

**TITLE:** Polytypy and systematics: diversification of *Papilio* swallowtail butterflies in the biogeographically complex Indo-Australian Region

**RUNNING HEAD:** Systematics of *Papilio* Swallowtail Butterflies

**AUTHORS:** Jahnavi Joshi<sup>1,2,\*</sup> and Krushnamegh Kunte<sup>1,\*</sup>

**AFFILIATION:** 1. National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bellary Road, Bengaluru 560065, India.

2. Current address: Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India.

**\*CO-CORRESPONDING AUTHORS:** [jahnavi@ccmb.res.in](mailto:jahnavi@ccmb.res.in); [krushnamegh@ncbs.res.in](mailto:krushnamegh@ncbs.res.in)

**ABSTRACT:** A long-standing problem in evolutionary biology and systematics is defining patterns of diversification and speciation, which is compounded by allopatric distributions of polytypic taxa in biogeographically fragmented landscapes. In this paper we revisit this enduring systematic challenge using Mormon swallowtail butterflies (*Papilio* subgenus *Menelaides*)—an evolutionary and genetic model system. *Menelaides* is speciose and intensively sampled, with nearly 260 years of systematic study complicated by polytypy resulting from discontinuous morphological variation. This variation is structured by the mainland-island matrix of the geologically complex Indo-Australian Region, where drawing species boundaries has been difficult. We sampled variation across the biogeographic range of *Menelaides*, covering 97% of currently recognized species and nearly half of all subspecies. We generated a well-supported mito-nuclear phylogeny, on which we delineated species based on two species delimitation methods (GMYC and mPTP) and strongly supported reciprocal monophyly. These analyses showed that the true species diversity in this group may be up to 25% greater than traditional taxonomy suggests, and prompts extensive taxonomic restructuring. Biogeographic analyses showed that *Menelaides* have diversified largely in allopatry in Indo-Australian subregions by repeated dispersals across key biogeographic barriers. These results provide critical insights into the diversification process in this morphologically diverse and taxonomically complicated model group. These results will also be informative in future studies on systematics, biogeography, speciation and morphological diversification in the Indo-Australian Region—arguably the most complex geological land/seascape in the world.

**KEYWORDS:** Molecular systematics; species delimitation methods; integrative taxonomy; Wallacean biogeography; tropical biodiversity; *Menelaides* Mormon swallowtails

## INTRODUCTION:

Species are fundamental units in ecology and evolutionary biology, yet defining species boundaries has long been challenging for systematic biologists. Specifically, a key difficulty in taxonomy and systematics is defining and delineating polytypic taxa. Polytypic taxa normally arise when geographically structured, morphologically distinct but incompletely sorted lineages are identified and named as subspecies of a typically widely distributed species (Wallace 1865; Poulton 1908; Mayr 1942, 1982; de Queiroz 2020). Thus, polytypic species consist of multiple types, each defining a subspecies. Defining allopatric populations either as species or subspecies of polytypic species is taxonomically problematic. This is because reproductive isolation in such isolated populations is usually unknown, and morphological characters may vary continuously or with subtle distinction among the different allopatric populations, where delineating a species may be subjective. Large proportions of taxa of birds, fish and butterflies that have been extensively collected, studied and named are polytypic, which has created long-standing problems for taxonomists, systematists and evolutionary biologists (Mallet 2004; Andersen et al. 2014). The problem of identifying and defining polytypic species is amplified in tropical landscapes, where empirical evidence increasingly suggests the presence of cryptic and unexplored diversity across taxonomic groups (Smith et al. 2008; Joshi and Karanth 2012; Agarwal et al. 2014; Barley et al. 2015; Toussaint et al. 2015; Maddison et al. 2020). An integrative taxonomic framework, especially with multiple species delimitation methods (Sites and Marshall 2003; Pons et al. 2006; Padial et al. 2010; Fujisawa and Barraclough 2013; Zhang et al. 2013; Bouckaert et al. 2014), offers a useful tool to enumerate species diversity and to develop robust evolutionary hypotheses for polytypic taxa (Condamine et al. 2012b; Joshi and Karanth 2012; Talavera et al. 2013b, 2013a; Schwarzfeld and Sperling 2015; Toussaint et al. 2015; Matos-Maraví et al. 2019). However, a large number of taxa even in some flagship and widely popular groups such as butterflies, beetles and odonates among the super-diverse insects, remain untouched by these approaches.

A classic group for studying geographic speciation, polytypic taxa and morphological evolution in relation to speciation is *Papilio* swallowtail butterflies, which have been extensively studied in evolutionary biology and systematics (Wallace 1865; Poulton 1908; Caterino and Sperling 1999; Caterino et al. 2001; Zakharov et al. 2004; Mallet 2004; Condamine et al. 2012b, 2012a, 2013; Allio et al. 2019; Owens et al. 2020). Sexual dimorphism, polymorphism and interbreeding of different forms of *Papilio* informed early ideas of reproductive isolation, speciation, and polytypy (Mallet 2004). Since then, the systematics of *Papilio* has been subjected to multiple, almost periodic revisions in line with the current methods and evolutionary thinking of the times, initially based purely on morphological data, but more recently with molecular data and phylogenetic methods (Munroe 1960; Hancock 1983; Caterino and Sperling 1999; Condamine et al. 2013). Recent molecular phylogenetic study on the butterfly family Papilionidae revealed that the genus *Papilio* has two distinct clades, one ranging over the Old World tropics and the other largely confined to the New World tropics (Condamine et al. 2012a, 2013). Among the seven subgenera in the Old World tropical clade of *Papilio*, subgenus *Menelaides* Hübner, [1819] is the most speciose with 56 described species and c. 200 subspecies (Häuser et al. 2005). In a recent Papilionidae phylogeny, *Menelaides* as traditionally defined was found to be non-monophyletic (Condamine et al. 2013), suggesting a need for detailed taxonomic work.

In this study, we use exhaustive taxon sampling of *Menelaides* (97% of currently recognised species and 50–60% of all known subspecies), mito-nuclear data, and detailed phylogenetic analysis to propose robust species hypotheses in an integrative taxonomic framework. We then infer biogeographic patterns and processes of diversification in this diverse Indo-Australian clade. The Indo-Australian Region is most remarkable in terms of its geological history and biological diversity among the tropical landscapes. It hosts 11 of the 33 globally recognised biodiversity hotspots consisting pieces of continental crust and oceanic islands either with Gondwanan or Eurasian origins (Hall 2012). Wallace's seminal work in Malay Archipelago assessing the influence of geography on species distributions was important in the development of biogeographic theory (Wallace 1860, 1863). Inspired by that work, we here test the relative roles of dispersal and vicariance in influencing the current distribution of *Menelaides* within the Indo-Australian Region.

## MATERIALS AND METHODS:

**Taxon Sampling and DNA Sequencing:** We sampled 240 individuals representing 54 of the 56 *Papilio* (*Menelaides*) species as recently listed (Häuser et al. 2005), including ~50% of all the described subspecies (SI Table 1). The two missing *Menelaides* species were *P. lampsacus* and *P. erskinei*, which could not be sampled fresh because of their extreme rarity, and older samples did not yield DNA that could be sequenced using Sanger sequencing. All samples were either collected fresh in the field by Adam Cotton and KK or purchased from collectors by Adam Cotton, and some were donated to this study by colleagues. We extracted genomic DNA from legs and/or thoracic muscle tissue using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Polymerase chain reactions (PCR) were performed using standard primer pairs from earlier butterfly phylogenetic work (Cho et al. 1995; Brower and DeSalle 1998; Caterino and Sperling 1999; Zakharov et al. 2004; Simonsen et al. 2011). We generated approx. 4,000 bp DNA sequence data for each specimen from four mitochondrial markers (*cytochrome c oxidase I*, *cytochrome c oxidase II*, *tRNA leucine*, and ribosomal *16S* genes; total 2,623 bp) and two nuclear markers (*elongation factor I-alpha* and *wingless*; total 1,386 bp). Table S2 summarises information pertaining to each locus and primer pair. We performed PCRs with the protocol: initial 2 min denaturation at 95°C, followed by 35 cycles of 30s at 94°C, 30s at 43-60°C (depending on the primer combinations) and 1 min at 72°C, then a 4 min final extension at 72°C. We cleaned PCR products using ExoSap, and sequenced using an ABI 310 Genetic Analyzer Version 3.1 (Applied Biosystems, Foster City, USA) in the NCBS sequencing facility. Table S1 provides taxonomic and sampling details of all individuals and GenBank accession numbers of all the sequences used in this work, including the newly generated sequences (GenBank accession numbers KX557490–KX558331).

For our phylogenetic and taxonomic comparisons, we used Häuser et al. (Häuser et al. 2005) as the reference since this is the latest comprehensive catalogue of *Papilio* species. However, it should be noted that other checklists and taxonomic treatments of *Papilio* (*Menelaides*) exist (D'Abrera 1982; Beccaloni et al. 2003), which may differentially treat taxa as species or subspecies. Consequently, what we have listed as new taxonomic proposals for a more stable taxonomic arrangement in Appendix 1 based on Häuser et al. (Häuser et al. 2005) may differ with other checklists but the overall magnitude of taxonomic changes proposed and our general analyses and conclusions will remain unaffected.

**Molecular Phylogenetic Analyses:** We reconstructed a molecular phylogeny of *Papilio* to ascertain the phylogenetic position of the subgenus *Menelaides* with: (a) *Menelaides* DNA sequences generated for this study (240 individuals), and (b) published sequences of 47 individuals (Zakharov et al. 2004; Condamine et al. 2013) representing all currently recognised *Papilio* subgenera: *Achillides*, *Chilasa*, *Eleppone*, *Heraclides*, *Menelaides*, *Papilio*, *Princeps*, *Pterourus*, and *Sinoprinceps*. Genus *Papilio* comprises tribe Papilionini, so we used representatives of sister tribes as outgroups: *Teinopalpini* (*Meandrusa payeni* and *Teinopalpus imperialis*), *Troidini* (*Troides helena* and *Pachliopta* spp.), and *Leptocircini* (*Graphium sarpedon*). We used this larger dataset with 287 individuals of all *Papilio* subgenera and seven outgroup taxa (total 294 individuals) to delineate the taxonomic position and monophyly of *Menelaides* (Fig. 2, SI Fig. 1). We subsequently used this for reconstructing the species-level phylogeny of *Menelaides* and species delimitation analyses (Fig. 3, SI Fig. 2).

We used Bayesian and maximum likelihood approaches to reconstruct phylogenetic trees. We used PartitionFinder to choose the best partition scheme and corresponding model of sequence evolution for mtDNA and nuclear gene sequence data, using the *greedy* algorithm, where branch lengths were linked, and models were searched for MrBayes, BEAST and RAxML. The Bayesian Information Criterion (BIC) was used to compare the fit of different models (Lanfear et al. 2012). PartitionFinder suggested a total of four partitions with likelihood score of lnL -59913.96 and BIC 124976.40, specifically the mitochondrial markers *16S* and *COI-COII-tRNA leu* with the GTR+I+G substitution model, the nuclear marker *EF1-alpha* with the SYM+I+G substitution model, and the nuclear marker *wingless* with the K80+I+G substitution model for MrBayes and RAxML model. For BEAST model it selected GTR+I+G substitution model for *16S* and *COI-COII-tRNA leu* and HKY model for the nuclear marker *EF1-alpha*.

We performed a partitioned Bayesian analysis in MrBayes 3.2 (Ronquist et al. 2012) and set four partitions based on PartitionFinder results. We estimated base frequencies, rates for the GTR, SYM and K80 models, and the gamma distribution shape parameter, for each partition separately in MrBayes 3.2. We kept the prior distribution over the tree topologies and branch lengths at default values. The run

was carried out for 50 million generations and sampled every 1,000 generations. We used split frequency below 0.01 to assess stationarity and set burn-in using MrBayes 3.2, and built a consensus tree with the remaining trees. We also checked for effective sample size (ESS) being more than 200 in Tracer v1.4.1. For maximum likelihood tree building, we used RAxML with four partitions and GTR+I, using a web-server (<http://embnet.vital-it.ch/raxml-bb/>) with 1,000 bootstraps. Lastly, we analysed nuclear data separately for the exhaustive dataset, using Bayesian and maximum likelihood approaches with settings as above.

**Species Delimitation Analyses:** We used two widely applied species delimitation methods to define species boundaries within *Menelaides*. First, we used the Generalized Mixed Yule Coalescent (GMYC) method to identify putative species in *Menelaides* using the mitochondrial *COI-tRNA leu-COII* dataset (Pons et al. 2006). We used the R package ‘splits’ and implemented the GMYC method using both single and multi-threshold methods (Species Limits by Threshold Statistics, <http://r-forge.r-project.org/projects/splits/>; (Ezard et al. 2009)) to detect a threshold value for the transition from interspecific to intraspecific branching patterns and identify clusters. We performed this analysis on two datasets, one where we retained multiple individuals representing the same haplotype and another where we removed all identical haplotypes to resolve all zero-length branches. We used BEAST 2 to generate a time-calibrated ultrametric input tree (Bouckaert et al. 2014). Then we used the single threshold model with the GMYC method.

Second, we used a Bayesian implementation of the Multi-rate Poisson Tree Process model (mPTP) (Kapli et al. 2017). Higher Bayesian support value on a node indicates that all descendants from this node are more likely to be from one species. We used a single RAxML tree based on mitochondrial data for mPTP on the web-server (<https://mptp.h-its.org>) to run the analysis with 50,000 MCMC generations.

We also calculated pairwise genetic distance ( $p$ -distance) among sister-species using *COI-tRNA-COII* data in MEGA7 (Kumar et al. 2016).

**Divergence Time Estimation:** We obtained unique haplotypes for each species defined by species delimitation analyses for mtDNA and nuDNA using Geneious 7.1.7, and used them to reconstruct a Bayesian species time tree using BEAST2 (Bouckaert et al. 2014). We used three recently estimated secondary calibrations from a well-sampled phylogenetic analysis of *Papilio*: (a) the most recent common ancestor (TMRCA) of the genus *Papilio* (mean=30.92 my, CI=25.74–38.28 my), (b) TMRCA of the Old World *Papilio* (mean=24.8 my, CI=20.69–30.68 my), and (c) TMRCA of the New World *Papilio* (mean=28.2 my, CI=23.35–35.2 my), with normal distributions (Condamine et al. 2013). In BEAST, we set the calibrations as follows: (a) *Papilio* mean=30.92 my, sigma=1, (b) Old World *Papilio* mean=24.8 my, sigma=1, and (c) New World *Papilio* mean=28.2 my, sigma=1.

We partitioned the dataset into four partitions based on PartitionFinder results, where *16S* and *COI-COII-tRNA leu* were set to GTR+I+G model and *EF1-alpha* and *wingless* to HKY substitution models. We used a relaxed molecular clock model with an uncorrelated exponential distribution and birth-death model, which has been suggested as the most appropriate model to define the speciation process (Drummond et al. 2006). For estimates of clock models, we treated the mitochondrial loci *16S* and *COI-COII-tRNA leu* as a single locus and two nuclear markers separately. We ran the program for 200 million generations and determined convergence of the chains to the stationary distribution using the program Tracer (v1.4.1) by evaluating the effective sample size. We constructed the consensus tree in TreeAnnotator (v1.4.8) and visualized it in FigTree (v1.2.2).

**Historical Biogeography Analyses:** We performed historical biogeography analyses on species time-tree of *Menelaides* generated in BEAST using the R package “BioGeography with Bayesian (and likelihood) Evolutionary Analysis of RangeS (BioGeoBEARS)” (Matzke 2013). BioGeoBEARS compares alternative biogeographic models and approaches in a hypothesis-testing framework using maximum likelihood. Specifically, we performed dispersal-vicariance analysis (DIVA-like), dispersal-extinction-cladogenesis (DEC) and BayArea-like analyses in which probabilistic inference of ancestral geographical ranges and range expansion is evaluated in a maximum likelihood framework. We estimated the parameter ‘J’ to evaluate the role of founder speciation (i.e., whether dispersal leads to speciation), which is critical in the Indo-Australian Region as it is a complex mosaic of islands and continents.

We assigned each species identified as such from our species delimitation results to the following biogeographic subregions based on the distributional ranges of *Menelaides*: (A) Peninsular



India, (B) Continental Asia (Himalaya, mountains of southwestern China, Myanmar, Indo-China, Malay Peninsula), (C) Japan and Taiwan, (D) Sundaland (Java, Borneo, Sumatra, Palawan and nearby islands), (E) The Philippines, (F) Wallacea (Sulawesi, the Moluccas, and the Lesser Sunda Islands), and (G) Sahul Land (Papua New Guinea, East Melanesian islands, New Caledonia, Polynesia-Micronesia) (Fig. 5). We set maximum number of subregions allowed for a species in our analysis to four based on the fact that most of the widespread species were distributed in four subregions or less. We implemented in BioGeoBEARS six models (DIVA, DIVA+J, DEC, DEC+J, BayArea, and BayArea+J) and compared likelihood scores under AIC. In this analyses, we explored the role of three processes (dispersal, vicariance and extinction), which we implemented as free parameters in a maximum likelihood framework and estimated from the data. These are maximum likelihood implementations of the original parsimony-based DIVA, likelihood-based DEC in Lagrange, and Bayesian implementation of BayArea, referred to as DIVAlike, DEClike and BayArealike (Matzke 2013). However, we will refer to them as DIVA, DEC and BayArea for simplicity.

## RESULTS:

**Circumscription of *Papilio* subgenera:** The subgenus *Menelaides* was polyphyletic in the partitioned Bayesian and maximum likelihood phylogenetic analyses based on 294 individuals of all the *Papilio* species, subgenera and outgroups with approx. 4kb of mitochondrial and nuclear gene sequences. Therefore, we first circumscribed *Menelaides* as the well-supported, most inclusive monophyletic clade containing the type-species, *Papilio polytes* (Fig. 2, SI Fig. 1). Six species (*antonio*, *demolion*, *noblei*, *liomedon*, *gigon* and *euchenor*) traditionally considered to be part of *Menelaides* formed two well-supported clades well outside *Menelaides* (Fig. 2, SI Fig. 1), which we propose to include in two available but rarely used subgenus names: (1) *Araminta* Moore, 1886; type-species: *Papilio demolion* Cramer, 1776 (containing *antonio*, *demolion*, *noblei*, *liomedon*, and *gigon*), and (2) the monobasic *Euchenor* Igarashi, 1979; type-species: *Papilio euchenor* Guérin-Méneville, [1830].

*Menelaides*—as delineated—was sister to the well-supported African subgenus *Princeps* Hübner, [1807]. However, *Princeps* as traditionally classified itself was polyphyletic, consisting of three distinct monophyletic clades, as reported earlier (Condamine et al. 2013). We therefore identified the sister of *Menelaides* as the true *Princeps*, consisting of the type-species *Papilio demodocus* Esper, 1799 (Fig. 2, SI Fig. 1). We designated the second monophyletic group that was previously under *Princeps* as the “*nireus-nobilis-epiphorbas* species group”, rather than assign it to a subgenus name. This is because the reduced dataset of nuclear markers did not support this relationship (SI Fig. 1), indicating that better taxon sampling may be required before this species group may be formally assigned to a subgenus. We propose the available but rarely used subgenus name *Melindopsis* Aurivillius, 1899 (type-species: *Papilio rex*), for the third monophyletic group that was previously under *Princeps* (Fig. 2). Overall, these relationships were congruent in two different analyses that were based on the comprehensive mitochondrial and nuclear dataset (Fig. 2) or a trimmed dataset containing only nuclear markers (SI Fig. 1). This congruence suggested that the monophyletic groups we delineated as subgenera *Menelaides*, *Princeps*, *Melindopsis*, and *nireus-nobilis-epiphorbas* species group—and their relationships with other *Papilio* subgenera—were robust against the choice of markers. This detailed molecular phylogenetic analysis and formal application of subgenus names now help stabilise the subgenera of *Papilio* and species within them.

**Phylogenetic Relationships and Systematics of *Menelaides*:** We generated a robust phylogenetic hypothesis for species groups, species and subspecies within *Menelaides* by analysing our dataset in two different ways. First, we analysed the complete nuclear and mitochondrial dataset of all 294 specimens using partitioned Bayesian and maximum likelihood phylogenetic methods, where delineating species was not always straightforward (SI Fig. 2). Second, we adopted an integrative taxonomic framework in which we first analysed the complete mito-nuclear dataset of all the specimens with two species delimitation methods (Fig. 3A), and then compared geographic distributions of each strongly supported monophyletic clade in relation to known biogeographic barriers (Figs. 3b, 4). This approach resulted in a well-supported phylogeny of *Menelaides* that defined all the ‘species’ clusters from SI Fig. 2, and which also agreed closely with the expectations of taxonomic divergence based on major biogeographic barriers (Fig. 3–4). We propose this arrangement as new hypotheses for species within *Menelaides*, which supports some classical groupings but at the same time provides novel insights into species diversity in this subgenus (Fig. 3, Appendix 1), as summarised below.

*Menelaides* was composed of six species groups, in which species assignments to species groups agreed with the traditional classification of Häuser *et al.* (Häuser *et al.* 2005), except in the case of *P. protenor* (see below) (Fig. 3). However, species hypotheses based on the species delimitation analysis and strongly supported reciprocal monophyly revealed a surprising cryptic species diversity within *Menelaides*, requiring extensive taxonomic revisions (Fig. 3). As per the consensus results of the more stringent species delimitation methods, seven taxa previously treated as subspecies were well supported as distinct genetic clades at the species level, whereas 13 taxa previously treated as species were nested within well-supported species clades, requiring a total of 20 taxonomic changes among the 42 species that were delineated (Fig. 3A, Appendix 1). In comparison, strongly supported reciprocal monophyly of biogeographically cohesive taxa suggested elevation of 14 subspecies to species level, and demotion of seven species to subspecies level, requiring a total of 21 taxonomic changes among the 54 species that were delineated (Fig. 3B, Appendix 1). The cryptic diversity revealed was spread across all the species groups and biogeographic zones (Fig. 3–4), indicating that there was widespread mischaracterisation of species and subspecies in traditional (morphological) taxonomic treatments. Two good cases in point for demotion of species were *P. polymnestor* and *P. inopinatus*. Both these taxa are widely considered to be strongly differentiated species that are endemic to Sri Lanka and peninsular India (*polymnestor*) and the islands of the Maluku province of Indonesia (*inopinatus*). All our analyses showed these to be embedded within *P. agenor* and *P. aegeus*, respectively, with strong support. On the other hand, *P. agenor* and *P. chaon* are widely treated as subspecies of *P. memnon* and *P. nephelus*, respectively, but our phylogenetic analyses (Fig. 3A and 3B) support their elevation to species level. In all, approx. 40% of the delineated species require a taxonomic change by elevating or demoting species and subspecies, which has important implications for the taxonomy and systematics of *Menelaides* (Fig. 3, Appendix 1).

**Molecular Divergence and Species Delimitation in *Menelaides*:** Molecular divergence (*p*-distance) between sister clades of species within *Menelaides* varied from <1% to 8% (Fig. 5). However, there was no break in the range of values for molecular divergence between species within *Menelaides*, and subspecies within any species, forming a gradual species-to-subspecies continuum in the range of values of molecular divergence (Fig. 5). Moreover, even the level of molecular divergence of the newly delimited species and/or strongly supported monophyletic clades was well within the range shown by traditionally well-regarded species (red data points in Fig. 5). This indicated that elevation from subspecies to the newly delimited *Menelaides* species was not due to poor performance of the newly developed species delimitation methods. Instead, this suggested that the traditional taxonomic assignments of species and subspecies were not congruent with the level of molecular divergence revealed by our analysis.

**Divergence Time Estimation, Diversification and Biogeography of *Menelaides* in the Indo-Australian Region:** In ancestral area reconstruction using DEC, DIVA and BayArea methods, the models with parameter ‘J’ evaluating founder speciation events were chosen in both the likelihood ratio test and the AIC framework (Table 1). The most recent common ancestor of *Menelaides* might have been widespread in the Indo-Australian Region but its precise origin was uncertain due to low probability in the three models: according to the DEC+J (LnL=-135.88) and DIVA+J (LnL=-133.87) models, the ancestor was confined to Sundaland, and according to the BayArea+J model (LnL=-139), it occurred in continental Asia, Sahul Land and Sundaland. According to DIVA+J and DEC+J, there were three vicariance and two dispersal events during the Miocene that led to four distinct *Menelaides* lineages in three distinct biogeographic zones: (1) Sundaland (leading to *helenus* and *memnon* species groups), (2) continental Asia (leading to *memnon*, *polytes* and *nephelus* species groups), and (3) Sahul Land (leading to *gambrisius* and *fuscus* species groups). Thus, these models reconstructed a congruent dispersal/vicariance scenario across the *Menelaides* species tree. Since DIVA+J was chosen as the best model by weighted AICs, we show only DIVA+J reconstruction of ancestral areas and dispersal/vicariance history of *Menelaides* (Fig 4), with remaining models compared in Table 1.

The diversification of *Menelaides* in the geologically complex Indo-Australian Region was influenced to a great extent by dispersal events, to a lesser extent by vicariance events and range expansions, and this was true across the biogeographic subregions and the six *Menelaides* species groups (Fig. 4). Of the 29 dispersal events, 28 were forward dispersals from ancestral areas to new subregions previously unoccupied by that clade, and only one was a back-dispersal into ancestral areas

of that clade. A significant amount of *in-situ* diversification occurred in two subregions: Sahul Land and Wallacea, largely within *memnon*, *gambrius* and *fuscus* species groups (Fig. 4).

The time of diversification varied considerably between the biogeographic subregions: species diversification occurred much more (10–16 species) near the origin—which broadly corresponds with current centres of diversity—of *Menelaides* over the past 12 million years (inset table in Fig. 4). Diversification occurred much more recently (3–4 million years) and to a much lesser extent (4–6 species) near the periphery of the *Menelaides* distribution, especially in Sri Lanka-Peninsular India and Japan-Taiwan subregions, following more recent dispersals into those subregions.

## DISCUSSION:

Considerable advances have taken place in the methods and applications of molecular systematics, phylogenetics, species delimitation methods, and more recently an integration of all of this with phylogenomic data (Hillis et al. 1996; Felsenstein 2004; de Queiroz 2007; McCormack et al. 2013; Misof et al. 2014; Rannala 2015; Kawahara et al. 2019; Hime et al. 2021). Much progress has been made since the importance of universal molecular markers and the use of mito-nuclear evidence, phylogenetic methods and integrative taxonomy was emphasized (Caterino et al. 2000; Wahlberg and Wheat 2008; Dupuis et al. 2012; Papakostas et al. 2016; Yang and Rannala 2017). As a result, taxonomy and the age-old problems of defining species and subspecies are beginning to be tackled with these new methods and tools. The use of molecular systematics and integrative taxonomy hold considerable promise for solving taxonomic problems such as those posed by *Papilio* and other diverse taxa, but efforts are just catching up to do so (Kunte et al. 2011, 2019; Andersen et al. 2014; Dupuis and Sperling 2015, 2016; Toussaint et al. 2015; Matos-Maraví et al. 2019). This may help taxonomists define species in manners that are systematically robust, and that will also provide well-informed species hypotheses for evolutionary biologists to study biogeography, population divergence, potential evolution of reproductive barriers, and diversification.

On this background, *Papilio* is an illuminating example. Systematics of *Papilio* in light of the significant polytypy and even more striking morphological (wing colour pattern) differentiation has traditionally been a problem that divided taxonomic opinion and practice. The phylogenetic and molecular systematic results presented above are among the strongest applications of modern methods to resolve age-old taxonomic issues of polytypy and species definitions in *Papilio*. The strong branch supports from a mito-nuclear dataset used in our phylogeny also suggest that our molecular sampling is likely adequate for this work. The use of larger genome-wide molecular datasets would be an important improvement on the existing work, especially where our work shows a disagreement between species delimitation methods (Fig. 3A) and phylogenetically strongly supported, biogeographically meaningful monophyly (Fig. 3B). Specifically, it will be important to attempt to resolve the following four cases with phylogenomic data. First, in our phylogenetic analysis, *P. protenor* was placed at the base of the *polytes* species group, although it is usually believed to be a member of the *memnon* species group on morphological grounds. Second, our analysis placed the island taxon *merapu* inside the predominantly mainland monophyletic group that we referred to as *P. agenor*, instead of inside the predominantly island clade (*P. memnon*) to which it might belong. Third, *P. alphenor* was placed in the subgroup (*alphenor* (*phestus*, *ambrax*)) when it is expected to be in the *polytes* + *javanus* subgroup based on wing patterns and mimetic female polymorphism. Finally, the cryptic species revealed by well-supported monophyly and widely allopatric distributions in the *helenus* species group are striking. Note that these phylogenetically well-supported findings are based on dense taxon sampling and inclusion of multiple samples of each taxon traditionally treated as species (SI Table 1), and a molecular dataset that hardly had any gaps, minimizing the possibility that these phylogenetic placements were artefacts. Phylogenomic approaches may be able to further resolve these unexpected results, in addition to providing further support for diversification in the highly diverse *memnon* species group.

The complex polytypy and phylogenetic distinctiveness described above is expected to unfold over an equally complex geological land/seascape. The Indo-Australian Region forms such a geologically complex landscape, setting stage for among the most complicated biogeographic scenarios for biotic diversification in the world (Lohman et al. 2011; Bacon et al. 2013; Brown et al. 2013). Its tall, ragged mountains, deep oceanic fissures and strong oceanic currents have influenced—and in some places prevented—dispersal and divergence of distinctive lineages within this region, giving rise to a unique biogeographic mix of various mainland and island faunas (Wallace 1860, 1876; Simpson 1961;

Mayr 1965; Lohman et al. 2011; Bacon et al. 2013; Brown et al. 2013). The various biogeographic barriers responsible for this isolation in some areas and mix in others have geologically formed from plate tectonics and geo-climatic cycles. These geo-climatic cycles, along with the history of dispersals and vicariance in a fragmented landscape, bear strong signatures on the patterns of diversification of *Papilio* swallowtails, which we summarize below. The synthetic view of the evolutionary history of Cenozoic period of the Indo-Australian Region is still emerging, when studies like ours with detailed phylogenies and fine-scale biogeographic analyses across the region will be crucial for understanding diversification processes in the landscape.

Sea level changes associated with geo-climatic cycles, and movement of the continental plates, have been critical in the formation of the Indo-Australian Region and the evolution of biodiversity therein. By early Miocene (25 mya), continental plates, many island arcs and major biogeographic barriers such as the Makassar Strait and the Wallace's Line had assumed their present positions (Hall 1998). Mid-Miocene period is also recognised as the 'climatic optimum' in this area as the warming phase occurred 17–15 mya and led to the extensive growth of tropical forests throughout the Sundaland. Subsequently, there was cooling that continued until 6 mya and led to the contraction of tropical forests. This was again followed by a warming phase that continued until ~3.2 mya. During this period, tectonic activity and uplift also affected both the positioning of several islands (e.g. the Philippines) and the shape and extent of individual islands (Hall 2002). These warm-wet climate cycles and geological changes in the Miocene-Pliocene appear to have played an important role in shaping some of the evolutionary divergences in the Indo-Australian Region across taxa including plants, birds, butterflies and mammals (Meijaard 2004; Lohman et al. 2011; Brown et al. 2013).

*Menelaides* species time tree with ancestral area reconstruction suggested that *Menelaides* started diversifying in the mid- to late-Miocene (approx. 11 mya; Fig. 4) with a widespread ancestor possibly across Sundaland, continental Asia and Sahul shelf. At the time, Sundaland was still one landmass (largely Borneo) and connected to the Asian mainland, while Sumatra probably consisted of an island arc and Java was still mostly submerged (Meijaard 2004). The initial three vicariance and two dispersal events led *Menelaides* diversification across the Indo-Australian Region. Four major *Menelaides* clades leading to six species groups had started to diversify in three distinct biogeographic zones (continental Asia, Sundaland and Sahul Land) towards the late Miocene (10–8 mya; Fig. 4).

After these initial vicariance events, dispersals across biogeographic subregions dominated *Menelaides* diversification, coupled to a lesser extent with *in situ* speciation in these biogeographic subregions (inset tables in Fig. 4). On the whole, *Menelaides* in the Sahul Land largely diversified through *in situ* speciation (9–2mya), although there was inter-island dispersal within this subregion. Diversification in the Sundaland, Wallacea and continental Asia appears to have been shaped by a combination of dispersals and *in situ* speciation from mid-Miocene to Pliocene (11–2.6 mya). This finding supports the emerging view that diversification in Wallacea and Asia was shaped by events in the Miocene and Pliocene (Stelbrink et al. 2012). Two isolated oceanic island chains (the Philippines and Japan) received *Menelaides* relatively recently (mostly 5–0.5 mya) through repeated dispersals from adjoining biogeographic subregions. Similarly, in peninsular India and Sri Lanka—which were far-flung from the centre of *Menelaides* diversity and therefore acted as habitat islands on the western edge of the Oriental Region—*Menelaides* dispersed recently (5–1 mya), and then diversified through some *in situ* speciation. Throughout this process of dispersal and *in situ* speciation across all the biogeographic subregions and in all the species groups, a predominant and peculiar pattern emerges: no two sister species of *Menelaides* occur on the same islands or within the same subregions without dispersal barriers between them. Indeed, very few species even within a species group occur within the same fine-scale island groups, and most of them occur in allopatry across major dispersal barriers, often separated on mainland and islands, or across major island groups. Thus, the predominant pattern of diversification and current species ranges is that *Menelaides* diversified almost exclusively in allopatry.

Detailed phylogenetic analyses revealed that species diversity in *Menelaides* was mischaracterised by traditional taxonomic treatments of this iconic butterfly subgenus. The unexpectedly high level of cryptic clade diversity, which was spatially strongly structured in the mainland-island mosaic, showed widespread reciprocal monophyly of biogeographically meaningful clades, in almost all species groups. This work provides critical insights into the diversification process of this important group of butterflies and also a useful case study of the biogeography, morphological



diversification and speciation in this dense cluster of endangered biodiversity hotspots of the Indo-Australian Region.

**AUTHOR CONTRIBUTIONS:** KK conceived and supervised the study; JJ designed and performed phylogenetic and biogeographic analyses; JJ and KK wrote the manuscript.

**FUNDING:** This research was funded by a Ramanujan Fellowship (Dept. of Science and Technology, Govt. of India) and a research grant from NCBS to KK, and an NCBS Campus Fellowship and a DST Young Investigator Fellowship (SB/YS/LS-191/2013) to JJ.

**ACKNOWLEDGEMENTS:** We thank Adam Cotton and Felix Sperling for generous contributions of specimens/tissue samples and advice on *Papilio* and molecular systematics; Manas Samant, Ravi Umadi, Rhucha Vatturkar, Anupama Prakash and the NCBS Sequencing Facility for DNA sequencing; and Felix Sperling, Adam Cotton and Rohit Naniwadekar for comments on the manuscript. New Indian material used in this work was collected under research and voucher specimen collection permits issued by the state forest departments in Kerala (permit no. WL 10-3781/2012 dated 18/12/2012, and GO (RT) No. 376/2012/F&WLD dated 26/07/2012) and Nagaland (permit no. CWL/GEN/240/522-39 dated 14 August 2012), which is now deposited in the Biodiversity Lab Research Collections at NCBS. Additional specimens were provided by Laurie Wills.

## REFERENCES:

- Agarwal I., Bauer A.M., Jackman T.R., Karanth P. 2014. Cryptic species and Miocene diversification of Palaearctic naked-toed geckos (Squamata: Gekkonidae) in the Indian dry zone. *Zool. Scr.* 43:455–471.
- Allio R., Scornavacca C., Benoit N., Clamens A.-L., Sperling F.A.H., Condamine F.L. 2019. Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail butterfly evolution. *Syst. Biol.*
- Andersen M.J., Nyári Á.S., Mason I., Joseph L., Dumbacher J.P., Filardi C.E., Moyle R.G. 2014. Molecular systematics of the world's most polytypic bird: the *Pachycephala pectoralis/melanura* (Aves: Pachycephalidae) species complex. *Zool. J. Linn. Soc.* 170:566–588.
- Bacon C.D., McKenna M.J., Simmons M.P., Wagner W.L. 2013. Evaluating multiple criteria for species delimitation: an empirical example using Hawaiian palms (Arecaceae: *Pritchardia*). *BMC Evol. Biol.* 12:23.
- Barley A.J., Datta-Roy A., Karanth K.P., Brown R.M. 2015. Sun skink diversification across the Indian–Southeast Asian biogeographical interface. *J. Biogeogr.* 42:292–304.
- Beccaloni G., Scoble M., Kitching I., Simonsen T., Robinson G., Pitkin B., Hine A., Lyal A. 2003. The Global Lepidoptera Names Index (LepIndex). v. 12.2. .
- Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10:e1003537.
- Brower A.V.Z., DeSalle R. 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Mol. Biol.* 7:73–82.
- Brown R.M., Siler C.D., Oliveros C.H., Esselstyn J.A., Diesmos A.C., Hosner P.A., Linkem C.W., Barley A.J., Oaks J.R., Sanguila M.B., Welton L.J., Blackburn D.C., Moyle R.G., Peterson A.T., Alcal A.C., Townsend P.A., Alcalá A.C. 2013. Evolutionary processes of diversification in a model island archipelago. *Annu. Rev. Ecol. Evol. Syst.* 44:411–435.
- Caterino M.S., Cho S., Sperling F.A.H. 2000. The current state of insect molecular systematics: a thriving tower of Babel. *Annu. Rev. Entomol.* 45:1–54.
- Caterino M.S., Reed R.D., Kuo M.M., Sperling F.A.H. 2001. A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). *Syst. Biol.* 50:106–127.
- Caterino M.S., Sperling F.A.H. 1999. Papilio phylogeny based on mitochondrial cytochrome oxidase I and II genes. *Mol. Phylogenet. Evol.* 11:122–137.
- Cho S., Mitchell A., Regier J.C., Mitter C., Poole R.W., Friedlander T.P., Zhao S. 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1 alpha recovers morphology-based tree for heliothine moths. *Mol. Biol. Evol.* 12:650–656.
- Condamine F.L., Sperling F.A.H., Kergoat G.J. 2013. Global biogeographical pattern of swallowtail diversification demonstrates alternative colonization routes in the northern and southern hemispheres. *J. Biogeogr.* 40:9–23.
- Condamine F.L., Sperling F.A.H., Wahlberg N., Rasplus J.-Y.Y., Kergoat G.J., Kergoat G.J. 2012a. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. *Ecol. Lett.* 15:267–277.
- Condamine F.L., Toussaint E.F.A., Cotton A.M., Genson G.S., Sperling F.A.H., Kergoat G.J. 2012b. Fine-scale biogeographical and temporal diversification processes of peacock swallowtails (Papilio subgenus *Achillides*) in the Indo-Australian Archipelago. *Cladistics.* 1:1–24.
- D'Abrera B. 1982. Butterflies of the Oriental Region. Part I: Papilionidae, Pieridae & Danaidae. Victoria: Hill House Publishing.
- Drummond A.J., Ho S.Y.W., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.
- Dupuis J.R., Roe A.D., Sperling F.A.H. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Mol. Ecol.* 21:4422–4436.
- Dupuis J.R., Sperling F.A.H. 2015. Repeated reticulate evolution in North American *Papilio machaon* group swallowtail butterflies. *PLoS One.* 10:e0141882.
- Dupuis J.R., Sperling F.A.H. 2016. Hybrid dynamics in a species group of swallowtail butterflies. *J.*

- Evol. Biol. 29:1932–1951.
- Ezard T., Fujisawa T., Barraclough T.G. 2009. SPLITS: SPecies' LImits by Threshold Statistics. R package version 1.0-11/r29. .
- Felsenstein J. 2004. Inferring Phylogenies. Sunderland (USA): Sinauer Associates.
- Fujisawa T., Barraclough T.G. 2013. Delimiting species using single-locus data and the generalized mixed Yule coalescent approach: a revised method and evaluation on simulated data sets. Syst. Biol. 62:707–724.
- Hall R. 1998. The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. Biogeography and geological evolution of SE Asia. Backhuys Publishers, Leiden, The Netherlands. p. 99–131.
- Hall R. 2002. Cenozoic geological and plate tectonic evolution of SE Asia and the SW Pacific: computer-based reconstructions, model and animations. J. Asian Earth Sci. 20:353–431.
- Hall R. 2012. Late Jurassic–Cenozoic reconstructions of the Indonesian region and the Indian Ocean. Tectonophysics. 570–571:1–41.
- Hancock D.L. 1983. Classification of the Papilionidae (Lepidoptera): a phylogenetic approach. Smithersia. 2:1–48.
- Häuser C.L., de Jong R., Lamas G., Robbins R.K., Smith C., Vane-Wright R.I. 2005. Papilionidae – revised GloBIS/GART species checklist (2nd draft). .
- Hillis D.M., Moritz C., Mable B.K. 1996. Molecular Systematics. Massachusetts, USA: Sinauer Associates.
- Hime P.M., Lemmon A.R., Lemmon E.C.M., Prendini E., Brown J.M., Thomson R.C., Kratovil J.D., Noonan B.P., Pyron R.A., Peloso P.L. V, Kortyna M.L., Keogh J.S., Donnellan S.C., Mueller R.L., Raxworthy C.J., Kunte K., Ron S.R., Das S., Gaitonde N., Green D.M., Labisko J., Che J., Weisrock D.W. 2021. Phylogenomics reveals ancient gene tree discordance in the amphibian Tree of Life. Syst. Biol. 70:49–66.
- Joshi J., Karanth K.P. 2012. Coalescent method in conjunction with niche modeling reveals cryptic diversity among centipedes in the Western Ghats of South India. PLoS One. 7:1–12.
- Kapli P., Lutteropp S., Zhang J., Kobert K., Pavlidis P., Stamatakis A., Flouri T. 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics. 33:1630–1638.
- Kawahara A.Y., Plotkin D., Espeland M., Meusemann K., Toussaint E.F.A., Donath A., Gimnich F., Frandsen P.B., Zwick A., dos Reis M., Barber J.R., Peters R.S., Liu S., Zhou X., Mayer C., Podsiadlowski L., Storer C., Yack J.E., Misof B., Breinholt J.W. 2019. Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. Proc. Natl. Acad. Sci. 116:22657–22663.
- Kumar S., Stecher G., Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33:1870–1874.
- Kunte K., Basu D.N., Girish Kumar G.S. 2019. Taxonomy, Systematics, and Biology of Indian Butterflies in the 21st Century. In: Ramani S., Mohanraj P., Yeshwanth H.M., editors. Indian Insects. Boca Raton, Florida, USA: CRC Press. p. 275–304.
- Kunte K., Shea C., Aardema M.L., Scriber J.M., Juenger T.E., Gilbert L.E., Kronforst M.R. 2011. Sex chromosome mosaicism and hybrid speciation among tiger swallowtail butterflies. PLoS Genet. 7:e1002274.
- Lanfear R., Calcott B., Guindon S., Lanfear R., Calcott B., Guindon S., Lanfear R., Calcott B., Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29:1695–1701.
- Lohman D.J., de Bruyn M., Page T., von Rintelen K., Hall R., Ng P.K.L., Shih H.T., Carvalho G.R., von Rintelen T. 2011. Biogeography of the Indo-Australian archipelago. Annu. Rev. Ecol. Evol. Syst. 42:205–226.
- Maddison W.P., Beattie I., Marathe K., Ng P.Y.C., Kanesharatnam N., Benjamin S.P., Kunte K. 2020. A phylogenetic and taxonomic review of baviine jumping spiders (Araneae, Salticidae, Baviini). Zookeys. 1004:27–97.
- Mallet J. 2004. Poulton, Wallace and Jordan: how discoveries in Papilio butterflies led to a new species concept 100 years ago. Syst. Biodivers. 1:441–452.
- Matos-Maravi P., Wahlberg N., Antonelli A., Penz C.M. 2019. Species limits in butterflies

- (Lepidoptera: Nymphalidae): reconciling classical taxonomy with the multispecies coalescent. Syst. Entomol. 44:745–756.
- Matzke N.J. 2013. BioGeoBEARS: Biogeography with Bayesian (and likelihood) evolutionary analysis in R scripts. R Packag. version 0.2. 1.
- Mayr E. 1942. Systematics and the Origin of Species. New York, USA: Columbia University Press.
- Mayr E. 1965. Animal Species and Evolution. Cambridge (USA): Harvard University Press.
- Mayr E. 1982. The Growth of Biological Thought: Diversity, Evolution, and Inheritance. Cambridge, USA: Harvard University Press.
- McCormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., Brumfield R.T. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66:526–538.
- Meijaard E. 2004. Solving mammalian riddles. A Reconstr. Tert. Quat. Distrib. Mamm. Their Palaeoenvironments Isl. South-east Asia (PhD Thesis). Sch. Archaeol. Anthropol. Aust. Natl. Univ. Canberra.
- Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B., Ware J., Flouri T., Beutel R.G., Niehuis O., Petersen M., Izquierdo-Carrasco F., Wappler T., Rust J., Aberer A.J., Aspöck U., Aspöck H., Bartel D., Blanke A., Berger S., Böhm A., Buckley T.R., Calcott B., Chen J., Friedrich F., Fukui M., Fujita M., Greve C., Grobe P., Gu S., Huang Y., Jermiin L.S., Kawahara A.Y., Krogmann L., Kubiak M., Lanfear R., Letsch H., Li Y.Y.Y., Li Z., Li J., Lu H., Machida R., Mashimo Y., Kapli P., McKenna D.D., Meng G., Nakagaki Y., Navarrete-Heredia J.L., Ott M., Ou Y., Pass G., Podsiadlowski L., Pohl H., von Reumont B.M., Schütte K., Sekiya K., Shimizu S., Slipinski A., Stamatakis A., Song W., Su X., Szucsich N.U., Tan M., Tan X., Tang M., Tang J., Timelthaler G., Tomizuka S., Trautwein M., Tong X., Uchifune T., Walz M.G., Wiegmann B.M., Wilbrandt J., Wipfler B., Wong T.K.F., Wu Q., Wu G., Xie Y., Yang S., Yang Q., Yeates D.K., Yoshizawa K., Zhang Q., Zhang R., Zhang W., Zhang Y.Y., Zhao J., Zhou C., Zhou L., Ziesmann T., Zou S., Li Y.Y.Y., Xu X., Zhang Y.Y., Yang H., Wang J.J.J., Wang J.J.J., Kjer K.M., Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science. 346:763–767.
- Munroe E. 1960. The classification of the Papilionidae (Lepidoptera). Mem. Entomol. Soc. Canada. 92:5–51.
- Owens H.L., Lewis D.S., Condamine F.L., Kawahara A.Y., Guralnick R.P. 2020. Comparative phylogenetics of *Papilio* butterfly wing shape and size demonstrates independent hindwing and forewing evolution. Syst. Biol. 69:813–819.
- Padial J., Miralles A., De la Riva I., Vences M. 2010. Review: The integrative future of taxonomy. Front Zool. 7:1–14.
- Papakostas S., Michaloudi E., Proios K., Brehm M., Verhage L., Rota J., Peña C., Stamou G., Pritchard V.L., Fontaneto D., Declerck S.A.J. 2016. Integrative taxonomy recognizes evolutionary units despite widespread mitonuclear discordance: evidence from a rotifer cryptic species complex. Syst. Biol. 65:508–524.
- Pons J., Barraclough T., Gomez-Zurita J., Cardoso A., Duran D., Hazell S., Kamoun S., Sumlin W., Vogler A. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55:595–609.
- Poulton E.B. 1908. Essays on Evolution. 1889-1907. Oxford, UK: Clarendon Press.
- de Queiroz K. 2007. Species concepts and species delimitation. Syst. Biol. 56:879–886.
- de Queiroz K. 2020. An updated concept of subspecies resolves a dispute about the taxonomy of incompletely separated lineages. Herpetol. Rev. 51:459–461.
- Rannala B. 2015. The art and science of species delimitation. Curr. Zool. 61:846–853.
- Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61:539–542.
- Schwarzfeld M.D., Sperling F.A.H. 2015. Comparison of five methods for delimitating species in *Ophion Fabricius*, a diverse genus of parasitoid wasps (Hymenoptera, Ichneumonidae). Mol. Phylogenet. Evol. 93:234–248.
- Simonsen T.J., Zakharov E. V., Djernaes M., Cotton A.M., Vane-Wright R.I., Sperling F.A.H. 2011. Phylogenetics and divergence times of Papilioninae (Lepidoptera) with special reference to the



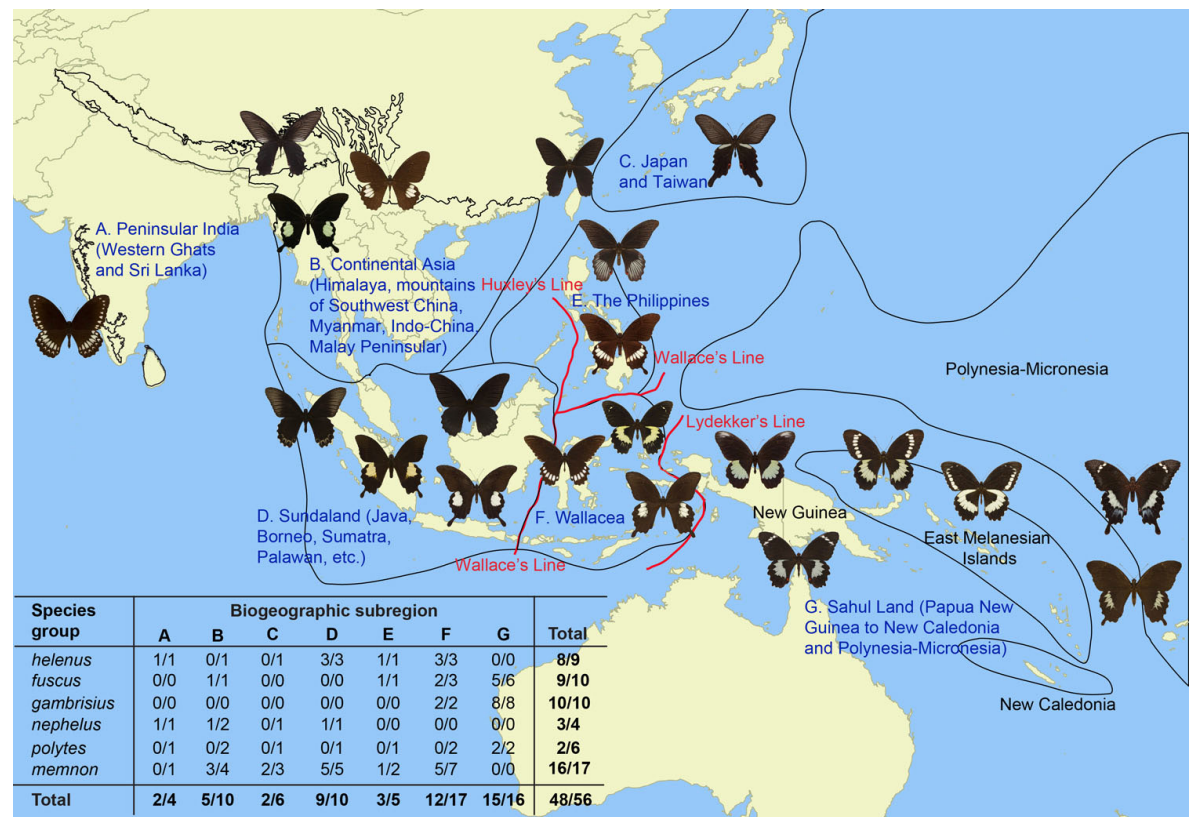
- enigmatic genera *Teinopalpus* and *Meandrusa*. *Cladistics*. 27:1–25.
- Simpson G.G. 1961. *Principles of Animal Taxonomy*. New York (USA): Columbia University Press.
- Sites J.W.J., Marshall J.C. 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends Ecol. Evol.* 18:462–470.
- Smith M.A., Rodriguez J.J., Whitfield J.B., Deans A.R., Janzen D.H., Hallwachs W., Hebert P.D.N. 2008. Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc. Natl. Acad. Sci. USA*. 105:12359–12364.
- Stelbrink B., Albrecht C., Hall R., von Rintelen T. 2012. The biogeography of Sulawesi revisited: is there evidence for a vicariant origin of taxa on Wallace’s “anomalous island”? *Evolution* (N. Y). 66:2252–2271.
- Talavera G., Dincă V., Vila R. 2013a. Factors affecting species delimitations with the GMYC model: Insights from a butterfly survey. *Methods Ecol. Evol.* 4:1101–1110.
- Talavera G., Lukhtanov V.A., Pierce N.E., Vila R. 2013b. Establishing criteria for higher-level classification using molecular data: The systematics of *Polyommatus* blue butterflies (Lepidoptera, Lycaenidae). *Cladistics*. 29:166–192.
- Toussaint E.F.A., Morinière J., Müller C.J., Kunte K., Turlin B., Hausmann A., Balke M. 2015. Comparative molecular species delimitation in the charismatic Nawab butterflies (Nymphalidae, Charaxinae, *Polyura*). *Mol. Phylogenet. Evol.* 91:194–209.
- Wahlberg N., Wheat C.W. 2008. Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Syst. Biol.* 57:231–242.
- Wallace A.R. 1860. On the zoological geography of the Malay Archipelago. *J. Proc. Linn. Soc. London. Zool.* 4:172–184.
- Wallace A.R. 1863. On the physical geography of the Malay Archipelago. *J. R. Geogr. Soc. London.*:217–234.
- Wallace A.R. 1865. On the phenomena of variation and geographical distribution as illustrated by the Papilionidae of the Malayan Region. *Trans. Linn. Soc. London*. 25:1–71.
- Wallace A.R. 1876. *The Geographical Distribution of Animals, with a Study of the Relations of Living and Extinct Faunas as Elucidating the Past Changes of the Earth’s Surface*. Vol I & II. London: Macmillan and Co.
- Yang Z., Rannala B. 2017. Bayesian species identification under the multispecies coalescent provides significant improvements to DNA barcoding analyses. *Mol. Ecol.* 26:3028–3036.
- Zakharov E. V., Caterino M.S., Sperling F.A.H. 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). *Syst. Biol.* 53:193–215.
- Zhang J., Kapli P., Pavlidis P., Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*. 29:2869–2876.

641 **Table 1:** *Menelaides* ancestral range reconstruction in BioGeoBEARS with three methods: (a)  
642 dispersal-extinction-cladogenesis (DEC), (b) dispersal-vicariance analysis (DIVA), and (c) BayArea,  
643 along with founder speciation parameter (+J). The  $\Delta$ AIC compares models with and without J (e.g.,  
644 DEC vs. DEC+J). The model with the lowest log likelihood score (DIVA + J) is marked bold.  
645

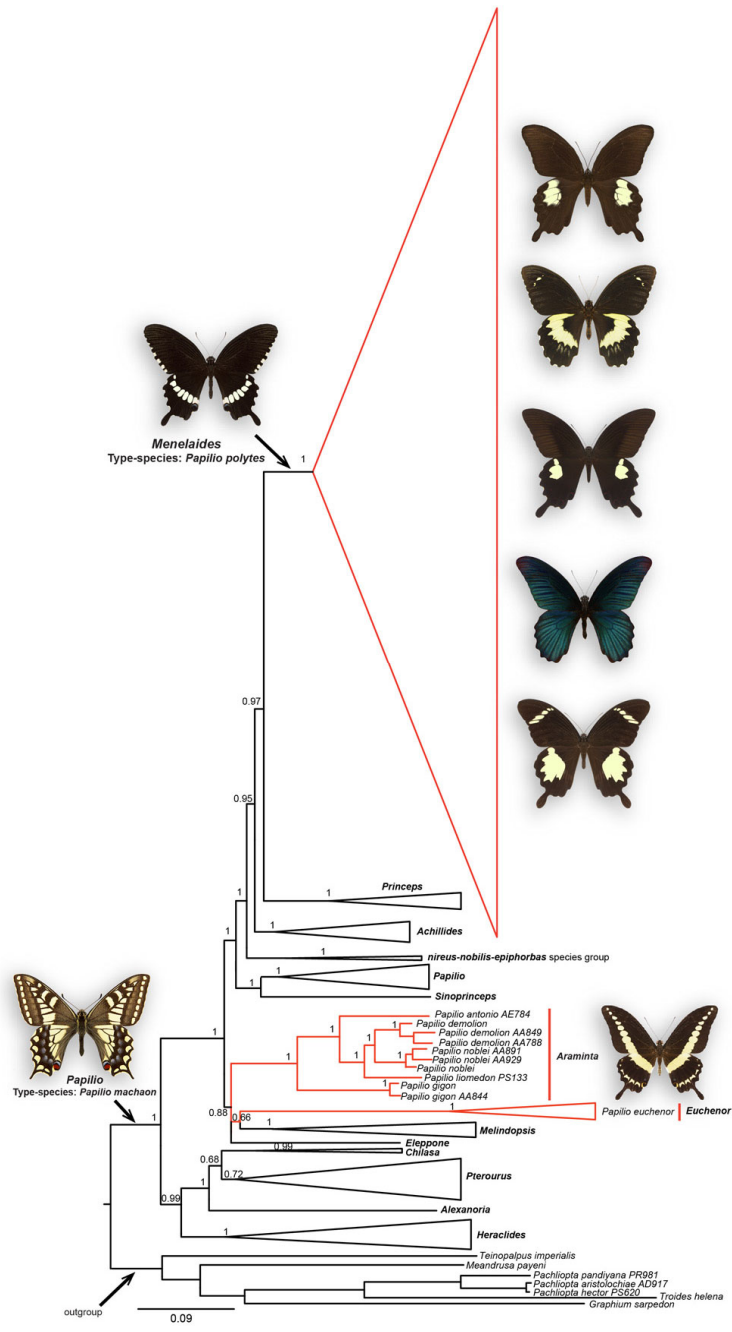
Models	Log Likelihood	No. of model parameters	AIC	$\Delta$ AIC	Dispersal rate (d)	Extinction rate (e)	Founder speciation parameter (J)
DEC	-156.8	2	317.6	39.8	0.0200059	1.76E-03	0
DEC + J	-135.9	3	277.8	0	0.0100188	1.00E-12	0.07539763
DIVA	-147.1	2	298.2	24.5	0.0242666	2.00E-09	0
<b>DIVA + J</b>	<b>-133.9</b>	<b>3</b>	<b>273.7</b>	<b>0</b>	<b>0.0123529</b>	<b>4.85E-09</b>	<b>0.06099658</b>
BAYAREA	-180	2	364.1	80.1	0.0259031	2.23E-01	0
BAYAREA + J	-139	3	284	0	0.0075915	1.00E-07	0.08948978

646

**Figure 1:** A map of the Indo-Australian Region showing globally recognised biodiversity hotspots (black outlines) and critical biogeographic barriers (red lines, and mainland-islands complexes) in the distributional range of the *Papilio* (*Menelaides*) swallowtail butterflies. We divided this landscape in seven biogeographic subregions (text in blue) as applied to distributional ranges of *Menelaides* species for biogeographic analysis. The inset table shows *Menelaides* diversity (number of endemic species/total number of species) by species groups in each biogeographic subregion, based on the total evidence presented in this paper (species delimitation analysis + strongly supported reciprocal monophyly). *Papilio erskinei* in the *gambrisius* species group (endemic to Sahul Land) and *P. lampsacus* in the *memnon* species group (endemic to Sundaland), which could not be included in this analysis, are also listed in the table.

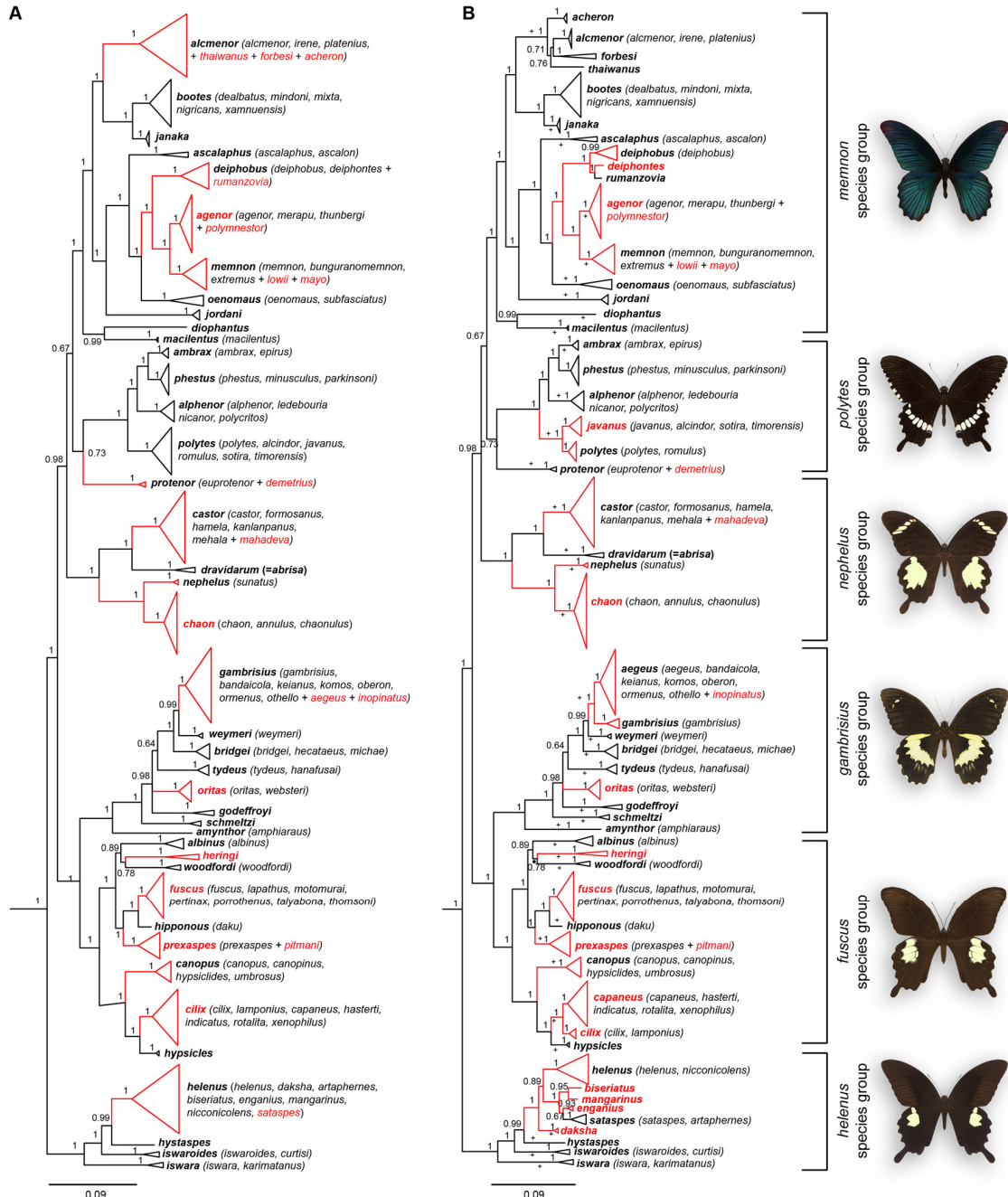


**Figure 2:** A partitioned Bayesian phylogram of *Papilio* based on four mitochondrial (*COI-tRNA<sup>Leu</sup>*, *COII* and *16S*) and two nuclear markers (*EF1-alpha* and *Wingless*), showing the arrangement of subgenera. Bayesian posterior probability is indicated at each node. The subgenus *Menelaides* as traditionally defined was polyphyletic, which required *Menelaides* to be delineated as the most inclusive monophyletic group containing the type-species, *Papilio polytes*. Species that are taken out of *Menelaides* and reassigned to two well-supported subgenera, *Araminta* and *Euchenor*, are also marked red.

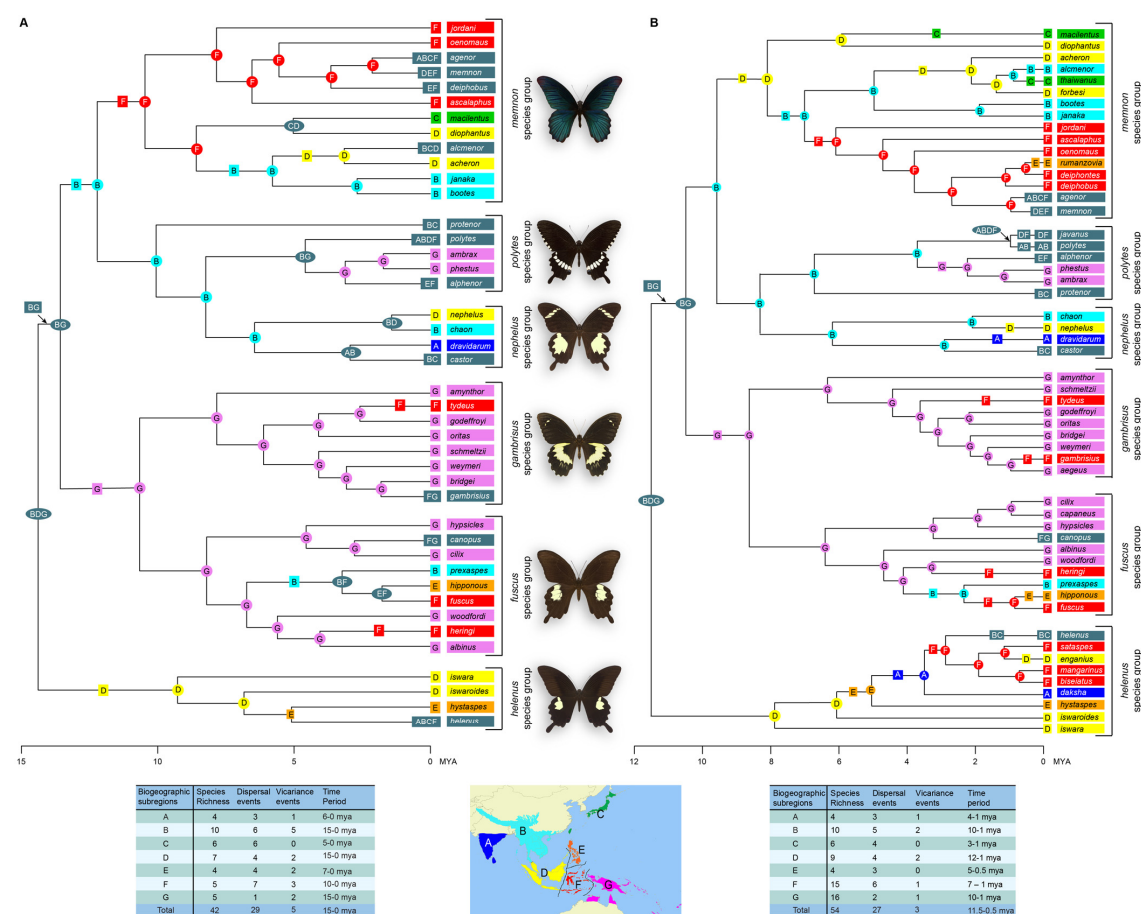




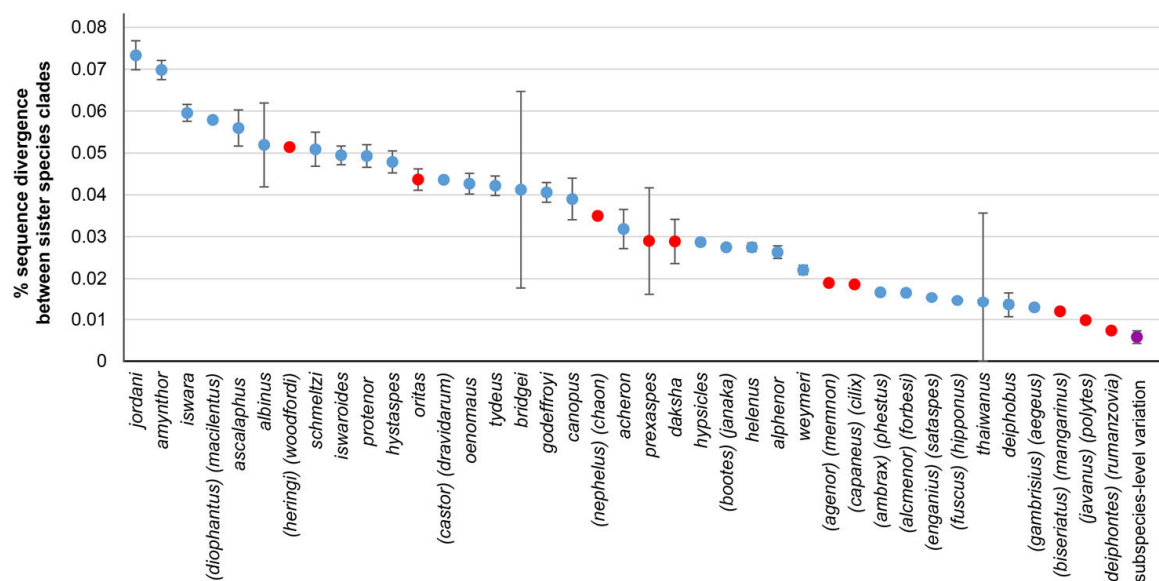
**Figure 3:** A partitioned Bayesian phylogram of the Indo-Australian *Menelaides*. **A:** Species relationships as supported by two species delimitation methods (Generalized Mixed Yule Process or GMYC, and Multi-rate Poisson Tree Process model or mPTP). **B:** Species relationships as supported by total evidence, i.e., a species delimitation method (GMYC, marked with '+') and/or strongly supported reciprocal monophyly. Bayesian posterior probability is indicated at each node. A change in taxonomic status—that is, proposals for either species demoted to subspecies status or subspecies reinstated as distinct species—is proposed for the taxa that are marked in red (also see Appendix 1). Names of terminal nodes, i.e., species names as suggested by our analyses, are marked bold, whereas subspecies names in parentheses are not marked bold. *Papilio erskinei* and *P. lampsacus* are not included in this phylogeny and in Figs. 4–5 due to the lack of molecular data.



**Figure 4:** Chronogram showing ancestral area reconstruction using DIVA+J method (the best model, see Table 1) for the Indo-Australian *Menelaides*. **A:** Ancestral area reconstruction for species as identified by species delimitation methods (Fig. 3A). **B:** Ancestral area reconstruction for species as identified by total evidence (Fig. 3B). Letters at internal nodes and their background coloured boxes/circles indicate the reconstructed ancestral area corresponding to respective biogeographic subregions. Letters at the terminal nodes indicate the current geographic distribution of species, with distribution in multiple subregions coloured dark grey. Letters in squares on branches are dispersal or vicariance events, and letters in circles at internal nodes represent estimated ancestral areas. Names of extant species and their distributions are colour-coded and boxed for easy reference. The number of dispersal and vicariance events are summarised in tables at the bottom. A=Peninsular India, B=Continental Asia, C=Japan and Taiwan, D=Sundaland, E=The Philippines, F=Wallacea, and G=Sahul Land.



**Figure 5:** Percentage sequence divergence between sister species clades of *Menelaides* (based on Fig. 3B). If specific taxa are reciprocal sister groups of each other, then both the species names are shown (e.g., (*diophantus*, *macilentus*)). If a species has a group of species as its sister, then only the name of the species which is being compared is shown (e.g., *jordani* for (*jordani*) (*heringi*, *woodfordi*)). Blue dots show percentage divergence between sister clades that are traditionally accepted as distinct species. Red dots represent taxa elevated or reinstated in this study from subspecies to species level based on Bayesian and Maximum Likelihood phylogenetic and species delimitation analyses (see Fig. 3 and Appendix 1). Error bars are computed from multiple comparisons in cases where a single species is sister to a group of species, e.g., *P. jordani*. Shown at the right end is subspecies-level molecular divergence averaged across all the species for which multiple subspecies were sequenced.



712 **Appendix 1:** Species checklist of *Menelaides* based on the analyses presented here. Taxon names that are underlined and marked bold are taxonomic changes  
713 with respect to the latest taxonomic checklist of *Menelaides* (Häuser et al. 2005).  
714

Sr. No.	Taxon name (as per original description)	Taxonomic status as given in Häuser et al. 2005	Proposed status based on species delimitation analysis using GMYC and mPTP methods (Fig. 3A)	Proposed status based on well-supported reciprocal monophyly (Fig. 3B)
<b><i>helenus</i> species group</b>				
1	<i>Papilio iswara</i> White, 1842	<i>P. iswara</i>	<i>P. iswara</i>	<i>P. iswara</i>
2	<i>Papilio iswaroides</i> Fruhstorfer, 1898	<i>P. iswaroides</i>	<i>P. iswaroides</i>	<i>P. iswaroides</i>
3	<i>Papilio hystaspes</i> C. & R. Felder, 1862	<i>P. hystaspes</i>	<i>P. hystaspes</i>	<i>P. hystaspes</i>
4	<i>Papilio (Charus) daksha</i> Hampson, 1888	<i>P. helenus daksha</i>	<i>P. helenus daksha</i>	<b><u><i>P. (helenus) daksha</i></u>*</b>
5	<i>Papilio sataspes</i> C. & R. Felder, [1865]	<i>P. sataspes</i>	<b><u><i>P. helenus sataspes</i></u></b>	<b><u><i>P. (helenus) sataspes</i></u>*</b>
6	<i>Papilio (Charus) helenus</i> var. <i>enganius</i> Doherty, 1891	<i>P. helenus enganius</i>	<i>P. helenus enganius</i>	<b><u><i>P. (helenus) enganius</i></u>*</b>
7	<i>Papilio helenus mangarinus</i> Rothschild, 1908	<i>P. helenus mangarinus</i>	<i>P. helenus mangarinus</i>	<b><u><i>P. (helenus) mangarinus</i></u>*</b>
8	<i>Papilio helenus biseriatus</i> Rothschild, 1895	<i>P. helenus biseriatus</i>	<i>P. helenus biseriatus</i>	<b><u><i>P. (helenus) biseriatus</i></u>*</b>
9	<i>Papilio helenus</i> Linnaeus, 1758	<i>P. helenus</i>	<i>P. helenus</i>	<i>P. helenus</i>
<b><i>fuscus</i> species group</b>				
10	<i>Papilio hypsicles</i> Hewitson, 1868	<i>P. hypsicles</i>	<i>P. hypsicles</i>	<i>P. hypsicles</i>
11	<i>Papilio cilix</i> Godman & Salvin, 1879	<i>P. canopus cilix</i>	<b><u><i>P. cilix</i></u></b>	<b><u><i>P. cilix</i></u></b>
12	<i>Papilio capaneus</i> Westwood, 1843	<i>P. canopus capaneus</i>	<b><u><i>P. cilix capaneus</i></u></b>	<b><u><i>P. (cilix) capaneus</i></u>*</b>
13	<i>Papilio canopus</i> Westwood, 1842	<i>P. canopus</i>	<i>P. canopus</i>	<i>P. canopus</i>
14	<i>Papilio prexaspes</i> C. & R. Felder, [1865]	<i>P. fuscus prexaspes</i>	<b><u><i>P. prexaspes</i></u></b>	<b><u><i>P. prexaspes</i></u></b>
15	<i>Papilio pitmani</i> Elwes & de Nicéville, [1887]	<i>P. pitmani</i>	<b><u><i>P. prexaspes pitmani</i></u></b>	<b><u><i>P. prexaspes pitmani</i></u></b>
16	<i>Papilio hipponous</i> C. & R. Felder, 1862	<i>P. hipponous</i>	<i>P. hipponous</i>	<i>P. hipponous</i>
17	<i>Papilio fuscus</i> Goeze, 1779	<i>P. fuscus</i>	<i>P. fuscus</i>	<i>P. fuscus</i>
18	<i>Papilio woodfordi</i> Godman & Salvin, 1888	<i>P. woodfordi</i>	<i>P. woodfordi</i>	<i>P. woodfordi</i>
19	<i>Papilio heringi</i> Niepelt, 1924	<i>P. fuscus x tydeus</i> (hybrid)	<b><u><i>P. heringi</i></u></b>	<b><u><i>P. heringi</i></u></b>
20	<i>Papilio albinus</i> Wallace, 1865	<i>P. albinus</i>	<i>P. albinus</i>	<i>P. albinus</i>
<b><i>gambrisius</i> species group</b>				
21	<i>Papilio amynthor</i> Boisduval, 1859	<i>P. amynthor</i>	<i>P. amynthor</i>	<i>P. amynthor</i>
22	<i>Papilio schmeltzi</i> Herrich-Schäffer, 1869	<i>P. schmeltzi</i>	<i>P. schmeltzi</i>	<i>P. schmeltzi</i>
23	<i>Papilio godeffroyi</i> Semper, 1866	<i>P. godeffroyi</i>	<i>P. godeffroyi</i>	<i>P. godeffroyi</i>



24	<i>Papilio oritas</i> Godman & Salvin, 1879	<i>P. aegeus oritas</i>	<u><i>P. oritas</i></u>	<u><i>P. oritas</i></u>
25	<i>Papilio tydeus</i> C. & R. Felder, 1860	<i>P. tydeus</i>	<i>P. tydeus</i>	<i>P. tydeus</i>
26	<i>Papilio bridgei</i> Mathew, 1886	<i>P. bridgei</i>	<i>P. bridgei</i>	<i>P. bridgei</i>
27	<i>Papilio weymeri</i> Niepelt, 1914	<i>P. weymeri</i>	<i>P. weymeri</i>	<i>P. weymeri</i>
28	<i>Papilio gambrisius</i> Cramer, [1777]	<i>P. gambrisius</i>	<i>P. gambrisius</i>	<i>P. gambrisius</i>
29	<i>Papilio aegeus</i> Donovan, 1805	<i>P. aegeus</i>	<u><i>P. gambrisius aegeus</i></u>	<u><i>P. (gambrisius) aegeus*</i></u>
30	<i>Papilio inopinatus</i> Butler, 1883	<i>P. inopinatus</i>	<u><i>P. gambrisius inopinatus</i></u>	<u><i>P. aegeus inopinatus</i></u>
31	<i>Papilio erskinei</i> Mathew, 1886	<i>P. erskinei</i>	Not sampled	Not sampled
<b><i>nephelus</i> species group</b>				
32	<i>Papilio chaon</i> Westwood, [1845]	<i>P. nephelus chaon</i>	<i>P. chaon</i>	<i>P. chaon</i>
33	<i>Papilio nephelus</i> Boisduval, 1836	<i>P. nephelus</i>	<i>P. nephelus</i>	<i>P. nephelus</i>
34	<i>Papilio dravidarum</i> Wood-Mason, 1880 <sup>1</sup>	<i>P. dravidarum</i>	<i>P. dravidarum</i>	<i>P. dravidarum</i>
35	<i>Papilio castor</i> Westwood, 1842	<i>P. castor</i>	<i>P. castor</i>	<i>P. castor</i>
36	<i>Papilio mahadeva</i> Moore, [1879]	<i>P. mahadeva</i>	<u><i>P. castor mahadeva</i></u>	<u><i>P. castor mahadeva</i></u>
<b><i>polytes</i> species group</b>				
37	<i>Papilio protenor</i> Cramer, [1775]	<i>P. protenor</i>	<i>P. protenor</i>	<i>P. protenor</i>
38	<i>Papilio demetrius</i> Stoll, [1782]	<i>P. demetrius</i>	<u><i>P. protenor demetrius</i></u>	<u><i>P. protenor demetrius</i></u>
39	<i>Papilio polytes</i> Linnaeus, 1758	<i>P. polytes</i>	<i>P. polytes</i>	<i>P. polytes</i>
40	<i>Papilio pammon</i> var. <i>javanus</i> C. Felder, 1862	<i>P. polytes javanus</i>	<i>P. polytes javanus</i>	<u><i>P. (polytes) javanus*</i></u>
41	<i>Papilio alphenor</i> Cramer, [1776]	<i>P. alphenor</i>	<i>P. alphenor</i>	<i>P. alphenor</i>
42	<i>Papilio phestus</i> Guérin-Ménéville, [1830]	<i>P. phestus</i>	<i>P. phestus</i>	<i>P. phestus</i>
43	<i>Papilio ambrax</i> Boisduval, 1832	<i>P. ambrax</i>	<i>P. ambrax</i>	<i>P. ambrax</i>
<b><i>memnon</i> species group</b>				
44	<i>Papilio macilentus</i> Janson, 1877	<i>P. macilentus</i>	<i>P. macilentus</i>	<i>P. macilentus</i>
45	<i>Papilio diophantus</i> Grose-Smith, 1883	<i>P. diophantus</i>	<i>P. diophantus</i>	<i>P. diophantus</i>
46	<i>Papilio jordani</i> Fruhstorfer, 1902	<i>P. jordani</i>	<i>P. jordani</i>	<i>P. jordani</i>
47	<i>Papilio oenomaus</i> Godart, 1819	<i>P. oenomaus</i>	<i>P. oenomaus</i>	<i>P. oenomaus</i>
48	<i>Papilio memnon</i> Linnaeus, 1758	<i>P. memnon</i>	<i>P. memnon</i>	<i>P. memnon</i>
49	<i>Papilio lowii</i> Druce, 1873	<i>P. lowii</i>	<u><i>P. memnon lowii</i></u>	<u><i>P. memnon lowii</i></u>
50	<i>Papilio mayo</i> Atkinson, [1874]	<i>P. mayo</i>	<u><i>P. memnon mayo</i></u>	<u><i>P. memnon mayo</i></u>
51	<i>Papilio agenor</i> Linnaeus, 1758	<i>P. memnon agenor</i>	<u><i>P. agenor</i></u>	<u><i>P. agenor</i></u>
52	<i>Papilio polymnestor</i> Cramer, [1775]	<i>P. polymnestor</i>	<u><i>P. agenor polymnestor</i></u>	<u><i>P. agenor polymnestor</i></u>

53	<i>Papilio deiphobus</i> Linnaeus, 1758	<i>P. deiphobus</i>	<i>P. deiphobus</i>	<i>P. deiphobus</i>
54	<i>Papilio deiphontes</i> C. & R. Felder, [1865]	<i>P. deiphobus deiphontes</i>	<i>P. deiphobus deiphontes</i>	<b><u>P. (deiphobus) deiphontes*</u></b>
55	<i>Papilio rumanzovia</i> Eschscholtz, 1821	<i>P. rumanzovia</i>	<b><u>P. deiphobus rumanzovia</u></b>	<i>P. (deiphobus) rumanzovia*</i>
56	<i>Papilio ascalaphus</i> Boisduval, 1836	<i>P. ascalaphus</i>	<i>P. ascalaphus</i>	<i>P. ascalaphus</i>
57	<i>Papilio janaka</i> Moore, [1858]	<i>P. janaka</i>	<i>P. janaka</i>	<i>P. janaka</i>
58	<i>Papilio bootes</i> Westwood, 1842	<i>P. bootes</i>	<i>P. bootes</i>	<i>P. bootes</i>
59	<i>Papilio alcmenor</i> C. & R. Felder, [1864]	<i>P. alcmenor</i>	<i>P. alcmenor</i>	<i>P. alcmenor</i>
60	<i>Papilio thaiwanus</i> Rothschild, 1898	<i>P. thaiwanus</i>	<b><u>P. alcmenor thaiwanus</u></b>	<b><u>P. (alcmenor) thaiwanus*</u></b>
61	<i>Papilio forbesi</i> Grose-Smith, 1883	<i>P. forbesi</i>	<b><u>P. alcmenor forbesi</u></b>	<b><u>P. (alcmenor) forbesi*</u></b>
62	<i>Papilio acheron</i> Grose-Smith, 1887	<i>P. acheron</i>	<b><u>P. alcmenor acheron</u></b>	<b><u>P. (alcmenor) acheron*</u></b>
63	<i>Papilio lampsacus</i> Boisduval, 1836	<i>P. lampsacus</i>	Not sampled	Not sampled

\* Taxon name with another name in parentheses before it, e.g., *P. (helenus) daksha\**, *P. (gambrisius) aegeus\** and *P. (polytes) javanus\**, is provisionally treated here as a subspecies based on species delimitation analyses. However, these taxa are well-supported monophyletic clades that are biogeographically meaningful; i.e., they are isolated across well-known major biogeographic barriers, which have evidently contributed to their molecular divergence and monophyly. These ‘subspecies’ clusters should be tested for phylogenetic cohesiveness and reciprocal monophyly using additional samples and a larger molecular dataset.

<sup>1</sup>: *Papilio dravidarum* Wood-Mason, 1880 is a well-established name that has been used widely in taxonomic literature over the past 150 years. *Papilio abrisa* Kirby, 1880 is considered its junior subjective synonym (Beccaloni et al. 2003), although as per the publication dates of the two descriptions, it is older and should take priority. However, *Papilio dravidarum* Wood-Mason, 1880 may be preserved for taxonomic stability due to its prevalent use.

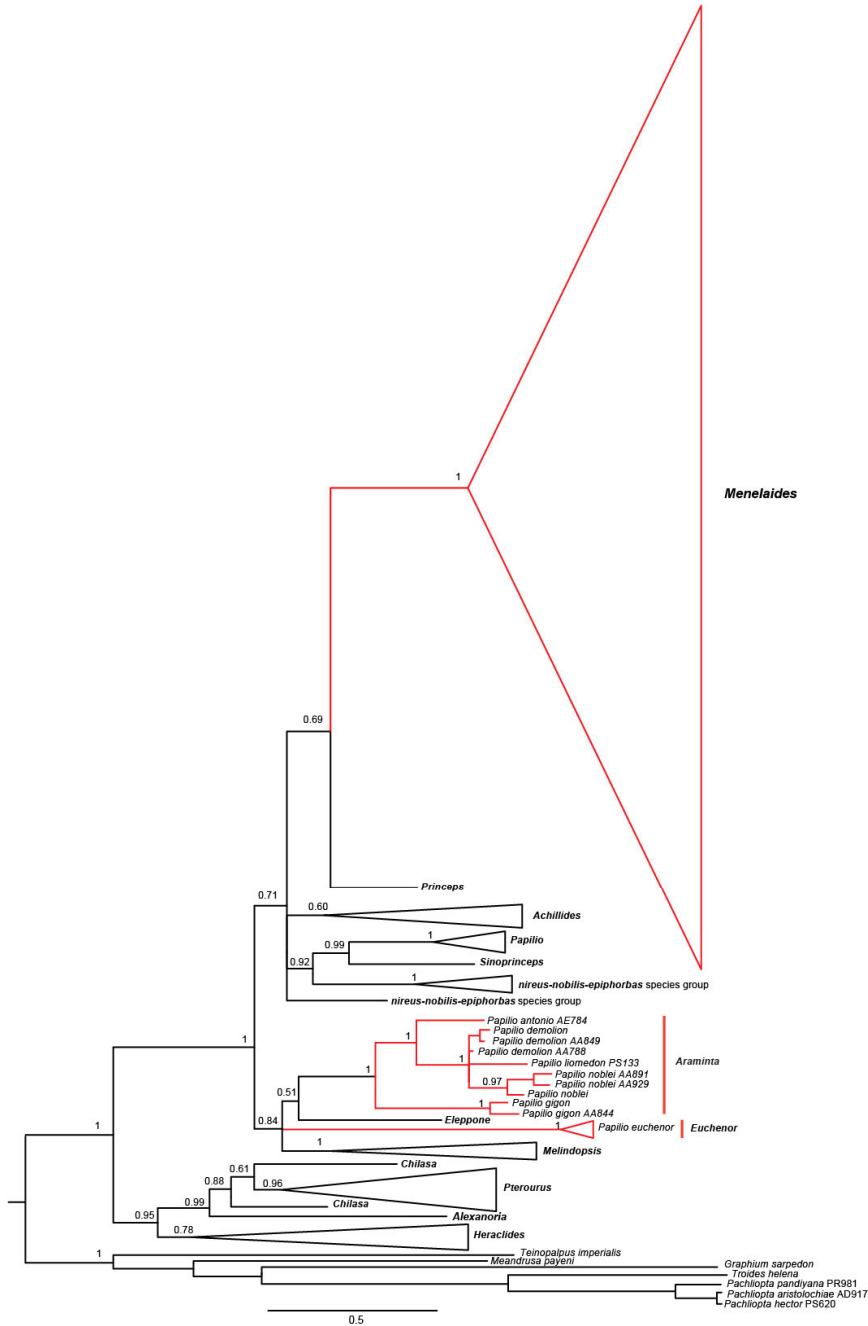
## SUPPLEMENTARY TABLES AND FIGURES

**SI Table 1:** This table follows classification of *Papilio* species as inferred from our molecular phylogenetic analyses. In other words, these are current hypotheses about species and higher taxonomic classification of *Papilio* (also see Appendix 1). Given under gene names are Genbank accession numbers. See Excel sheet, SITable1\_TaxaDetails.xlsx.

**SI Table 2:** Information about the primers and loci used in this study. See Word document, SITable2\_PrimersInfo.docx.

**SI Table 3:** Phylogenetic support (posterior probability) from the species delimitation methods—Generalized Mixed Yule Coalescent (GMYC) and Multi-rate Poisson Tree Process model (mPTP)—to delineate *Menelaides* species. Species names follow Appendix 1. See Excel sheet, SITable3\_SpeciesDelimitationResults.xlsx.

**SI Figure 1:** A partitioned Bayesian phylogram of *Papilio* based on two nuclear markers (*EF1-alpha* and *wingless*). Bayesian posterior probability is indicated at each node. Taxonomic position and monophyly of *Menelaides* as fixed by our phylogenetic analyses is marked red. Taxa that were previously treated under *Menelaides* but are now found to be outside the subgenus are also marked red.



**SI Figure 2:** A detailed Bayesian phylogram of *Menelaides* species of the Indo-Australian Region based on three mitochondrial and two nuclear markers along with voucher codes, taxa names and geographic distribution.

