus *Eurema sensu lato* and treat *Abaeis*, *Pyrisitia*, *Teriocolias* and *Leucidia* all as subgenera of *Eurema*. However, given the topology of the tree that we recovered, solution (1) would still leave *Eurema* paraphyletic. Thus, we recognise a single genus with two subgenera as follows.

Genus *Eurema* Hübner, [1819] (type species: *Papilio delia* Cramer, 1780, junior homonym; valid name is *Pieris daira* Godart, 1819)

Subgenus *Eurema* (*Eurema*) Hübner, [1819].

Terias Swainson, 1821 (type species: *Papilio hecabe* Linnaeus, 1758) **syn. nov.**

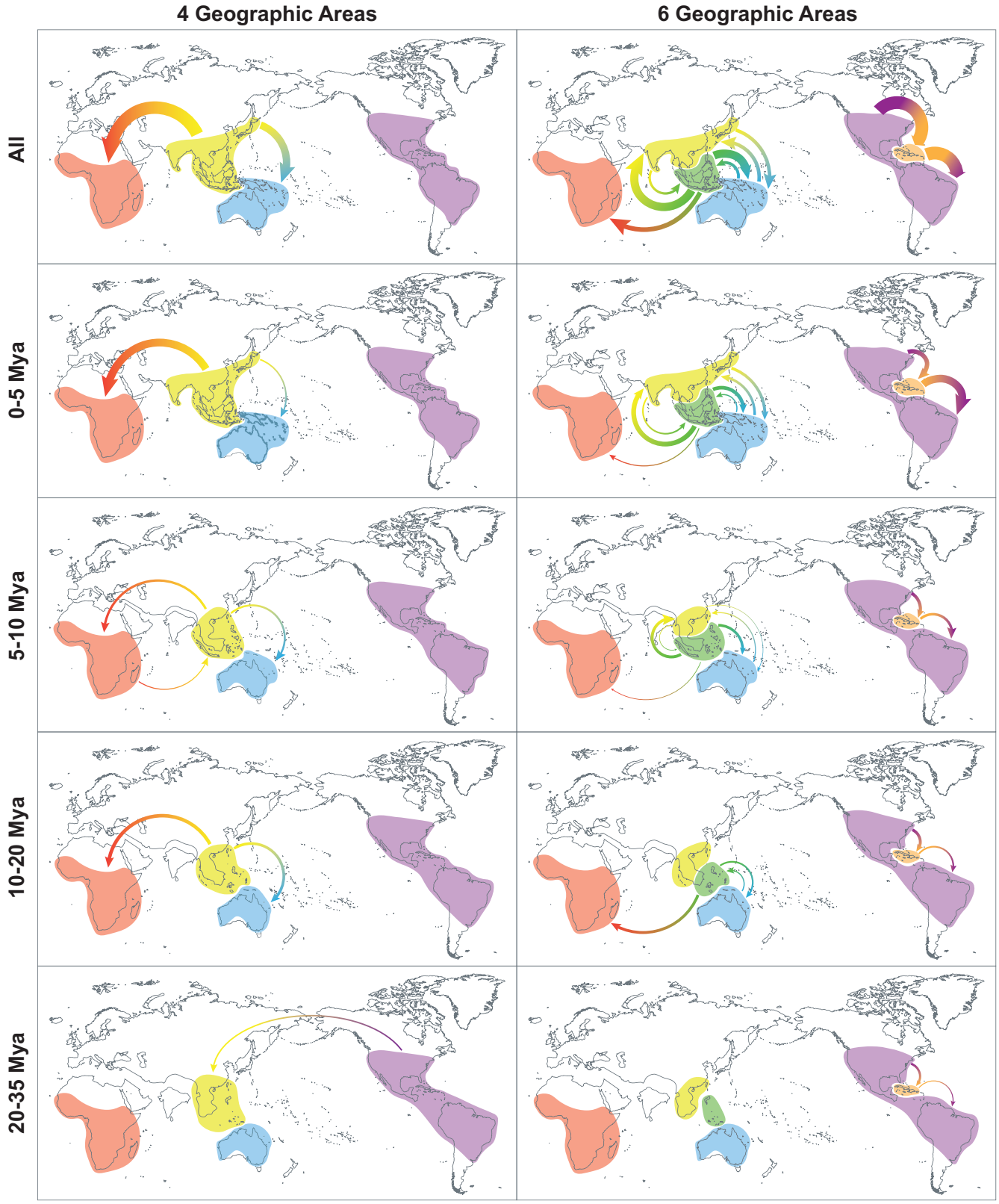

FIGURE 3 | Grass-yellow dispersal events estimated with biogeographical stochastic mapping during different time periods. An analysis with four geographic areas (New World, Africa, Asia, and Australasia) is shown on the left, and an analysis with six areas (separating the Caribbean archipelago and the IAA from the New World and Asia, respectively) is shown on the right. The top two maps represent dispersal events among four or six areas throughout the grass-yellows' entire evolutionary history. Arrows indicate the direction of each dispersal event, and the thickness of each arrow is scaled to the relative dispersal rate between regions through time; there is no significance to the precise placement of each arrow within a region. Maps reflect the changing geography over time at (from the bottom) 35, 20, 10, and 5 Mya and the present.

Heurema Herrich-Schaeffer, 1867 (type species: *Eurema impura* Vollenhoven, 1865)

Pyrisitia Butler, 1870 (type species: *Papilio proterpia* Fabricius, 1775) **syn. nov.**

Maiva Grose-Smith & Kirby, 1893 (type species: *Maiva sulphurea* Grose-Smith & Kirby, 1893, junior synonym of *Papilio brigitta* Stoll, 1780) **syn. nov.**

Kibreeta Moore, 1906 (type species: *Papilio libythea* Fabricius, 1798, unavailable name; valid name is *Terias brigitta rubella* Wallace, 1867)

Nirmula Moore, 1906 (type species: *Terias venata* F. Moore, 1858, junior synonym of *Terias laeta* Boisduval, 1836)

Lirinia Grishin, 2023 (type species: *Terias lirina* H. Bates, 1861) **syn. nov.**

Subgenus *Eurema* (*Abaeis*) Hübner, [1819] (type species: *Papilio nicippe* Cramer, 1779) **syn. nov. et. stat. rev.**

Xanthidia Boisduval & Leconte, 1829 (type species: *Papilio nicippe* Cramer, 1779)

Lucidia Lacordaire, 1833 (type species: *Papilio albula* Cramer, 1775) **syn. nov.**

Leucidia Doubleday, 1847 (type species: *Pieris elvina* Godart, 1819) **syn. nov.**

Sphaenogona Butler, 1870 (type species: *Terias bogotana* C. & R. Felder, 1861) **syn. nov.**

Teriocolias Röber, 1909 (type species: *Terias atinas* Hewitson, 1874, junior synonym of *Terias zelia* Lucas, 1852) **syn. nov.**

4.3.1 | Rationale for Taxonomic Changes

In this work, we recognise a single, species-rich pantropical genus, *Eurema sensu lato*, with two subgenera, *Eurema* (*Eurema*) and *Eurema* (*Abaeis*), based on the topology of our extensively sampled phylogenomic tree. Here, we discuss the rationale for these changes in the context of previous taxonomic and systematic investigations. In a pioneering study, (Klots 1931 [1933]) undertook a detailed morphological revision of all pierid genera. He recognised two genera in the *Eurema* group: *Leucidia* Doubleday, 1847, and *Eurema* Hübner, 1819, with seven subgenera, which were based largely on slight differences in the male genitalia and androconial patches (sex-brands). Importantly, he stated, 'The genus [*Eurema*] as it stands here is quite homogeneous. None of the subgenera possess characters essentially different from those of the others ...Further research may very possibly show that some of these subgenera are not worthy of even that rank' (Klots 1931 [1933], 190). Further, he stated that 'the relationships of *Leucidia* are rather hard to trace. It may be an offshoot from the ancestral stem of *Eurema*, to some of the species of which the genitalia are very similar' (Klots 1931 [1933], 184–185), implying a close relationship between the two taxa. Nevertheless, Klots (1931 [1933]) maintained *Leucidia* as a distinct genus because of the peculiar patches of sex scales located near the base of dorsum of forewing underside and near the base of costa of hindwing upper side. However, he cautioned that this trait may be a secondary character under sexual selection.

Klots' classification was followed by subsequent workers (Bridges 1988; Field 1950; Yata 1989), some with slight modification regarding the treatment of subgenera. Yata (1989) noted that *Eurema* was paraphyletic with respect to *Leucidia* and suggested that *Leucidia* should be synonymised under *Eurema*. However, in a major checklist of the New World butterfly fauna, Lamas (2004) provisionally treated *Leucidia*, as well as the subgenera *Eurema* (*Abaeis*), *Eurema* (*Pyrisitia*) and *Eurema* (*Teriocolias*) from Central and South America, as distinct genera, and Pelham (2008) treated *Eurema* (*Abaeis*) and *Eurema* (*Pyrisitia*) from North America also as distinct genera in accordance with Opler and Warren (2002). Neither checklist provided the rationale for these systematic changes, only that subgenera were treated as full genera as a matter of 'convenience' (G. Lamas, personal communication). Braby (2005) tentatively followed these classifications and thus recognised five genera, namely, *Eurema*, *Abaeis* Hübner, [1819], *Pyrisitia* Butler, 1870, *Teriocolias* Röber, 1909 and *Leucidia* pending further investigation of their monophyly. However, subsequent molecular phylogenies revealed strong evidence that *Eurema* was paraphyletic, with *Leucidia*, *Teriocolias* and *Pyrisitia* all nested within *Eurema* (Braby et al. 2006), indicating that the generic classification for the New World was far from satisfactory. Wahlberg et al. (2014) recovered the same pattern based on a larger data set (eight genetic loci) and called for further taxon sampling and systematic revision of the *Eurema* group. Zhang et al. (2021), however, proposed to treat the subgenus *Eurema* (*Terias*) from the Old World as a valid genus, and thus all species of *Eurema* from Africa, through Asia to Australasia, were placed in *Terias* Swainson, 1821. Thus, they recognised six valid genera, despite the issues of non-monophyly identified earlier by Braby et al. (2006) and Wahlberg et al. (2014). Subsequently, Zhang et al. (2023) proposed further divisions of the *Eurema* group based on a larger taxon dataset. All species were placed in the reinstated subtribe Euremina Grote, 1898, which was composed of five genera in two sections: the *Eurema* section with three valid genera (*Terias*, *Eurema*, *Pyrisitia*) and the *Abaeis* section with two valid genera (*Abaeis*, *Teriocolias*). The genus *Leucidia* was placed in synonymy and treated as a subgenus of *Abaeis* (i.e., *Abaeis* [*Leucidia*]), and a further seven subgenera were recognised, including one newly described, *Pyrisitia* (*Lirinia*).

The topology of our phylogenomic tree is broadly similar to that of Zhang et al. (2023) in that there are two reciprocally monophyletic lineages. Zhang et al. (2023) referred to these two lineages as 'sections', that is, the *Eurema* section and the *Abaeis* section, whereas we formally treat them as subgenera: *Eurema* (*Eurema*) and *Eurema* (*Abaeis*). The composition of taxa in each of these lineages or sections in our study is similar to that of Zhang et al. (2023). Recognising these lineages or sections as distinct genera or recognising a plethora of genera and subgenera within these sections is simply not warranted for such a homogeneous group of organisms in which the lineage (stem group) is of similar age to its sister taxa (*Nathalis* Boisduval, [1836]; *Kricogonia* Reakirt, 1863 and *Prestonia* Schaus, 1920). *Eurema* shows little evidence of deep phylogenetic structure, with short internal branches subtending the early major branches (clades). Moreover, as pointed out long ago by Klots, there are no

consistent morphological characters to diagnose these lineages. The two lineages (subgenera) recognised in our work comprise a composite of concepts as originally proposed by Klots—according to our topology, *Eurema* (*Abaeis*) *sensu lato* includes the subgenera *Eurema* (*Teriocolias*) and *Eurema* (*Abaeis*), as well as the genus *Leucidia*, whereas *Eurema* (*Eurema*) *sensu lato* includes the subgenera *Eurema* (*Eurema*), *Eurema* (*Pyrisitia*), *Eurema* (*Maiva*), *Eurema* (*Nirmula*) and *Eurema* (*Terias*). Hence, the taxa *Terias*; *Maiva* Grose-Smith & Kirby, 1893; *Pyrisitia*; *Lirinia* Grishin, 2023; *Leucidia*; *Lucidia* Lacordaire, 1833; *Sphaenogona* Butler, 1870 and *Teriocolias* recognised as valid by Zhang et al. (2023) are treated as junior synonyms of *Eurema*.

5 | Conclusion

We present the first globally sampled, time-calibrated phylogenetic and biogeographic reconstruction of the pantropical grass-yellow butterflies and inferred events that led to their diversification and cosmopolitan distribution. We found that the genus *Eurema sensu stricto* was not monophyletic and rectified this situation by synonymizing the genera *Abaeis*, *Leucidia*, *Terias* and *Teriocolias* into *Eurema* and recognising the subgenera *Eurema* (*Eurema*) and *Eurema* (*Abaeis*). After originating in the Americas during the Late Oligocene (*ca*. 26 Mya for the crown age), one lineage dispersed across Beringia to Asia in the Early Miocene. This was followed by dispersal to Australasia and Africa. Diversification accelerated during the Late Miocene, especially in the IAA and Caribbean archipelago. Dispersal between continents and between archipelagoes and continents seems to have promoted diversification and led to the group's pantropical distribution. Countering traditional assumptions regarding island biogeography, the islands of the Caribbean archipelago and IAA are not biogeographic dead ends or 'sinks', where lineages become stranded. They can be 'sources' from which new lineages disperse elsewhere.

Author Contributions

A.Y.K., D.J.L., N.E.P.: funding. A.R.P.-A., A.B.M., D.J.L., D.P., D.P.M., J.A.G., J.V.L., K.K., K.A.-P., M.F.B., N.E.P., N.W., Ra.N., Y.-F.H., Y.I.: specimen collection and identification. D.J.L., J.V.L.: data assembly. D.J.L., J.V.L., K.A.-P., P.M.-M., Re.N., T.D., W.L.: analyses and interpretation. D.J.L., J.V.L., M.F.B., P.M.-M., W.L.: writing original draft. All authors: editing.

Acknowledgements

We are grateful to Paola de Lima Ferreira for her suggestions regarding the biogeographic analyses, Caroline Storer for her assistance with bioinformatic processing and to Y-Lan Nguyen and Rebecca L. Messcher for their assistance with lab work. We are grateful to Hélène Morlon for offering advice on implementing the ClaDS2 diversification analysis and thank two anonymous reviewers for their helpful comments on previous versions of this paper. The computing cluster at CCNY is supported by the Graduate Research Training Initiative of the State of New York. Computational resources were also provided by the e-INFRA CZ project (ID: 90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic. We thank Athulya Girish Kizhakke and NCBS Sequencing Facility for sequencing Indian *Eurema*, and state forest departments in Kerala, Karnataka and Arunachal Pradesh for research and collection permits. Permission to collect in Thailand was authorised by the National Research Council of Thailand and the Department of National Parks, Wildlife and Plant Conservation. Fieldwork in Vietnam was conducted in collaboration with Cat Tien National Park, and fieldwork in Indonesia was conducted under various permits from RISTEK and the Department of Forestry. Specimen collection in the Philippines was done with permits from the Department of Environment and Natural Resources and in collaboration with the National Museum of the Philippines as authorised by the National Cultural Heritage Act of 2009. Open access publishing facilitated by Biologicke centrum Akademie ved Ceske republiky, as part of the Wiley - CzechELib agreement.

Ethics Statement

The authors have nothing to report.

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

NCBI BioSample numbers for the data analysed in this study and repositories of each specimen are given in Table S1.

References

Aberer, A. J., K. Kobert, and A. Stamatakis. 2014. "ExaBayes: Massively Parallel Bayesian Tree Inference for the Whole-Genome Era." *Molecular Biology and Evolution* 31: 2553–2556. https://doi.org/10.1093/molbev/msu236.

Alfimov, A. V., and D. I. Berman. 2001. "Beringian Climate During the Late Pleistocene and Holocene." *Quaternary Science Reviews* 20: 127–134. https://doi.org/10.1016/S0277-3791(00)00128-1.

Andermann, T., Á. Cano, A. Zizka, C. Bacon, and A. Antonelli. 2018. "SECAPR—A Bioinformatics Pipeline for the Rapid and User-Friendly Processing of Targeted Enriched Illumina Sequences, From Raw Reads to Alignments." *PeerJ* 6: e5175. https://doi.org/10.7717/peerj.5175.

Andrews, S. 2010. "FastQC: A Quality Control Tool for High Throughput Sequence Data." http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Anisimova, M., and O. Gascuel. 2006. "Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative." *Systematic Biology* 55: 539–552. https://doi.org/10.1080/10635150600755453.

Baele, G., P. Lemey, T. Bedford, A. Rambaut, M. A. Suchard, and A. V. Alekseyenko. 2012. "Improving the Accuracy of Demographic and Molecular Clock Model Comparison While Accommodating Phylogenetic Uncertainty." *Molecular Biology and Evolution* 29: 2157–2167. https://doi.org/10.1093/molbev/mss084.

Barido-Sottani, J., and H. Morlon. 2023. "The ClaDS Rate-Heterogeneous Birth–Death Prior for Full Phylogenetic Inference in BEAST2." *Systematic Biology* 72: 1180–1187. https://doi.org/10.1093/sysbio/syad027.

Bickford, D., D. J. Lohman, N. S. Sodhi, et al. 2007. "Cryptic Species: A New Window on Diversity and Conservation." *Trends in Ecology & Evolution* 22: 148–155. https://doi.org/10.1016/j.tree.2006.11.004.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Braby, M. F. 2005. "Provisional Checklist of Genera of the Pieridae (Lepidoptera: Papilionoidea)." *Zootaxa* 832: 1–16. https://doi.org/10.5281/zenodo.170669.

Braby, M. F., R. Vila, and N. E. Pierce. 2006. "Molecular Phylogeny and Systematics of the Pieridae (Lepidoptera: Papilionoidea): Higher

Classification and Biogeography." *Zoological Journal of the Linnean Society* 147: 239–275. https://doi.org/10.1111/j.1096-3642.2006.00218.x.

Breinholt, J. W., C. Earl, A. R. Lemmon, E. M. Lemmon, L. Xiao, and A. Y. Kawahara. 2018. "Resolving Relationships Among the Megadiverse Butterflies and Moths With a Novel Pipeline for Anchored Phylogenomics." *Systematic Biology* 67: 78–93. https://doi.org/10.1093/sysbio/syx048.

Bridges, C. A. 1988. *Catalogue of Family-Group and Genus-Group Names (Lepidoptera: Rhopalocera)*. Champaign-Urbana: Privately Published.

Carvalho, A. P. S., H. L. Owens, R. A. St Laurent, et al. 2024. "Comprehensive Phylogeny of Pieridae Butterflies Reveals Strong Correlation Between Diversification and Temperature." *iScience* 27: 109336. https://doi.org/10.1016/j.isci.2024.109336.

Chen, I. C., J. K. Hill, R. Ohlemüller, D. B. Roy, and C. D. Thomas. 2011. "Rapid Range Shifts of Species Associated With High Levels of Climate Warming." *Science* 333: 1024–1026. https://doi.org/10.1126/science.1206432.

Compeau, P. E. C., P. A. Pevzner, and G. Tesler. 2011. "How to Apply de Bruijn Graphs to Genome Assembly." *Nature Biotechnology* 29: 987–991. https://doi.org/10.1038/nbt.2023.

Corbet, A. S. 1949. "The Linnaean Names of Indo-Australian Rhopalocera. Part 7. Summary of Determinations." *Proceedings of the Royal Entomological Society of London B* 18: 191–200. https://doi.org/10.1111/j.1365-3113.1949.tb01447.x.

Crews, S. C., and L. A. Esposito. 2020. "Towards a Synthesis of the Caribbean Biogeography of Terrestrial Arthropods." *BMC Evolutionary Biology* 20: 12. https://doi.org/10.1186/s12862-019-1576-z.

Darlington, P. J. 1957. *Zoogeography: The Geographical Distributions of Animals*. New York: Wiley.

de Jong, R. 2017. "Fossil Butterflies, Calibration Points and the Molecular Clock (Lepidoptera: Papilionoidea)." *Zootaxa* 4270: 1–63. https://doi.org/10.11646/zootaxa.4270.1.1.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. "Relaxed Phylogenetics and Dating With Confidence." *PLoS Biology* 4: e88. https://doi.org/10.1371/journal.pbio.0040088.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. "Bayesian Phylogenetics With BEAUti and the BEAST 1.7." *Molecular Biology and Evolution* 29: 1969–1973. https://doi.org/10.1093/molbev/mss075.

Dupin, J., N. J. Matzke, T. Särkinen, et al. 2017. "Bayesian Estimation of the Global Biogeographical History of the Solanaceae." *Journal of Biogeography* 44: 887–899. https://doi.org/10.1111/jbi.12898.

Edgar, R. C. 2004. "MUSCLE: Multiple Sequence Alignment With High Accuracy and High Throughput." *Nucleic Acids Research* 32: 1792–1797. https://doi.org/10.1093/nar/gkh340.

Epskamp, S., A. O. J. Cramer, L. J. Waldorp, V. D. Schmittmann, and D. Borsboom. 2012. "Qgraph: Network Visualizations of Relationships in Psychometric Data." *Journal of Statistical Software* 48: 1–18. https://doi.org/10.18637/jss.v048.i04.

Espeland, M., J. Breinholt, K. R. Willmott, et al. 2018. "A Comprehensive and Dated Phylogenomic Analysis of Butterflies." *Current Biology* 28: 770–778. https://doi.org/10.1016/j.cub.2018.01.061.

Espeland, M., J. P. W. Hall, P. J. DeVries, et al. 2015. "Phylogeny and Biogeography of the Riodinidae (Lepidoptera: Papilionoidea)." *Molecular Phylogenetics and Evolution* 93: 296–306. https://doi.org/10.1016/j.ympev.2015.08.006.

Field, W. D. 1950. "A Revision of *Eurema* Hübner Subgenus *Teriocolias* Röber." *Acta Zoologica Lilloana* 9: 359–374.

Gamble, T., A. M. Bauer, E. Greenbaum, and T. R. Jackman. 2007. "Evidence for Gondwanan Vicariance in an Ancient Clade of Gecko Lizards." *Journal of Biogeography* 35: 88–104. https://doi.org/10.1111/j.1365-2699.2007.01770.x.

Goebel, T., M. R. Waters, and D. H. O'Rourke. 2008. "The Late Pleistocene Dispersal of Modern Humans in the Americas." *Science* 319: 1497–1502. https://doi.org/10.1126/science.1153569.

Hall, R. 2017. "Southeast Asia: New Views of the Geology of the Malay Archipelago." *Annual Review of Earth and Planetary Sciences* 45: 331–358. https://doi.org/10.1146/annurev-earth-063016-020633.

Harris, R. S. 2007. "Improved Pairwise Alignment of Genomic DNA." (Ph.D. Thesis), The Pennsylvania State University, State College, PA. https://www.bx.psu.edu/~rsharris/rsharris_phd_thesis_2007.pdf.

Heads, M., J. R. Grehan, J. Nielsen, and B. Patrick. 2023. "Biogeographic–Tectonic Calibration of 14 Nodes in a Butterfly Timetree." *Cladistics* 39: 293–336. https://doi.org/10.1111/cla.12537.

Helmstetter, A. J., S. Glemin, J. Käfer, et al. 2022. "Pulled Diversification Rates, Lineages-Through-Time Plots and Modern Macroevolutionary Modelling." *Systematic Biology* 71: 758–773. https://doi.org/10.1093/sysbio/syab083.

Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. 2017. "UFBoot2: Improving the Ultrafast Bootstrap Approximation." *Molecular Biology and Evolution* 35: 518–522. https://doi.org/10.1093/molbev/msx281.

Höhna, S., W. A. Freyman, Z. Nolen, J. P. Huelsenbeck, M. R. May, and B. R. Moore. 2019. "A Bayesian Approach for Estimating Branch-Specific Speciation and Extinction Rates." https://doi.org/10.1101/555805.bioRxiv.

Höhna, S., M. J. Landis, T. A. Heath, et al. 2016. "RevBayes: Bayesian Phylogenetic Inference Using Graphical Models and an Interactive Model-Specification Language." *Systematic Biology* 65: 726–736. https://doi.org/10.1093/sysbio/syw021.

Irungbam, M., J. S. Irungbam, M. Rindos, J. P. Maresova, and Z. F. Fric. 2023. "Phylogeography of the Small Grass Yellow *Eurema brigitta* (Lepidoptera: Pieridae) Unveils the Existence of Distinct Taxa Within the Paleotropics." *Austral Entomology* 62: 410–417.

Jiang, D., S. Klaus, Y.-P. Zhang, D. M. Hillis, and J.-T. Li. 2019. "Asymmetric Biotic Interchange Across the Bering Land Bridge Between Eurasia and North America." *National Science Review* 6: 739–745. https://doi.org/10.1093/nsr/nwz035.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. "ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates." *Nature Methods* 14: 587–589. https://doi.org/10.1038/nmeth.4285.

Kato, Y. 2000. "Overlapping Distribution of Two Groups of the Butterfly *Eurema hecabe* Differing in the Expression of Seasonal Morphs on Okinawa-Jima Island." *Zoological Science* 17: 539–547.

Kawahara, A. Y., J. W. Breinholt, M. Espeland, et al. 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., D. Plotkin, M. Espeland, et al. 2019. "Phylogenomics Reveals the Evolutionary Timing and Pattern of Butterflies and Moths." *Proceedings of the National Academy of Sciences* 116: 22657–22663. https://doi.org/10.1073/pnas.1907847116.

Kawahara, A. Y., C. Storer, A. P. S. Carvalho, et al. 2023. "A Global Phylogeny of Butterflies Reveals Their Evolutionary History, Ancestral Hosts and Biogeographic Origins." *Nature Ecology & Evolution* 7: 903–913. https://doi.org/10.1038/s41559-023-02041-9.

Klots, A. B. 1928. "A Revision of the Genus *Eurema* (Lep. Pieridae). Part I. New World Species, Morphology and Phylogeny." *Journal of the New York Entomological Society* 35: 61–72.

Klots, A. B. 1929. "A Revision of the Genus *Eurema* Hübner (Lepidoptera, Pieridae). Part II, New World Species, Taxonomy and Synonymy." *Entomologica Americana* 9: 99–171.

Klots, A. B. 1931. "A Generic Revision of the Pieridae (Lepidoptera): Together With a Study of the Male Genitalia." *Entomologica Americana* 12: 139–240.

Lamas, G. 2004. "Checklist: Part 4A. Hesperioidea-Papilionoidea." In *Atlas of Neotropical Lepidoptera*, edited by J. B. Heppner, 1–430. Gainesville, Florida: Association for Tropical Lepidoptera/Scientific Publishers.

Lamas, G. 2015. "Catalog of the Butterflies (Papilionoidea)".

Landis, M. J., N. J. Matzke, B. R. Moore, and J. P. Huelsenbeck. 2013. "Bayesian Analysis of Biogeography When the Number of Areas Is Large." *Systematic Biology* 62: 789–804. https://doi.org/10.1093/sysbio/syt040.

Larsson, A. 2014. "AliView: A Fast and Lightweight Alignment Viewer and Editor for Large Datasets." *Bioinformatics* 30: 3276–3278. https://doi.org/10.1093/bioinformatics/btu531.

Lin, R. J., Y. C. Lin, M. F. Braby, A. Zwick, and Y. F. Hsu. 2024. "Phylogenetic Relationships and Historical Biogeography of Silkmoths (Lepidoptera: Bombycidae) Suggest an Origin in Southern Gondwana." *Molecular Phylogenetics and Evolution* 200: 108176. https://doi.org/10.1016/j.ympev.2024.108176.

Lohman, D. J., M. de Bruyn, T. Page, et al. 2011. "Biogeography of the Indo-Australian Archipelago." *Annual Review of Ecology, Evolution, and Systematics* 42: 205–226. https://doi.org/10.1146/annurev-ecolsys-102710-145001.

Louca, S., and M. W. Pennell. 2020. "Extant Timetrees Are Consistent With a Myriad of Diversification Histories." *Nature* 580: 502–505. https://doi.org/10.1038/s41586-020-2176-1.

Maliet, O., F. Hartig, and H. Morlon. 2019. "A Model With Many Small Shifts for Estimating Species-Specific Diversification Rates." *Nature Ecology and Evolution Şerban Procheş* 3: 1086–1092. https://doi.org/10.1038/s41559-019-0908-0.

Marincovich, L., and A. Y. Gladenkov. 1999. "Evidence for an Early Opening of the Bering Strait." *Nature* 397: 149–151. https://doi.org/10.1038/16446.

Matos-Maraví, P., N. Wahlberg, A. V. L. Freitas, P. Devries, A. Antonelli, and C. M. Penz. 2021. "Mesoamerica Is a Cradle and the Atlantic Forest Is a Museum of Neotropical Butterfly Diversity: Insights From the Evolution and Biogeography of Brassolini (Lepidoptera: Nymphalidae)." *Biological Journal of the Linnean Society* 133: 704–724. https://doi.org/10.1093/biolinnean/blab034.

Matzke, N. J. 2013. "Probabilistic Historical Biogeography: New Models for Founder-Event Speciation, Imperfect Detection, and Fossils Allow Improved Accuracy and Model-Testing." *Frontiers of Biogeography* 5: 242–248. https://doi.org/10.21425/F5FBG19694.

Matzke, N. J. 2022. "Statistical Comparison of DEC and DEC+J Is Identical to Comparison of Two ClaSSE Submodels, and Is Therefore Valid." *Journal of Biogeography* 49: 1805–1824. https://doi.org/10.1111/jbi.14346.

Mayr, E. 1954. "Change of Genetic Environment and Evolution." In *Evolution as a Process*, edited by J. Huxley, A. C. Hardy, and E. B. Ford. London: Allen and Unwin.

Meyer, A. L. S., and J. J. Wiens. 2018. "Estimating Diversification Rates for Higher Taxa: BAMM Can Give Problematic Estimate of Rates and Rate Shifts." *Evolution* 72: 39–53. https://doi.org/10.1111/evo.13378.

Miller, M. A. 2019. "CIPRES: A Gateway to the Tree of Life." Scientia. https://doi.org/10.33548/scientia381.

Minh, B. Q., H. A. Schmidt, O. Chernomor, et al. 2020. "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era." *Molecular Biology and Evolution* 37: 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Mitchell, K. J., R. C. Pratt, L. N. Watson, et al. 2014. "Molecular Phylogeny, Biogeography, and Habitat Preference Evolution of Marsupials." *Molecular Biology and Evolution* 31: 2322–2330. https://doi.org/10.1093/molbev/msu176.

Morin, A., B. Eisenbraun, J. Key, et al. 2013. "Collaboration Gets the Most out of Software." *eLife* 2: e01456. https://doi.org/10.7554/eLife.01456.

Morlon, H., E. Lewitus, F. L. Condamine, M. Manceau, J. Clavel, and J. Drury. 2016. "RPANDA: An R Package for Macroevolutionary Analyses on Phylogenetic Trees." *Methods in Ecology and Evolution* 7: 589–597. https://doi.org/10.1111/2041-210X.12526.

Müller, R. D., J. Cannon, X. Qin, et al. 2018. "GPlates: Building a Virtual Earth Through Deep Time." *Geochemistry, Geophysics, Geosystems* 19: 2243–2261. https://doi.org/10.1029/2018GC007584.

Nunes, R., C. Storer, T. Doleck, A. Y. Kawahara, N. E. Pierce, and D. J. Lohman. 2022. "Predictors of Sequence Capture in a Large-Scale Anchored Phylogenomics Project." *Frontiers in Ecology and Evolution* 10: 943361. https://doi.org/10.3389/fevo.2022.943361.

Opler, P. A., and A. D. Warren. 2002. *Butterflies of North America*. Vol. 2. Fort Collins: Contributions of the C. P. Gillette Museum of Arthropod Diversity, Colorado State University.

Paradis, E., and K. Schliep. 2019. "Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R." *Bioinformatics* 35: 526–528. https://doi.org/10.1093/bioinformatics/bty633.

Pelham, J. P. 2008. "A Catalogue of the Butterflies of the United States and Canada, With a Complete Bibliography of the Descriptive and Systematic Literature." *Journal of Research on the Lepidoptera* 40: 652.

Quental, T. B., and C. R. Marshall. 2011. "Diversity Dynamics: Molecular Phylogenies Need the Fossil Record." *Trends in Ecology & Evolution* 25: 434–441.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. "Posterior Summarisation in Bayesian Phylogenetics Using Tracer 1.7." *Systematic Biology* 67: 901–904. https://doi.org/10.1093/sysbio/syy032.

Ree, R. H. 2005. "Detecting the Historical Signature of Key Innovations Using Stochastic Models of Character Evolution and Cladogenesis." *Evolution* 59: 257–265. https://doi.org/10.1111/j.0014-3820.2005.tb00986.x.

Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. "A Likelihood Framework for Inferring the Evolution of Geographic Range on Phylogenetic Trees." *Evolution* 59: 2299–2311. https://doi.org/10.1111/j.0014-3820.2005.tb00940.x.

Ree, R. H., and I. Sanmartín. 2018. "Conceptual and Statistical Problems With the DEC+J Model of Founder-Event Speciation and Its Comparison With DEC via Model Selection." *Journal of Biogeography* 45: 741–749. https://doi.org/10.1111/jbi.13173.

Ree, R. H., and S. A. Smith. 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.

Ribeiro, P. G., M. F. Torres Jiménez, T. Andermann, A. Antonelli, C. D. Bacon, and P. Matos-Maraví. 2021. "A Bioinformatic Platform to Integrate Target Capture and Whole Genome Sequences of Various Read Depths for Phylogenomics." *Molecular Ecology* 30: 6021–6035. https://doi.org/10.1111/mec.16240.

Scott, J. A. 1986. *The Butterflies of North America: A Natural History and Field Guide*. Stanford: Stanford University Press.

Simon, C. 2022. "An Evolving View of Phylogenetic Support." *Systematic Biology* 71: 921–928. https://doi.org/10.1093/sysbio/syaa068.

Simpson, J. T., K. Wong, S. D. Jackman, J. E. Schein, S. J. M. Jones, and İ. Birol. 2009. "ABySS: A Parallel Assembler for Short Read Sequence Data." *Genome Research* 19: 1117–1123. https://doi.org/10.1101/gr.089532.108.

Talavera, G., V. Lukhtanov, N. E. Pierce, and R. Vila. 2022. "DNA Barcodes Combined With Multi-Locus Data of Representative Taxa

Can Generate Reliable Higher-Level Phylogenies." *Systematic Biology* 71: 382–395. https://doi.org/10.1093/sysbio/syab038.

Toussaint, E. F. A., F. M. S. Dias, O. H. H. Mielke, et al. 2019. "Flight Over the Proto-Caribbean Seaway: Phylogeny and Macroevolution of Neotropical Anaeini Leafwing Butterflies." *Molecular Phylogenetics and Evolution* 137: 86–103. https://doi.org/10.1016/j.ympev.2019.04.020.

Tseng, H.-Y., H. Chiba, D. J. Lohman, et al. 2022. "Out of Asia: Intercontinental Dispersals After the Eocene-Oligocene Transition Shaped the Zoogeography of Limenitidinae Butterflies (Lepidoptera: Nymphalidae)." *Molecular Phylogenetics and Evolution* 170: 107444. https://doi.org/10.1016/j.ympev.2022.107444.

Vaidya, G., D. J. Lohman, and R. Meier. 2011. "SequenceMatrix: A User-Friendly Gene Concatenation Tool for Assembling Phylogenetic Character Matrices." *Cladistics* 27: 171–180. https://doi.org/10.1111/j.1096-0031.2010.00329.x.

Vila, R., C. D. Bell, R. Macniven, et al. 2011. "Phylogeny and Palaeoecology of *Polyommatus* Blue Butterflies World Show Beringia Was a Climate-Regulated Gateway to the New World." *Proceedings of the Royal Society B* 282: 2737–2744. https://doi.org/10.1098/rspb.2010.2213.

Wahlberg, N., J. Rota, M. F. Braby, N. E. Pierce, and C. W. Wheat. 2014. "Revised Systematics and Higher Classification of Pierid Butterflies (Lepidoptera: Pieridae) Based on Molecular Data." *Zoologica Scripta* 43: 641–650. https://doi.org/10.1111/zsc.12075.

Wahlberg, N., and C. W. Wheat. 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.

Wang, X.-Q., and J.-H. Ran. 2014. "Evolution and Biogeography of Gymnosperms." *Molecular Phylogenetics and Evolution* 75: 24–40. https://doi.org/10.1016/j.ympev.2014.02.005.

Williams, M. C. 2023. "Afrotropical Butterflies and Skippers: A Digital Encyclopaedia." https://www.metamorphosis.org.za/?p=articles&s=Results&pt=166.

Wolfe, J. A. 1994. "An Analysis of Neogene Climates in Beringia." *Palaeogeography, Palaeoclimatology, Palaeoecology* 108: 207–216. https://doi.org/10.1016/0031-0182(94)90234-8.

Yang, Z. 1994. "Maximum Likelihood Phylogenetic Estimation From DNA Sequences With Variable Rates Over Sites: Approximate Methods." *Journal of Molecular Evolution* 39: 306–314. https://doi.org/10.1007/bf00160154.

Yata, O. 1989. "A Revision of the Old World Species of the Genus *Eurema* Hübner (Lepidoptera, Pieridae) Part I. Phylogeny and Zoogeography of the Subgenus *Terias* Swainson and Description of the Subgenus *Eurema* Hübner." *Bulletin of the Kitakyushu Museum of Natural History* 9: 1–103.

Yata, O. 1991. "A Revision of the Old World Species of the Genus *Eurema* Hübner (Lepidoptera, Pieridae) Part II. Description of the *smilax*, the *hapale*, the *ada* and the *sari* (Part) Groups." *Bulletin of the Kitakyushu Museum of Natural History* 10: 1–51.

Yata, O. 1992. "A Revision of the Old World Species of the Genus *Eurema* Hübner (Lepidoptera, Pieridae) Part III. Description of the *sari* Group (Part)." *Bulletin of the Kitakyushu Museum of Natural History* 11: 1–77.

Yata, O. 1994. "A Revision of the Old World Species of the Genus *Eurema* Hübner (Lepidoptera, Pieridae) Part IV. Description of the *hecabe* Group (Part)." *Bulletin of the Kitakyushu Museum of Natural History* 13: 59–105.

Yata, O. 1995. "A Revision of the Old World Species of the Genus *Eurema* Hübner (Lepidoptera, Pieridae) Part V. Description of the *hecabe* Group (Part)." *Bulletin of the Kitakyushu Museum of Natural History* 14: 1–54.

Yata, O., and K. Morishita. 1985. *Butterflies of the South East Asian Islands, Volume 2: Pieridae and Danaidae*. Tokyo: Plapac Co. Ltd.

Zhang, J., Q. Cong, J. Shen, P. A. Opler, and N. V. Grishin. 2021. "Genomics-Guided Refinement of Butterfly Taxonomy." *Taxonomic Report of the International Lepidoptera Survey* 9: 1–55.

Zhang, w., Q. Cong, J. Shen, L. Song, and N. V. Grishin. 2023. "Butterfly Classification and Species Discovery Using Genomics." *Taxonomic Report of the International Lepidoptera Survey* 11: 1–94.

Zink, R. M., R. C. Blackwell-Rago, and F. Ronquist. 2000. "The Shifting Roles of Dispersal and Vicariance in Biogeography." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 267: 497–503. https://doi.org/10.1098/rspb.2000.1028.

Supporting Information

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