Systematic Entomology

Systematic Entomology (2020), **45**, 481–492 DOI: 10.1111/syen.12408

Species delimitation and evolutionary relationships among *Phoebis* New World sulphur butterflies (Lepidoptera, Pieridae, Coliadinae)

RAYNER NÚÑEZ[1], JULIO A. GENARO[2], ANTONIO PÉREZ-ASSO[3], LEIDYS MURILLO-RAMOS[4], DANIEL H. JANZEN[5], WINNIE HALLWACHS[5], NIKLAS WAHLBERG[6] and AXEL HAUSMANN[1]

[1]Bavarian State Collection of Zoology, Section Lepidoptera, Munich, Germany, [2]Florida State Collection of Arthropods, Division of Plant Industry, Gainesville, FL, U.S.A., [3]National Museum of Natural History, Miami, FL, U.S.A., [4]Universidad de Sucre, Sucre, Colombia, [5]Department of Biology, University of Pennsylvania, Philadelphia, PA, U.S.A. and [6]Department of Biology, Lund University, Lund, Sweden

Abstract. The most accepted taxonomic treatment of the New World sulphurs of the genus *Phoebis* Hübner, [1819] recognizes 16 species including those in the current synonyms *Aphrissa* and *Rhabdodryas*. This total conflicts with the results of several recent pierid DNA barcode studies across the Neotropics. We used a five-locus dataset to carry out species delimitation analyses using the coalescence-based method implemented in BPP software. After testing the resulting species hypotheses using marginal likelihood estimates, we inferred their phylogenetic relationships and performed an ancestral range reconstruction with BIOGEOBEARS. Our analyses recovered two different hypotheses, 26 and 24 species, that scored the highest marginal likelihood estimate. Differences between these two hypotheses, when reconciled with barcode clusters and morphology, indicated that 24 is the most likely number of species. *Phoebis neocypris* **stat. rev.**, *Phoebis rurina* **stat. rev.**, *Phoebis virgo* **stat. rev.**, *Phoebis marcellina* **stat. rev.**, *Phoebis thalestris* **stat. rev.**, and *Phoebis rorata* **stat. rev.** are raised to the species rank. We dated the crown age of *Phoebis* to the mid-Miocene, with the islands of the Greater Antilles as the most probable ancestral range. Three main clades of *Phoebis* diverged early in the evolutionary history of the genus, but most extant species-level diversity arose after the Pliocene–Pleistocene boundary. Our analyses recovered alternate range expansions and contractions, and dispersal from the islands to the continent and back, in the three main clades. Both sympatric and allopatric speciation seem to have shaped the current species richness.

Introduction

Coliadinae is a worldwide subfamily of Pieridae currently containing 18 genera and 220 species (Ackery & de Jong, 1999; Braby, 2005). The subfamily has been regarded as a natural group (Klots, 1931; Geiger, 1981; Janz & Nylin, 1998), and, indeed, molecular studies have recovered it as monophyletic (Braby *et al.*, 2006; Wahlberg *et al.*, 2014; Chazot *et al.*, 2019). These studies found a division into two major clades. One of them includes *Eurema* Hübner and related genera, while the other contains all the large sulphur butterfly genera, including the largely Neotropical *Phoebis* Hübner (Fig. 1).

The taxonomy of *Phoebis* has been assessed mainly by classical morphological approaches, through the study of external features and genitalia (Brown, 1929, 1933; Klots, 1929; d'Almeida, 1940). Recent large-scale screenings of Central and South American butterfly faunas using cytochrome oxidase subunit I (COI) barcodes have suggested the existence of cryptic

Correspondence: Rayner Núñez, Bavarian State Collection of Zoology, Section Lepidoptera, Munich, Germany. E-mail: rayner-nunez75@gmail.com

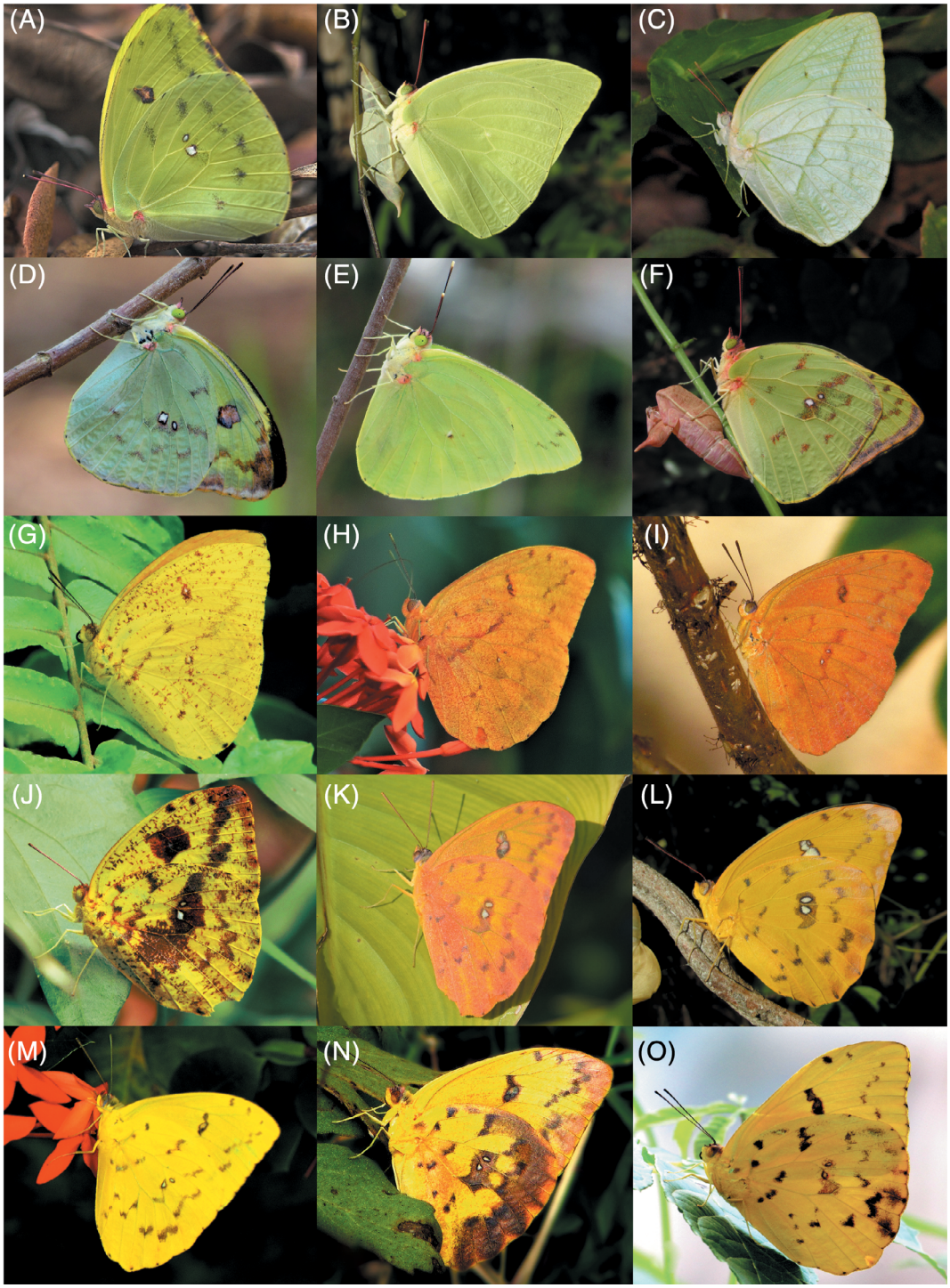

Fig. 1. Adults of *Phoebis* butterflies. (a) *Phoebis neleis*, female Cuba; (b) *P. neleis*, male Cuba; (c) *P. trite watsoni*, female Hispaniola; (d) *P. godartiana godartiana*, female Hispaniola; (e) *P. godartiana godartiana*, male Hispaniola; (f) *P. orbis browni*, female Hispaniola; (g) *P. argante rorata*, male Hispaniola; (h) *P. avellaneda*, male Cuba; (i) *P. editha*, female Hispaniola; (j) *P. argante rorata*, female Hispaniola; (k, l) *P. philea philea*, female Cuba; (m) *P. philea philea*, male Cuba; (n) *P. philea thalestris*, female Hispaniola; (o) *P. philea thalestris*, male Hispaniola. Photographs: (a) Francisco Rodríguez; (b, f, l) Douglas M. Fernández; (k, m) Rayner Núñez; (n) Julio A. Genaro; (c–g, i, j, o) Antonio R. Pérez–Asso. [Colour figure can be viewed at wileyonlinelibrary.com].

diversity in the genus (Janzen *et al.*, 2009; Basset *et al.*, 2015; Lavinia *et al.*, 2017). More recently, the genus has been revised by Murillo-Ramos *et al.* (2018) using 20 morphological characters and four genes. They demonstrated that *Aphrissa* Butler and *Rhabdodryas* Godman & Salvin, 1889 should be treated as junior synonyms of *Phoebis*. These results confirmed the classification by Klots (1929, 1931), which was based on morphology. However, most subsequent authors preferred to keep *Aphrissa* and *Rhabdodryas* as separate genera.

The phylogenetic position of *Phoebis* has been assessed in several molecular studies (Braby *et al.*, 2006; Wahlberg *et al.*, 2014; Chazot *et al.*, 2019). The genus belongs to one of the two main clades within Coliadinae and is a sister of a clade grouping the genera *Anteos* Hübner, *Catopsilia* Hübner, *Colias* Fabricius, *Zerene* Hübner and *Dercas* Doubleday. Our understanding of the relationships among *Phoebis* species remains incomplete. *Phoebis sennae* (Linnaeus) and *Phoebis statira* (Cramer) were the only species included in molecular phylogenies until Murillo-Ramos *et al.* (2018) added six more. The position of the remaining species, half of the genus, is still unclear and this figure does not include suspected cryptic taxa.

Here we perform the first molecular species delimitation analyses of nearly all species of *Phoebis*. We use a multilocus dataset covering all but one species of the genus, *P. schausi* (Avinoff), and employ a coalescent-based method to infer species hypothesis for *Phoebis* and contrast the results with current traditional taxonomy. The current species delimitation methods implementing the multispecies coalescent model (MSC; Degnan & Rosenberg, 2009) recognize genealogical incongruence and incomplete lineage sorting and include several priors allowing a variety of scenarios regarding the number of species, ancestral population sizes and divergence times (Fujita *et al.*, 2012). With these features, coalescent-based species delimitation methods outperform threshold-based methods (Yu *et al.*, 2017; Luo *et al.*, 2018). We also reconstruct the phylogenetic relationships among species of *Phoebis* and perform the first divergence time and ancestral range estimation analyses.

Materials and methods

Taxon sampling and DNA sequencing

Our sampling covers all known species of *Phoebis* except *P. schausi*, and 20 of the 39 named subspecies (Lamas, 2004; Warren *et al.*, 2018). *Phoebis schausi* is apparently related to *P. statira* and is restricted to southern Mexico and Guatemala. Current knowledge only allows us to distinguish males and our attempts to secure tissue samples were fruitless. The taxa included in the dataset inhabit a broad geographical range from southern United States to central South America, including both sides of the Andes Cordillera and the West Indies.

DNA was extracted from leg tissues of dried specimens using a DNeasy Kit (Qiagen, Hilden, Germany). We amplified and sequenced five gene fragments for *Phoebis* specimens totalling 3154 bp: COI and the nuclear gene regions elongation factor 1α (EF1α), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ribosomal protein S5 (RpS5) and cytosolic malate dehydrogenase (MDH). Primer pair sequences and PCR protocols followed standard protocols described in Wahlberg & Wheat (2008). Sequence assembling and editing were performed using SEQUENCHER 4.10.1 (Gene Codes, Ann Arbor, MI, U.S.A.). We also retrieved from GenBank sequences of *Phoebis* species and employed as outgroups representatives of all other genera in the Coliadinae subclade in which the genus belongs. All new sequences produced were deposited in GenBank.

We prepared three datasets for our analyses: (i) a dataset containing seven outgroup and 56 *Phoebis* specimens using the five gene regions mentioned earlier was used to obtain the calibrated species tree and, excluding the outgroups, a species tree for the species delimitation analyses (Table S1); (ii) a reduced 5098 bp dataset with eight gene regions was used to confirm the relationships among the main *Phoebis* clades (Table S2); the three additional genes were wingless (WG), carbamoylphosphate synthase domain protein (CAD) and isocitrate dehydrogenase (IDH); and (iii) to assign specimens to the putative species in the delimitation analyses, we used a dataset of 425 COI barcodes of *Phoebis* for specimen clustering (Table S3).

Molecular species delimitation

We applied the Bayesian Phylogenetics and Phylogeography (BPP) method as implemented in BPP 3.4 software (Yang, 2015). The method uses the MSC to compare different models of species delimitation and species phylogeny in a Bayesian framework.

BPP requires the a *priori* individual assignment to putative species. We employed three input configurations to assess their effects on the results. In our first input, we introduced each of the 56 available *Phoebis* specimens as a putative species. In the second input, we assigned them to the 27 identified species and subspecies plus three divergent unnamed entities totalling 30 putative species. In the third, we assigned the specimens to 24 putative species combining the most recent findings using DNA data (Janzen *et al.*, 2009; Cong *et al.*, 2016; Lavinia *et al.*, 2017) to cluster the 425 available barcodes and check the allopatric/sympatric status among hypothetical species, each represented by its own cluster. The results of the 30 and 24 species assignments are available in Table S4.

In our BPP analyses, the heredity scalar was set to 0.25 for the mitochondrial locus and to 1 for all nuclear loci. The ancestral population size (θ) parameters α and β are assigned to an inverse gamma distribution (IG[α, β]), and we evaluated three different scenarios following the recommendations of Leaché & Fujita (2010): large ancestral population size, IG[3, 0.2]; medium ancestral population size, IG[3, 0.1]; and small ancestral population size, IG[3, 0.02]. The prior distribution of $\tau\theta$, the parameter controlling the divergence time of the root in the species trees, was diffuse ($\alpha = 3$) and we specified $\beta = 0.066$, enforcing a sequence divergence mean of 3.3% which translates into absolute times of *c*. 11.27 Ma for the crown age of *Phoebis* assuming a butterfly mutation rate of 2.9×10^{-9} (Keightley *et al.*, 2015). We used the A11 analysis, joint species

delimitation, and species tree inference of unguided species delimitation, employing 500 000 generations with sampling every 50 generations and a burn-in of 10 000 generations. We ran the analysis twice using a different seed number for each parameter combination to confirm the consistency of the results.

We compared the different species hypothesis by comparing their marginal likelihood estimates (MLEs) using path sampling (Baele *et al.*, 2013) in BEAST2 (Bouckaert *et al.*, 2014). We used 50 steps with chains running for 100 000 generations with sampling every 10 000 cycles. We always checked that the final effective sample size was > 200. We evaluated the fit of two tree models, Yule and birth–death, for each species delimitation hypothesis but show in our results only the one with the highest MLE. The support for the species delimitation hypotheses was assessed via Bayes factors (ln Bf) (Kass & Raftery, 1995).

Phylogenetic relationships and divergence time estimation

We used *BEAST2 available in BEAST 2.5.2 (Ogilvie *et al.*, 2017) which implements the Bayesian multispecies coalescent method to infer calibrated species trees. Substitution models were inferred for each locus with PARTITIONFINDER v.0.1 (Lanfear *et al.*, 2012). We assigned to the mitochondrial COI locus a gene ploidy of 0.5, and to the remaining nuclear loci a gene ploidy of 2.0 (diploid). Uncorrelated relaxed-clock models were chosen for all loci, and we estimated nuclear clock rates relative to the COI mean clock rate fixed to 1.0. The relative clock mean priors were all lognormal (M = 0, S = 1). The analyses ran for five billion generations with sampling every 500 000th generation on CIPRES (Miller *et al.*, 2010). As no fossils of *Phoebis* or its close relatives are known, we had to rely on secondary calibration points. Two recent studies estimated the age of Coliadinae at *c*. 42 and 48 Ma (in Espeland *et al.*, 2018 and Chazot *et al.*, 2019, respectively). We took two secondary calibration points from the latter which covered most Coliadinae genera. We used uniform priors encompassing the 95% highest posterior density (HPD) intervals for the age of the crown of the Coliadinae clade where *Phoebis* belongs (27.7–44.6 Ma) and the split between *Phoebis* and the *Anteos* clade (21.6–31.6 Ma). Convergence, effective sample sizes, and divergence times with upper and lower 95% HPD bounds were assessed in TRACER 1.6. TREEANNOTATOR was used to summarize the results with a 10% burn-in. The tree was visualized and edited in FIGTREE 1.4.3.

For the purposes of comparison, we also reconstructed the phylogenetic relationships of *Phoebis* using maximum likelihood. We ran RAXML 8.2.10 (Stamatakis, 2014) locally using 1000 replicates with the concatenated dataset with the GTRGAMMA model. We also ran analyses for each individual gene.

Ancestral range estimation

The ancestral range reconstruction of *Phoebis* was estimated with the R package BIOGEOBEARS (Matzke, 2013). The range of the genus was subdivided into five geographical areas: A, the Greater Antilles, including the Bahamas and South Florida; B southeastern United States excluding South Florida; C, the Lesser Antilles; D, southwestern United States and Central America; E, South America (Table S5). BIOGEOBEARS implements the DEC, DIVALIKE and BAYAREALIKE models within a maximum likelihood framework (Matzke, 2013). We avoided the use of the founder-event *j* parameter that was recently criticized (Ree & Sanmartín, 2018). Ancestral area probability was inferred using the time calibrate species tree obtained in StarBEAST2. For comparing models, the Akaike information criterion corrected for sample size (AICc) was used. The AICc weights were recalculated after the analysis, to compare only the three models without the *j* parameter, and use them to obtain relative model probabilities as percentages (Burnham & Anderson, 2002).

Results

Species delimitation with BPP

Our analyses yielded six different hypotheses, containing between 21 and 39 species (Fig. 2). Both runs of each parameter combination converged except the one with 56 putative species as input and the most conservative prior, $\tau\theta$ =IG[3, 0.2]. In the latter case, different lineages were recovered within *Phoebis philea* (Linnaeus) and *P. sennae*, with the arrays varying in each run.

The 26-species hypothesis had the highest MLE, only 8.5 units higher than the 24-species hypothesis, which we do not regard as conclusively different (Table 1). These two hypotheses were obtained with the less conservative priors, $\tau\theta$ =IG[3, 0.1] and $\tau\theta$ =IG[3, 0.02], with the one yielding 24 species obtained by three different combinations of input species assignments and $\tau\theta$ priors (Fig. 2). The differences between the two hypotheses are related to the recognition of all subspecies of *P. sennae* and *P. orbis* (Poey) as separate lineages. We consider the 24-species hypothesis to be the most likely one (see Discussion). The posterior probability (PP) values for the 24-species hypothesis ranged from 0.53 to 1.

The results achieved with the most conservative prior had low MLE values in the three species assignments. The assignment of specimens to species following the traditionally most accepted taxonomic treatment (Lamas, 2004) scored the worst MLE (Table 1).

Species assignment and the $\tau\theta$ prior value had a profound impact on the number of delimited species and their PP. When each specimen was set as a putative species, the analyses resulted in 22, 30 or 39 species hypotheses from the most to the least conservative prior with mean PP values for the number of delimited species of 0.24, 0.20 and 0.24 respectively (Fig. 2). In the analyses with the specimens a priori assigned to 30 species, the output was 21, 24 and 26 hypothetical species with PP = 0.33, 0.35 and 0.48 respectively. With the specimens assigned to 24, species, the resulting hypotheses were 21 species with the most conservative $\tau\theta$ prior (PP = 0.35) and 24 with the others (PP = 0.43 and 0.97, respectively).

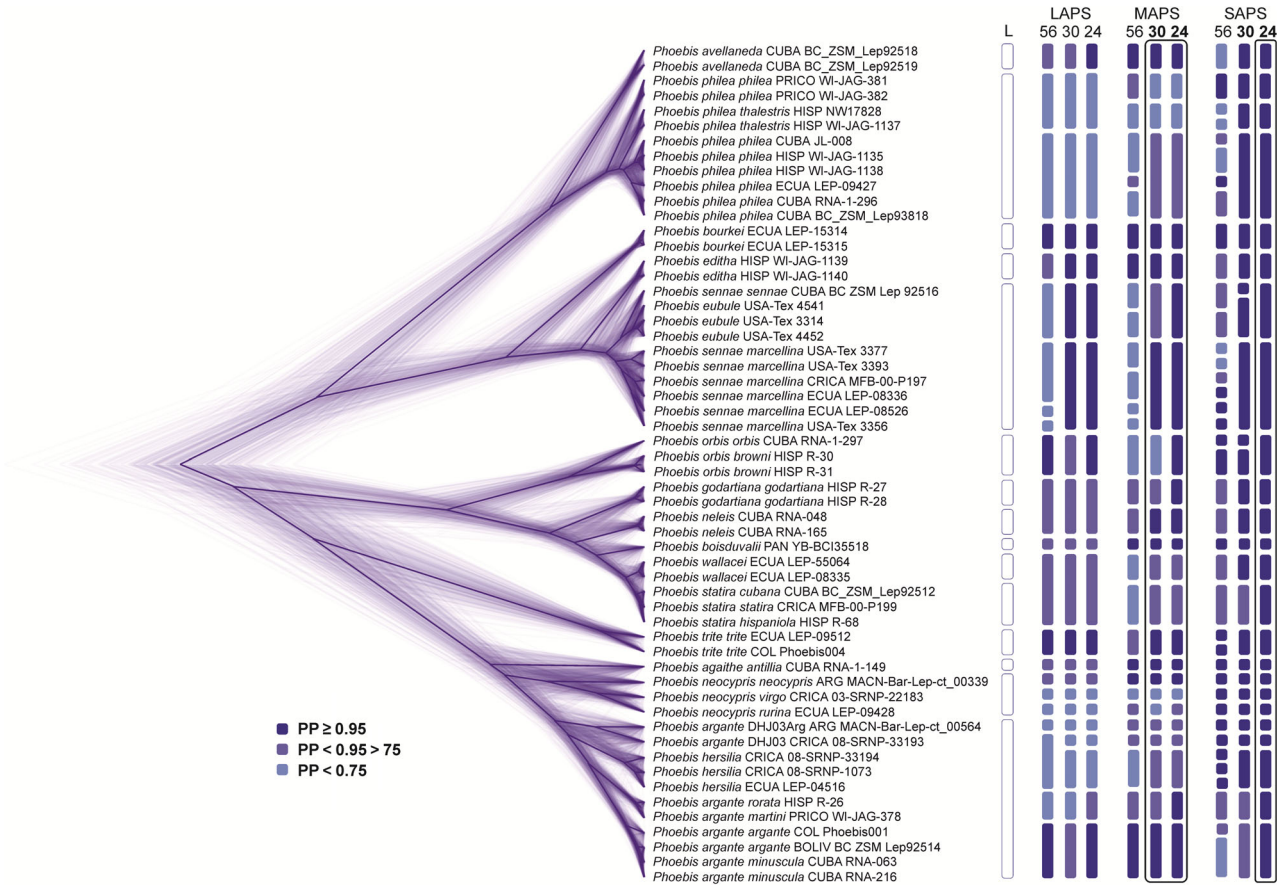

Fig. 2. Results of BPP analyses. Nine combinations of specimens' assignment and prior: specimens assigned to 56, 30 and 24 putative species and each assignment assessed under large, medium and small ancestral population size priors (LAPS, MAPS, SAPS). Species hypothesis in bold scored the best marginal likelihood estimate. Empty bars highlight the preferred 24 species hypothesis (see text for details). Empty bars under 'L' represent the current taxonomy by Lamas (2004). The cloudogram is based on 500 posterior trees from *BEAST. [Colour figure can be viewed at wileyonlinelibrary.com].

Table 1. Marginal likelihood estimate (MLE) values using path sampling and Bayes factor model testing for each species hypothesis obtained in BPP and the most accepted taxonomic treatment (Lamas, 2004). Log Bayes factor (ln Bf) values in the range 2–10 represent positive but not conclusive support and ln Bf > 10 represents decisive support for the species hypothesis with the highest MLE

Species hypothesis	Specimens' assignment	τ0 prior	Tree model	MLE	Ln Bf
26 species	30	3, 0.02	Birth–death	−10 097.39	–
24 species	24/30	3, 0.1/3, 0.02	Yule	−10 105.89	8.5
39 species	56	3, 0.02	Yule	−10 115.62	18.23
22 species	56	3, 0.2	Yule	−10 116.85	19.46
30 species	56	3, 0.1	Birth–death	−10 118.22	20.83
21 species	24/30	3, 0.2	Yule	−10 124.77	27.38
15 species (Lamas, 2004)	–	–	Birth–death	−10 142.24	44.85

Regardless of the assignment of specimens and the $\tau\theta$ prior used, *Phoebis neocypris* (Hübner) and its two related taxa from central South America, as well as *Phoebis argante*DHJ03Arg from central South America were identified as separate lineages. The Hispaniolan and Puerto Rican subspecies of *P. argante* (Fabricius), *rorata* (Butler) and *martini* (Comstock), respectively, were recovered by all analyses as a single lineage separated from other *P. argante* populations, the allied *P. hersilia* (Cramer), and the Central American *P. argante*DHJ03.

Our analyses also yielded evidence of three lineages within *P. philea*, although not for all combinations of specimens' assignments and $\tau\theta$ prior values. Similarly, the Greater Antillean nominal subspecies of *P. sennae* was recovered together with the recently raised *Phoebis eubule* (Linnaeus) in a lineage separated from *P. sennae marcellina* (Cramer).

Table 2. Divergence time estimates for key nodes of the *Phoebis* phylogeny using two tree models, Yule and birth–death. Marginal likelihood estimate (MLE) values using path sampling and Bayes factor model testing. Log Bayes factor (ln Bf) values in the range 2–10 represent positive but not conclusive support

Tree model	MLE	Crown *Phoebis*	Clade C1	Clade C2	Clade C3
Yule	−14 114.07	11.1 (8.55–14.12)	6.1 (4.27–8.12)	4.25 (2.88–5.86)	6.85 (4.62–9.07)
Birth–death	−14 115.12	11.44 (8.56–14.47)	6.84 (4.43–8.46)	4.42 (2.94–6.07)	7.1 (4.97–9.5)

The analyses with the most conservative prior merged the species pairs *P. statira* and *Phoebis wallacei* (Felder and Felder) and *Phoebis neleis* (Boisduval) and *Phoebis godartiana* (Swainson) in single lineages. The analyses with the priors accounting for medium and small ancestral population sizes recovered all four species (Fig. 2).

All other *Phoebis* species included in the analyses were delimited regardless of the prior and species assignment.

Phylogenetic relationships and divergence time estimates within Phoebis

Our calibrated species tree with the preferred 24-species hypothesis using the Yule tree model scored the best MLE value, but only differed by 1.05 units from that tree inferred with a birth–death model (Tables 2, S6). Their divergence time estimates are similar and from here on we use only the estimates obtained with the Yule tree model. There are three major clades within *Phoebis*, which we refer to as clades C1, C2 and C3. All are strongly supported (Fig. 3). The analyses recovered the crown of *Phoebis* in the mid-Miocene *c*. 11 Ma. Clade C1 is sister to clades C2 and C3 but with poor support (PP = 0.64; Fig. 3). The concatenated maximum likelihood tree had a similar topology, but the sister relationship of the C2 and C3 clades was only weakly supported (Fig. S2). The trees obtained with the reduced eight-gene dataset showed the same relationships among the three clades with strong support at all nodes (Fig. S3). The five individual gene trees yielded the same three main clades but the relationships among them varied; in most cases it was only weakly supported (Fig. S4).

The genus diverged into three clades early in the Miocene but, with a few exceptions, the divisions leading to most extant species diverged during the Pliocene–Pleistocene boundary *c*. 2.5 Ma (dotted grey line in Fig. 3).

The clade C1 contains two clades. One includes *Phoebis bourkei* (Dixey) as sister to the Hispaniolan *Phoebis editha* (Butler) and the two lineages recovered within *P. sennae*. The other clade groups *Phoebis avellaneda* (Herrich–Schäffer) and *Phoebis philea* together with their relatives (Fig. 3). Most deep nodes within C1 are strongly supported.

Clade C2 groups all species that were previously included in the now synonymized *Aphrissa* (Fig. 3). The crown age of the C2 clade is the youngest, estimated at 4.42 Ma. *Phoebis orbis* is sister to all other members. The next node leads to the island endemic species *P. neleis* and *P. godartiana*, as sister to a clade containing all the continental species. Overall nodal support in clade C2 is strong.

Clade C3 includes *Phoebis trite* (Linnaeus, 1758) as sister to all other species, an early split during the late Miocene, 7.1 Ma (Fig. 3). Following this, our reconstruction obtained the clade grouping the tailed species, *P. neocypris* and relatives, as sister to *P. argante* and its allies (Fig. 3).

Ancestral range estimation

The DEC model scored the highest relative probabilities: 99.18% based on the AICc values and their weights (Table 3). The original BIOGEOBEARS output is shown in Table S7. The most likely ancestral range for the ancestor of *Phoebis* is the Greater Antilles, but the probability was low, only 24% (Fig. S5).

After the initial separation of clade C1, our analysis inferred vicariant events for one subclade, starting in the late Pliocene, leading to speciation in all areas: *P. bourkei* and *P. marcellina* in South America and Central and South America, respectively, while *P. editha* arose in the Greater Antilles and the range of *P. sennae* includes both archipelagos and southeastern North America (Figs 3, S5). Ancestors of the other subclade gave rise to several Greater Antillean endemics, including an apparently unnamed Puerto Rican taxon currently recognized as *P. philea*. Range expansion back to the continent was recovered for the recent most split within this clade in which a vicariant event also led to the Antillean endemic *P. thalestris*.

The ancestor of clade C2 evolved in the Greater Antilles, leading to several endemics, until the Pliocene–Pleistocene boundary, when dispersers reached the continent and expanded through most areas south of southwestern North America (Fig. 3).

During the Miocene and most of the Pliocene, populations of clade C3 gave rise to the widespread *P. trite* and *P. agarithe*. Starting in the late Pliocene and throughout the Pleistocene, ancestral South American populations led to two clades, one containing all tailed species, and *P. argante* and its allies (Figs 3, S5). Members of the latter rapidly expanded their range to Central America and the Greater Antilles.

Discussion

Species delimitation and its taxonomic implications

The performance of BPP was affected when each specimen was assumed to be a putative species, with oversplitting occurring in analyses in which the ancestral population

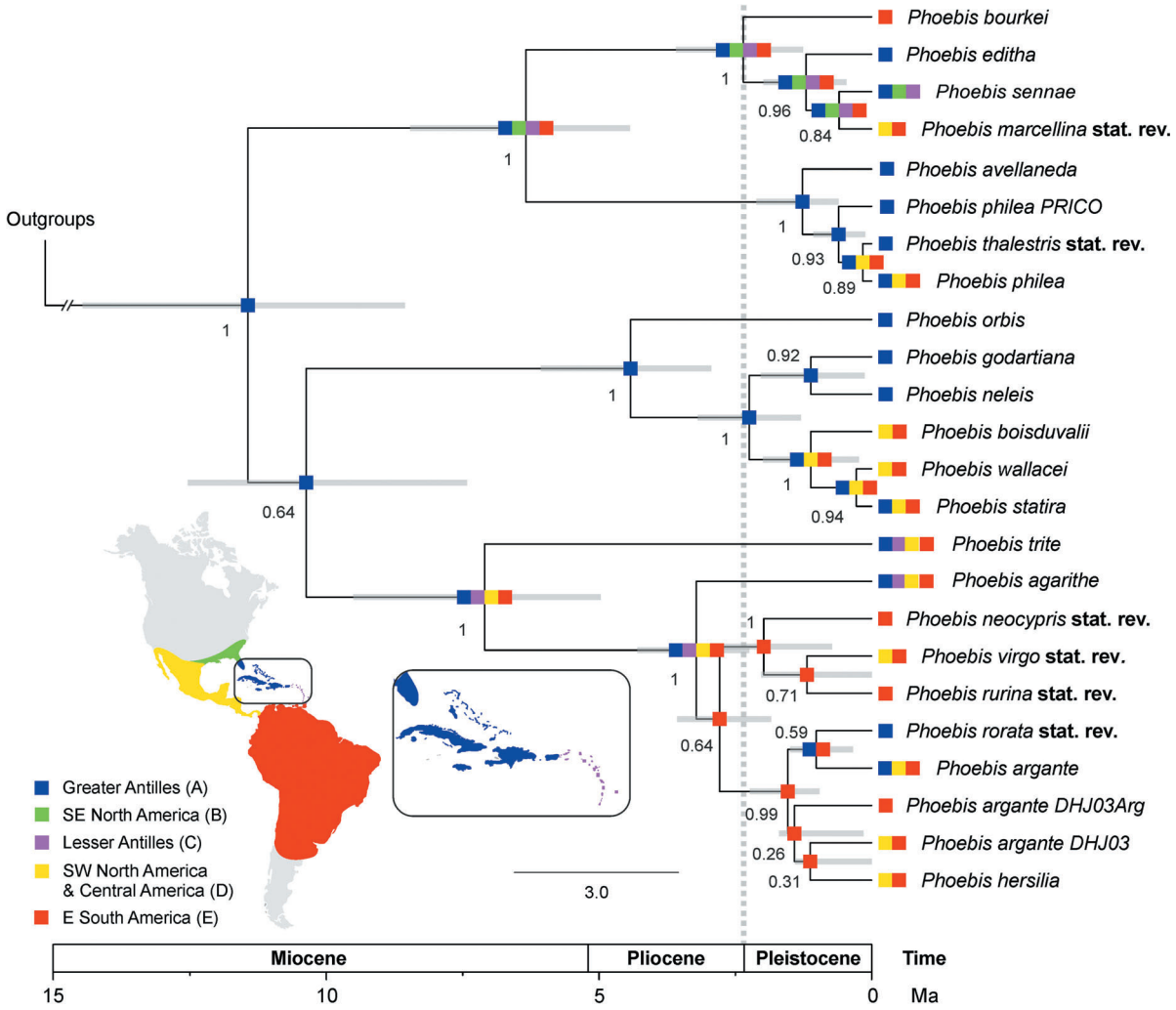

Fig. 3. Calibrated species tree, based on secondary calibrations from Chazot *et al.* (2019), and historical biogeography of *Phoebis* butterflies. Horizontal bars indicate the 95% highest posterior density at each node. The geographical distribution is presented to the right of each species' name in the five areas. The results of the best BIOGEOBEARS model, dispersal–extinction–cladogenesis (DEC), are presented. Only the most likely ancestral area is shown at each node (refer to Fig. S5 for the relative probabilities at selected nodes). [Colour figure can be viewed at wileyonlinelibrary.com].

Table 3. Results of the BIOGEOBEARS model comparison

	Ln *L*	No. of parameters	*d*	*e*	AICc	AICc weight[a]	Relative probability
DEC	−65.17	2	0.075	0.000	134.9	0.9988	99.88
DIVALIKE	−71.91	2	0.091	0.000	148.4	0.0012	0.12
BAYAREALIKE	−76.37	2	0.049	0.130	157.3	0.0000	0.00

Ln *L*, log-likelihood of the model; *d*, dispersal; *e*, extinction; relative probability, relative probability of the model in comparison to the others; AICc, Akaike information criterion corrected for sample size; DEC, dispersal–extinction–cladogenesis model.
[a]AICc weights and their corresponding relative probabilities are calculations to exclude the *j* parameter.

prior was set to medium and small size. We think these results reflect the weak phylogenetic signal of some of the markers among closely related taxa (e.g. *P. sennae*–*P. marcellina*, *P. philea*–*P. thalestris*–*P. avellaneda*) (Fig. S4). The performance of the analyses improved when the specimens were assigned to the known species and subspecies plus the three interim-named entities recognized by barcode clustering, and taking into account their allopatric/sympatric status.

Although analyses with the most conservative priors have been regarded as the most appropriate (Leaché & Fujita, 2010; Matos-Maraví *et al.*, 2019), in our case they failed to recognize young species pairs such as *P. neleis* and *P. godartiana*, and *P.*

statira and *P. wallacei*, at least with the markers we employed. The taxonomic status of the *statira-wallacei* species pair is still unclear; at present only two external male characters enable them to be distinguished (Monroe, 2016). The situation is quite different for the first pair. *Phoebis neleis* and *P. godartiana* are allopatric but differ in their phenotypes, genitalia and host plants (Brown, 1929; d'Almeida, 1939;Smith *et al.*, 1994; Monroe, 2016). The status of the Jamaican subspecies *P. godartiana hartonia* (Butler), which shows some phenotypic and genitalic differences from the nominotypic *P. godartiana godartiana* (Brown, 1931; Monroe, 2016), remains unchanged until molecular data become available.

At first, the favoured 24-species hypothesis might be considered as taxonomic oversplitting. However, it is in fact a combination of old species concepts based on phenotypic and genitalic differences (Illiger, 1801; Butler, 1869; Brown, 1929), along with the cryptic species discovered in the last decade by barcode studies (Janzen *et al.*, 2009; Basset *et al.*, 2015; Lavinia *et al.*, 2017).

The three *Phoebis* with tails were initially regarded as valid species (Brown, 1929; d'Almeida, 1940) but treated as subspecies in recent times (Lamas, 2004), probably due the allopatry of populations and an inconsistent application of the subspecies concept. Monroe (2016) resurrected the views of Brown (1929), who acknowledged phenotypic differences, including size and wing pattern, but also in their genitalia. Their barcodes are distinctive (Fig. S1) and they were recovered as separated lineages by all our BPP analyses, regardless the combination of species assignment and priors used (Fig. 2), leading us to recognize *P. neocypris* **stat. rev.**, *P. rurina* (Felder and Felder) **stat. rev.** and *P. virgo* (Butler) **stat. rev.** as valid species.

Phoebis sennae, as subspecies *P. s. marcellina*, and *P. eubule* were distinguished as separated species by Cong *et al.* (2016) in a whole-genome study. *Phoebis eubule*'s barcode differs only by two barcode nucleotides from *P. sennae marcellina* and *P. sennae sennae*, but Cong *et al.* (2016) found a higher divergence among nuclear markers. In all our BPP analyses, *P. marcellina* **stat. rev.** was always recovered as a distinct lineage, whereas *P. eubule* and *P. sennae sennae* were merged in most cases (Fig. 2). Specimens from southeastern North American and Greater Antillean populations exhibit a less marked underside pattern, differentiating them from *P. marcellina* (Smith *et al.*, 1994; Monroe, 2016). In addition, southeastern North American populations may be the result of successful repeated colonizations from the Greater Antilles. Supporting this hypothesis is the fact that all other *Phoebis* species inhabiting southeastern North America are restricted to subtropical Florida, along with many other butterflies endemic to that region, and to the Bahamas and the Greater Antilles (Smith *et al.*, 1994). Floridan *P. philea* and *P. agarithe* are separated by hundreds of kilometres from conspecific southwestern North American populations, indicating a probable Greater Antillean origin. Even so, a Central American–southwestern North American origin for *P. sennae*, expanding across southeastern North American and reaching the Antilles, cannot be ruled out. With the available data and since *P. sennae* has priority over *P. eubule*, the name *P. sennae* should be used for these populations, as in treatments prior to Cong *et al.* (2016). The status of *P. sennae amphitrite* (Feisthamel, 1839) from Chile, the last named subspecies, remains to be verified but is very likely to be a separated species altogether.

Another instance suggesting *Phoebis* speciation in the Greater Antilles is found in *P. philea* and its relatives. *Phoebis philea* and *P. thalestris* (Illiger) were described as separate species but since the beginning of the 20th century only the first has been recognized by most authors (Röber in Röber, 1909; Brown, 1929; Bates, 1935; Torre, 1954; Brown & Heineman, 1972; Riley, 1975; Smith *et al.*, 1994; Lamas, 2004; Monroe, 2016), with only Comstock (1944) and Schwartz (1989) accepting the second. Most of our BPP analyses recovered three lineages related to *P. philea*: *P. philea* represented by specimens from Cuba, Ecuador and Hispaniola; *P. thalestris* from Hispaniola; and the third from Puerto Rico. The presence of two sympatric lineages in Hispaniola supports *P. thalestris* **stat. rev.** as a valid species. *Phoebis philea* and *P. thalestris* have different phenotypes (Smith *et al.*, 1994; Monroe, 2016; Fig. 1k–o), different UV-reflective areas on their wings (Allyn & Downey, Allyn Jr. & Downey, 1977), and similar but not identical male genitalia (Brown, 1929; Schwartz, 1989). On the other hand, the Puerto Rican population phenotype is identical to that of *P. philea* from other territories. More data (molecular, ecological, genitalia, etc.) are needed before any status change for the Puerto Rican taxon can be proposed. This situation is complicated by evidence pointing towards a reverse colonization from the Greater Antilles to the continent and back (Figs 3, S5). Historical records from the islands (Illiger, 1801; Dewitz, 1877; Gundlach, 1881, 1891; Fruhstorfer, 1907) only referred to the existence of *P. thalestris* and *P. huebneri* (Fruhstorfer), the second described from Cuba, although it is probably conspecific with the first and absent from our sampling. *Phoebis philea* was first recorded from south Florida and Cuba in the 20th century (Klots, 1951 and Zayas & García, 1965, respectively. Alayo & Hernández (1987) acknowledge that the newly arrived *P. philea philea* started hybridizing with Cuban *P. philea huebneri*, with intergrades being collected in eastern Cuba since then. Nominate *P. philea philea* had never been recorded from Hispaniola, meaning its expansion through the east has continued and that our record is the first from that island. The results of our ancestral range reconstruction (see later) seem to imply that Puerto Rican *philea* could be direct descendants of the ancestral stock of these closely related lineages.

Phoebis argante is another example in which the traditional morphology-based taxonomy does not fit. Its variability across the Neotropics gave rise to more than a dozen names for species, subspecies, varieties, aberrations, etc. until the first decades of the 20th century (Lamas, 2004) when a consensus to accept a single species was gradually reached and prevailed for *c*. 100 years. However, with the advent of barcode studies, Janzen *et al.* (2009) revealed three sympatric lineages inhabiting Costa Rica. One of these was identified as *P. hersilia* Cramer, 1777, recognized by its silvery white dots on the underside of the wings and the enlarged continuous black borders on forewings of the females, as well as genitalia differences. With no obvious differences between the other two entities, the assignment of one of them to *P. argante* is impossible until DNA from holotypes

becomes available. The other two taxa have been referred to provisionally as *P. argante*DHJ01 and *P. argante*DHJ03 (Janzen *et al.*, 2009; Basset *et al.*, 2015; Lavinia *et al.*, 2017). Rearing records show that *P. hersilia* larvae feed on legumes of the genus *Inga*, which is used also by *P. argante*DHJ01 and *P. argante*DHJ03, but so far there are no hosts in common between the last two (*P. argante*DHJ03 adults have only been collected by net, and hundreds of *P. argante*DHJ01 have been reared from larvae in the same forest). With regard to ecology, the data gathered so far point towards habitat segregation between the three lineages although they may co-occur in the ecotone between rainforests and dry forests, and of course flying adults can easily be found together. As larvae, *P. hersilia* is primarily a rainforest species while oviposition by *P. argante*DHJ01 seems to be mostly in dry forest. The scarce records of *P. argante*DHJ03 are of specimens netted in rainforest. Minor genitalic differences only distinguish *P. hersilia* males from the others. Barcode sequences from Panamian specimens (Basset *et al.*, 2015) matched the three Costa Rican clusters (Fig. S1). More recently, barcodes matching *P. argante*DHJ03 were recovered from northern Argentina together with a fourth distinct barcode from sympatric specimens, here named as *P. argante*DHJ03Arg (Lavinia *et al.*, 2017; Fig. S1). Our BPP analyses included representatives of all these entities plus new sequences from Cuba, Ecuador, Hispaniola and Puerto Rico. For seven of the nine input combinations, the analyses recovered the four entities mentioned and a fifth cluster grouping the representatives of Hispaniola and Puerto Rico, subspecies *P. argante rorata* and *P. argante martini*, respectively. Females of these populations differ from all other relatives in their heavily marked wings with larger post-discal bands on the upper side of the forewings and distinctive enlarged patches on the disc of both wings on the underside (Fig. 1j). Therefore, we reinstate *P. rorata* **stat. rev.** to its original specific status. The Jamaican subspecies *P. argante comstocki* Avinoff awaits confirmation as a member of the latter species.

Evolutionary relationships and historical biogeography of Phoebis

Our results match those of Murillo-Ramos *et al.* (2018) in supporting *Phoebis* as monophyletic after including the species formerly placed in *Aphrissa* and *Rhabdodryas*. Long ago, Klots (1929, 1931) acknowledged the morphological uniformity of the group by proposing a single genus, *Phoebis*.

According to our estimates, the crown clade of *Phoebis* started diversifying during the mid-Miocene (11.1–11.44 Ma), close to the Chazot *et al.* (2019) estimate of 12.22 Ma based on an analysis including only *P. sennae* and *P. statira*. Our analyses suggest a late Miocene–early Pliocene origin (7.1–4.42 Ma) for the three clades containing the extant species of the genus.

Our biogeographical reconstruction inferred an origin of *Phoebis* in the Greater Antilles. Large blocks of the archipelago had already emerged by the middle Miocene (Iturralde-Vinent, 2006). The ancestor was likely to have inhabited eastern Cuba–northern Hispaniola, one the largest blocks that began to break apart 16–14 Ma, close to the upper 95% HPD limit of the age estimated for the first split in the genus. As suggested by present-day behaviour of some species (Urquhart & Urquhart, 1976; Oliveira *et al.*, 1998; Srygley & de Oliveira, 2001), some ancestral populations were probably strongly vagile and colonized the continent across the open ocean. Reverse colonization from the Greater Antilles to the continent has been inferred for other groups, including Urocoptidae land snails (Uit de Weerd *et al.*, 2016), *Heraclides* swallowtail butterflies (Lewis *et al.*, 2015), *Exophthalmus* weevils (Zhang *et al.*, 2017), *Anolis* lizards (Nicholson *et al.*, 2005), *Eleutherodactylus* frogs (Heinicke *et al.*, 2007) and *Amazona* parrots (Russello & Amato, 2004).

Several island endemics might have diverged allopatrically during periods of higher sea levels in the Pleistocene (Iturralde-Vinent, 2003; Iturralde-Vinent, 2006). Examples include *P. orbis* (Cuba and Hispaniola), the closely related pair *P. neleis* (Cuba, Bahamas and south Florida) and *P. godartiana* (Hispaniola and Jamaica), *P. avellaneda* (Cuba) and *P. editha* (Hispaniola). Allopatric speciation may also have shaped the continental diversity, e.g. the members of the tailed clade, but sympatric speciation seems to have played an important role. Several closely related species are found in sympatry across the continent, including *P. statira*, *P. boisduvalii*, *P. wallacei* and probably *P. schausi* in clade C2 and the species of the *P. argante* complex in clade C3.

Our findings also suggest a recent colonization of the islands from the continent, e.g. the lineage leading to *P. rorata*. These events can also explain the coexistence of a widespread species with closely related single island endemics, *P. philea* with *P. avellaneda* in Cuba and with *P. thalestris* in Hispaniola. That might be also the case with the Hispaniolan endemic *P. editha* and its relative *P. sennae*, although our analysis only inferred a broad ancestral range for the latter. This mix of widespread and restricted endemics is analogous to the situation in the genus *Vanessa*, which also has island endemics in Southeast Asia, despite having highly dispersive widespread species (Wahlberg & Rubinoff, 2011).

Future perspectives

Our analyses support a much greater species richness in *Phoebis* than traditionally thought. Genetic divergence together with sympatry in some cases (*P. argante*DHJ01/*P. argante*DHJ03, *P. argante*DHJ03/DHJ03Arg and *P. philea/thalestris*) or distinctive allopatric phenotypes in others (*P. argante/rorata*, *Phoebis* tailed species) support this view. Further work should verify the status of some populations included here (e.g. Puerto Rican *P. philea* and *P. wallacei*), while the complete clarification of the *P. argante* complex status will require the DNA study of holotype specimens. Additional sampling from Jamaica, Chile and Central America is needed to check the status of other populations. A global biogeographic study of the subfamily Coliadinae, including the large widespread genera *Colias* and *Eurema*, will shed more light on the evolution of this group of butterflies.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Neighbour-joining tree for all 425 barcode COI sequences available of species in the genus *Phoebis*.

Figure S2. Maximum likelihood five-gene concatenated tree of *Phoebis*.

Figure S3. Relationships among *Phoebis* main clades reconstructed with the reduced eight-gene dataset.

Figure S4. *Phoebis* maximum likelihood individual gene trees.

Figure S5. Ancestral range probabilities of *Phoebis* species.

Table S1. *Phoebis* and Coliadinae outgroup specimens utilized in the five-gene dataset.

Table S2. *Phoebis* and Coliadinae outgroup specimens utilized in the eight-gene dataset.

Table S3. *Phoebis* specimens included in the COI barcode analysis.

Table S4. Specimens' assignment to 30 and 24 species for the BPP analyses input.

Table S5. Subdivision in five geographical areas of the distribution of *Phoebis* species for the BIOGEOBEARS analysis.

Table S6. BEAST2 analyses, marginal likelihood (MLE) and divergence time estimates.

Table S7. Results of the BIOGEOBEARS model comparison.

Acknowledgements

We thank Feliberto Bermúdez, Douglas M. Fernández, Tim Norriss, Sheyla Yong and Joel Lastra for their help during the fieldwork and collection of specimens in Cuba. We are grateful to Carlos Suriel and Celeste Mir (National Museum of Natural History, Santo Domingo, Dominican Republic), Yvonne Arias and Héctor Andujar (Grupo Jaragua Inc.) and Ruth Bastardo (Autonomous University of Santo Domingo, Dominican Republic) for their support during the fieldwork and study of collections in the Dominican Republic. We thank Keith R. Willmott, in particular, for kindly supplying tissue samples from several *Phoebis* species. We also thank Jerome Moriniere and Isabella Stoeger for their work and help in the molecular laboratory at the ZSM. The parataxonomists (Janzen & Hallwachs, 2011) of Area de Conservacion Guanacaste (ACG) in northwestern Costa Rica reared and processed all the Costa Rican specimens barcoded, and the ACG staff protected the forest in which their caterpillars live; funding for this came from the government of Costa Rica and the Guanacaste Dry Forest Conservation Fund. The Centre for Biodiversity Genomics at the University of Guelph did all of the ACG barcoding. Funding for this project was available through a Georg Foster Research Fellowship from the Alexander von Humboldt Foundation (1162549-CUB-GFHERMESE). Niklas Wahlberg acknowledges funding from the Swedish Research Council (2015-04441). Barry Goater and Tim Norriss helped to improve the grammar. We also thank David Lohman and another anonymous reviewer for their comments on the manuscript. The authors declare there are no conflicts of interest.

References

Ackery, P.R., de Jong, R. & Vane-Wright, R.I. (1999) The butterflies: Hedyloidea, Hesperioidea and Papilionoidea. *Lepidoptera, Moths and Butterflies. Evolution, Systematics and Biogeography*, Vol. **1** (ed. by N.P. Kristensen), pp. 263–300. de Gruyter, Berlin.

Alayo, P. & Hernández, L.R. (1987) *Atlas de las Mariposas Diurnas de Cuba*. Cientifico–Tecnica, La Habana.

Allyn, A.C. Jr & Downey, J.C. (1977) Observations on male U–V reflectance and scale ultrastructure in *Phoebis* (Pieridae). *Bulletin of the Allyn Museum*, **42**, 1–20.

Baele, G., Li, W.L.S., Drummond, A.J., Suchard, M.A. & Lemey, P. (2013) Accurate model selection of relaxed molecular clocks in Bayesian phylogenetics. *Molecular Biology and Evolution*, **30**, 239–243.

Basset, Y., Barrios, H., Segar, S. *et al.* (2015) The butterflies of Barro Colorado Island, Panama: local extinction since the 1930s. *PLoS ONE*, **10**, e0136623.

Bates, D.M. (1935) The butterflies of Cuba. *Bulletin of the Museum of Comparative Zoology*, **78**, 63–258.

Bouckaert, R., Heled, J., Kuehnert, D. *et al.* (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10**, e1003537.

Braby, M.F. (2005) Provisional checklist of genera of the Pieridae (Lepidoptera: Papilionidae). *Zootaxa*, **832**, 1–16.

Braby, M.F., Vila, R. & Pierce, N.E. (2006) Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. *Zoological Journal of the Linnean Society*, **147**, 239–275.

Brown, F.M. (1929) A revision of the genus *Phoebis* (Lepidoptera). *American Museum Novitates*, **368**, 1–22.

Brown, F.M. (1931) A revision of the genus *Aphrissa*. *American Museum Novitates*, **454**, 1–14.

Brown, F.M. (1933) Notes on the genus *Phoebis* and the description of a new species. *American Museum Novitates*, **653**, 1–5.

Brown, F.M. & Heineman, B.W. (1972) *Jamaica and its Butterflies*. E. W. Classey Ltd, London.

Burnham, K.P. & Anderson, D.R. (2002) Model Selection and Multimodel Inference: a Practical Information–Theoretic Approach. Springer Science and Business Media.

Butler, A.G. (1869) Descriptions of three new species of *Callidryas*. *Annals and Magazine of Natural History*, **4**, 202–203.

Chazot, N., Wahlberg, N., Freitas, A.V.L. *et al.* (2019) Priors and posteriors in Bayesian timing of divergence analyses: the age of butterflies revisited. *Systematic Biology*, **65**, 797–813.

Comstock, J.A. (1944) Insects of Puerto Rico and the Virgin Islands. Rhopalocera or butterflies. *Scientific Survey of Porto Rico and the Virgin Islands*, **12**, 419–622.

Cong, Q., Shen, J., Warren, A.D., Borek, D., Otwinowski, Z. & Grishin, N.V. (2016) Speciation in cloudless sulphurs gleaned from complete genomes. *Genome Biology and Evolution*, **8**, 915–993.

d'Almeida, R.F. (1939) Revisão do gênero *Aphrissa* Butl. (Lepid., Pieridae). *Boletim biológico (São Paulo)*, **4**, 423–443.

d'Almeida, R.F. (1940) Revisão do gênero *Phoebis* Hübn. (Lepidopt. Pieridiidae). *Arquivos de Zoologica do Estado de São Paulo*, **1**, 67–152.

Degnan, J.H. & Rosenberg, N.A. (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, **24**, 332–340.

Dewitz, H. (1877) Tagschmetterlinge von Portorico gesammelt von Herrn Consul Krug. *Stettiner entomologische Zeitung*, **38**, 233–245.

Espeland, M., Breinholt, J., Willmott, K.R. *et al.* (2018) A comprehensive and dated phylogenomic analysis of butterflies. *Current Biology*, **28**, 770–778.

Fruhstorfer, H. (1907) Verzeichniss der von Herrn Dr. Theodor Koch–Grünberg am oberen Waupes 1903–1905 gesammelten Rhopaloceren mit Besprechung verwandter Arten. *Stettiner entomologische Zeitung*, **68**, 207–309.

Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A. & Moritz, C. (2012) Coalescent–based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution*, **27**, 480–488.

Geiger, H.J. (1981) Enzyme electrophoretic studies on the genetic relationships of pierid butterflies (Lepidoptera: Pieridae) I. European taxa. *Journal of Research on the Lepidoptera*, **19**, 181–195.

Godman, F.D. & Salvin, O. (1889) Insecta. Lepidoptera–Rhopalocera. *Biologia Centrali Americana*. (ed. by Bernard Quaritch). Dulau and Co, London.

Gundlach, J.C. (1881) *Contribución a la Entomología Cubana. Parte primera. Lepidópteros*. G. Montiel, La Habana.

Gundlach, J.C. (1891) Apuntes para la fauna puerto–riqueña. *Anales de la Sociedad española de Historia natural*, **20**, 109–128.

Heinicke, M.P., Duellman, W.E. & Hedges, B. (2007) Major Caribbean and Central American frog faunas originated by ancient oceanic dispersal. *Proceedings of the Natural Academy of Sciences*, **104**, 10092–10097.

Illiger, J.K.W. (1801) Neue Insekten. *Magazin für Insektenkunde*, **1**, 163–208.

Iturralde-Vinent, M.A. (2003) Ensayo sobre la paleogeografía del Cuaternario de Cuba. Memorias V Congreso Cubano de Geología y Minería. La Habana.

Iturralde-Vinent, M.A. (2006) Meso–Cenozioic Caribbean palaeogeography, implications for the historical biogeography of the region. *International Geology Review*, **48**, 791–827.

Janz, N. & Nylin, S. (1998) Butterflies and plants: a phylogenetic study. *Evolution*, **52**, 486–502.

Janzen, D.H. & Hallwachs, W. (2011) Joining inventory by parataxonomists with DNA barcoding of a large complex tropical conserved wildland in northwestern Costa Rica. *PLoS ONE*, **6**, e18123. https://doi.org/10.1371/journal.pone.0018123.

Janzen, D.H., Hallwachs, W.D., Blandin, P. *et al.* (2009) Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Molecular Ecology Resources*, **9**(Suppl. 1), 1–26.

Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of theAmerican Statistical Association*, **90**, 773–795.

Keightley, P.D., Pinharanda, A., Ness, R.W. *et al.* (2015) Estimation of the spontaneous mutation rate in Heliconius melpomene. *Molecular Biology and Evolution*, **32**, 239–243.

Klots, A.B. (1929) The generic status of *Catopsilia* Hübner and *Phoebis* Hübner, with a discussion of the relationship of the species and the homologies of the male genitalia (Lepidoptera, Pieridae). *Bulletin of the Brooklyn Entomological Society*, **24**, 203–214.

Klots, A.B. (1931) A generic revision of the Pieridae (Lepidoptera), together with a study of the male genitalia. *Entomologica Americana*, **12**, 139–204.

Klots, A.B. (1951) *A Field Guide to the Butterflies of North America East of the Great Plains*. Houghton–Mifflin, Boston.

Lamas, G. (2004) Pieridae. *Atlas of Neotropical Lepidoptera. Checklist: Part 4A. Hesperioidea – Papilionoidea*, Vol. **5A** (ed. by J.B. Heppner), pp. 99–117. Association for Tropical Lepidoptera–Scientific Publishers, Gainesville.

Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012) Partition finder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.

Lavinia, P.D., Núñez, E.O., Kopuchian, C., Lijtmaer, D.A., García, N.C., Hebert, P.D.N. & Tubaro, P.L. (2017) Barcoding the butterflies of southern South America: species delimitation efficacy, cryptic diversity and geographic patterns of divergence. *PLoS ONE*, **12**, e0186845.

Leaché, A.D. & Fujita, M.K. (2010) Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B: Biological Sciences*, **277**, 3071–3077.

Lewis, D.S., Sperling, F.A.H., Nakahara, S., Cotton, A.M., Kawahara, A.Y. & Condamine, F.L. (2015) Role of Caribbean Islands in the diversification and biogeography of Neotropical *Heraclides* swallowtails. *Cladistics*, **31**, 291–314.

Luo, A., Ling, C., Ho, S.Y.W. & Zhu, C.–.D. (2018) Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, **67**, 830–846.

Matos-Maraví, P., Wahlberg, N., Antonelli, A. & Penz, C.M. (2019) Species limits in butterflies (Lepidoptera: Nymphalidae): reconciling classical taxonomy with the multispecies coalescent. *Systematic Entomology*, **44**, 745–756.

Matzke, N.J. (2013) *BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts*. R package, version 2.1. [WWW document]. URL http://CRAN.R-project.org/package=BioGeoBEARS [accessed on 15 April 2015].

Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES science gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)* (ed. by J. Towns), pp. 1–8. New Orleans, LA.

Monroe, J.L. (2016) The large Sulphurs of the Americas. The International, Gainesville Diversity Foundation.

Murillo-Ramos, L., Hernández-Torres, R., Núñez, R. & Ayazo, R. (2018) New insights on the taxonomy and phylogenetic relationships of the Neotropical genus *Phoebis* (Pieridae: Coliadinae) revealed by molecular and morphological data. *Zootaxa*, **4457**, 179–188.

Nicholson, K.E., Glor, R.E., Kolbe, J.J., Larson, A., Hedges, S.B. & Losos, J.B. (2005) Mainland colonization by Island lizards. *Journal of Biogeography*, **32**, 929–938. https://doi.org/10.1111/j.1365-2699.2004.01222.x.

Ogilvie, H.A., Bouckaert, R.R. & Drummond, A.J. (2017) StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Molecular Biology and Evolution*, **34**, 2101–2114.

Oliveira, E.G., Srygley, R.B. & Dudley, R. (1998) Do neotropical migrant butterflies navigate using a solar compass? *Journal of Experimental Biology*, **201**, 3317–3331.

Ree, R.H. & Sanmartín, I. (2018) Conceptual and statistical problems with the DEC+ J model of founder–event speciation and it comparison with DEC via model selection. *Journal of Biogeography*, **45**, 741–749.

Riley, N.D. (1975) *A Field Guide to the Butterflies of the West Indies*. Collins, London.

Röber, J.K.M. (1909) Gattung *Catopsilia*. *Die Gross–Schmetterlinge der Erde*, Vol. **5** (ed. by Alfred Seitz), pp. 85–88. Alfred Kernen Verlag, Sttutgart.

Russello, M.A. & Amato, G. (2004) *Ex situ* population management in the absence of pedigree information. *Molecular Ecology*, **13**, 2829–2840. https://doi.org/10.1111/j.1365-294X.2004.02266.x.

Schwartz, A. (1989) *The Butterflies of Hispaniola*. University of Florida Press, Gainesville.

Smith, D.S., Miller, L.D. & Miller, J.Y. (1994) *The Butterflies of the West Indies and South Florida*. Oxford University Press, New York.

Srygley, R.B. & de Oliveira, E.G. (2001) Sun compass and wind drift compensation in migrating butterflies. *Journal of Navigation*, **54**, 405–427.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post–analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

de la Torre, S.L. (1954) An annotated list of the butterflies and skippers of Cuba (Lepidoptera, Rhopalocera). *Journal of the New York Entomological Society*, **62**, 113–128.

Uit de Weerd, D.R., Robinson, D.G. & Rosenberg, G. (2016) Evolutionary and biogeographical history of the land snail family Urocoptidae (Gastropoda: Pulmonata) across the Caribbean region. *Journal of Biogeography*, **43**, 763–777. https://doi.org/10.1111/jbi.12692.

Urquhart, F.A. & Urquhart, N.R. (1976) Migration of butterflies alomg the Gulf coast of northern Florida. *Journal of the Lepidopterists' Society*, **30**, 59–61.

Wahlberg, N., Rota, R., Braby, M.F., Pierce, N.E. & Wheat, C.W. (2014) Revised systematics and higher classification of pierid butterflies (Lepidoptera: Pieridae) based on molecular data. *Zoologica Scripta*, **43**, 641–650.

Wahlberg, N. & Rubinoff, D. (2011) Vagility across *Vanessa* (Lepidoptera: Nymphalidae): mobility in butterfly species does not inhibit the formation and persistence of isolated sister taxa. *Systematic Entomology*, **36**, 362–370.

Wahlberg, N. & Wheat, C.W. (2008) Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Systematic Biology*, **57**, 231–242.

Warren, A.D., Davis, K.J., Stangeland, E.M., Pelham, J.P., Willmott, K.R. & Grishin, N.V. (2018) Illustrated Lists of American Butterflies [WWW document]. URL http://www.butterfliesofamerica.com/ [accessed on 21 November 2017].

Yang, Z. (2015) The BPP program for species tree estimation and species delimitation. *Current Zoology*, **61**, 854–865.

Yu, G., Rao, D., Matsui, M. & Yang, J. (2017) Coalescent–based delimitation outperforms distance–based methods for delineating less divergent species: the case of *Kurixalus odontotarsus* species group. *Scientific Reports*, **7**, 16124.

de Zayas, F.A. & García, I. (1965) Tres nuevas mariposas (Lepidoptera: Rhopalocera) para la fauna de Cuba. *Poeyana (A)*, **11**, 1–3.

Zhang, G., Basharat, U., Matzke, N. & Franz, N.M. (2017) Model selection in statistical historical biogeography of Neotropical insects – the *Exophthalmus* genus complex (Curculionidae: Entiminae). *Molecular Phylogenetics and Evolution*, **109**, 226–239.

Accepted 6 November 2019
First published online 4 December 2019