WILEY-BLACKWELL

Cladistics 29 (2013) 88–111

Cladistics

10.1111/j.1096-0031.2012.00412.x

Fine-scale biogeographical and temporal diversification processes of peacock swallowtails (*Papilio* subgenus *Achillides*) in the Indo-Australian Archipelago

Fabien L. Condamine[a,*], Emmanuel F.A. Toussaint[a], Adam M. Cotton[b], Gwenaelle S. Genson[a], Felix A.H. Sperling[c] and Gael J. Kergoat[a]

[a]*INRA, UMR 1062 Centre de Biologie pour la Gestion des Populations (INRA/IRD/CIRAD/Montpellier SupAgro), Campus International de Baillarguet, CS30016, 34988 Montferrier-sur-Lez, France;* [b]*86/2 Moo 5, Ban Hua Tung, Tambon Nong Kwai, Amphoe Hang Dong, Chiang Mai 50230 Thailand;* [c]*Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada*

Accepted 7 May 2012

Abstract

Explanations for the high species diversity of the Indo-Australian Archipelago are often challenged by the region's complex climatic and geological histories. Here, we investigated the evolutionary history of swallowtail butterflies of the *Papilio* subgenus *Achillides*, comprising up to 25 recognized species and about 100 subspecies distributed across the Indo-Australian Archipelago. To estimate the relative contributions of factors influencing their biodiversity, we used DNA sequences to infer the phylogeny and species limits of 22 species including most of their subspecies. We recovered a highly resolved and well-supported phylogeny for the subgenus, and clarified some taxonomic ambiguities at the species level. The corresponding DNA-based species phylogeny was then employed to reconstruct their historical biogeography using relaxed-clock and parametric-based analyses. Molecular dating and biogeographical analyses showed that *Achillides* originated around 19 Ma in Sunda + Wallacea. Biogeographical reconstructions indicated that geological vicariance shaped the early evolutionary history of *Achillides* whereas dispersal influenced late diversification. Birth–death likelihood analyses allowed exploration of their tempo and mode of diversification. We detected several shifts in diversification rates that are attributable to past climate-induced biogeographical events. By assessing both regional and fine-scale biodiversity patterns, this study brings new findings to a biogeographical understanding of the Indo-Australian Archipelago.

The Indo-Australian Archipelago encompasses a large proportion of the Earth's biodiversity, comprising 11 biodiversity hotspots and one of the world's three remaining tropical wilderness areas (Myers et al., 2000; Mittermeier et al., 2004). Biologists and naturalists have long been fascinated by the exceptional diversity of organisms in this region (e.g. birds of paradise) and its island systems have served as natural laboratories to study the evolutionary dynamics of island colonizations and radiations. The biogeography of the Indo-Australian Archipelago has been of great interest to evolutionary biologists since Wallace (1860, 1863), who first noted a major faunal discontinuity that is today well known as Wallace's line (Lohman et al., 2011). Surprisingly, the mechanisms responsible for the region's diversification remain poorly understood (Woodruff and Turner, 2009; Klaus et al., 2010; Müller and Beheregaray, 2010; Müller et al., 2010; for a review see Lohman et al., 2011). Understanding how organisms have diversified through time in the Indo-Australian Archipelago remains a significant challenge because numerous interdependent factors have probably been involved in shaping its present-day pattern of biodiversity.

Perhaps more than any other factor, geological history has been a major driving force of biogeographical diversification (Wiens and Donoghue, 2004;

*Corresponding author:
E-mail address: fabien.condamine@gmail.com

Lomolino et al., 2010). It is difficult to assess the effects of plate tectonics on biogeographical patterns because the origin, spread and radiation of many taxa took place when the Earth's surface had a different configuration (Ree and Sanmartín, 2009). The present-day geography of the Indo-Australian Archipelago is the result of complex geological events created by the convergence of the Australian, Eurasian and Indian plates (Hall, 2002; Metcalfe, 2005). India and Australia separated from Antarctica in the Early Cretaceous and Late Cretaceous, respectively (Scotese, 2001; Sanmartín and Ronquist, 2004). The two plates gradually drifted northward toward Eurasia (Scotese, 2001). Subsequently, there were three important periods in the geological development of the Indo-Australian Archipelago. First, the collision of India into Eurasia modified plate boundaries and initiated the formation of the Himalayas and the Tibetan Plateau in the Eocene about 45 Ma (Scotese, 2001). Consequently, mountain formation resulting from the collision led to major changes in habitats, climate and drainage systems, and promoted dispersal from Gondwana via India into Southeast Asia (e.g. Klaus et al., 2010). Second, the continued indentation of Asia by India further modified Sundaland, and the geographical changes of the Malay Peninsula began in the late Oligocene around 25 Ma (Hall, 1998, 2002; Metcalfe, 2005). Meanwhile, the Australian plate collided with Southeast Asia, resulting in major changes in the configuration of plate boundaries (Hall, 1998, 2002; Metcalfe, 2005). The progressive arrival of the Australian plate also provided possible pathways for dispersal between Asia and Australia, as well as forming new barriers (Hall, 1998; Lohman et al., 2011). Moreover, this collision favoured the creation of numerous islands in the region and New Guinea formed as a result of the collision of the Australian plate with the Pacific plate (Hall, 2002; Metcalfe, 2005). Third, the Sulawesi archipelago began to form following the movements of the tectonic plates and the collision of different terranes between 24 and 13 Ma (Wilson and Moss, 1999). Some of its biota may have vicariant origins from Sundaland in the west or New Guinea/Australia to the east (Wilson and Moss, 1999). The Philippines archipelago, whose complexity is well recognized (Honza and Fujioka, 2004), formed by accumulation of terranes and creation of oceanic islands during the middle Miocene approximately 15 Ma, and was completed 5 Ma. New Caledonia was part of the eastern margin of Gondwana until the Cretaceous (Cluzel et al., 2001; Pelletier, 2006) and split from Australia around 70–65 Ma (Grandcolas et al., 2008) but remained connected to New Zealand by a remnant Gondwanan fragment called Zealandia. While drifting eastwards, New Caledonia was completely immersed (Cluzel et al., 2001; Pelletier, 2006). Near 34–37 Ma, New Caledonia engaged the subduction zone initiated by the Australian and Pacific plates, which started the obduction of the first over the second, resulting in the formation of the current archipelago (Cluzel et al., 2001; Pelletier, 2006; Grandcolas et al., 2008).

These major tectonic events and geological changes were concomitant with a burst in the diversification of many groups of organisms (e.g. Bininda-Emonds et al., 2007; Hunt et al., 2007; Wahlberg et al., 2009), suggesting that geological history probably had a profound impact on the biotic diversification of the Indo-Australian Archipelago (Lohman et al., 2011). Nevertheless, studies documenting how the historical geology of the Indo-Australian Archipelago may have shaped the evolutionary history of a variety of taxa are scarce or have generally focused on particular areas (Yagi et al., 1999; Woodruff and Turner, 2009) with few studies recognizing the implications of Wallacea itself (Müller et al., 2010). Even if the region's highly complex biogeographical pattern can be accounted for by plate tectonic activity, other factors could have contributed to the diversification of biotas. Biogeographical consequences of these factors are manifold, but most stem from the effects of plate movements (Lomolino et al., 2010). Global circulation patterns of ocean and wind currents are not only controlled by equatorial-to-polar gradients of solar radiation and temperature, but also by the relative proportion of landmasses and water (Hall, 1998; de Queiroz, 2005). Oceanic currents and regional climates have also varied through time (Zachos et al., 2001; Miller et al., 2005) due to numerous past geological events (Hall, 2002). Although such changes have probably affected biotic diversification and distribution, identifying the involved processes remains an open question (Currie et al., 2004) particularly in the Indo-Australian Archipelago (Lohman et al., 2011). Pleistocene glaciations and linked sea-level fluctuations are often invoked as an important driving factor in recent speciation or extinction events (Voris, 2000; Woodruff and Turner, 2009; Lomolino et al., 2010, Lohman et al., 2011). Finally the geographical distributions of taxa can also be influenced by biotic factors such as interactions between species. It is especially true for phytophagous organisms (Winkler et al., 2009). When herbivores are highly specific, their distributions must depend in part on the availability of appropriate host plants, and consequently the geographical range of many specific phytophagous organisms correspond to those of their plant hosts (Becerra and Venable, 1999).

Swallowtail butterflies (Papilionidae) represent an ideal group of organisms to address biogeographical questions in the Indo-Australian Archipelago (Wallace, 1865). They exhibit high species richness especially in Sundaland and Wallacea, as well as a high level of endemism in all the corresponding biodiversity hotspots (e.g. Vane-Wright and de Jong, 2003). Despite being a fascinating group, since the pioneering work of Wallace

(1865) few studies have focused on how swallowtails originated and diversified in the Indo-Australian Archipelago (but see Zeuner, 1943; Yagi et al., 1999). The genus *Papilio*, in particular, is well represented in the region with about 100 species (Häuser et al., 2005) and corresponds to a highly complex and diverse assemblage. Although the genus *Papilio* has been the subject of many molecular phylogenetic studies (e.g. Caterino and Sperling, 1999; Reed and Sperling, 1999; Zakharov et al., 2004), these studies have never included comprehensive sampling of species occurring in the Indo-Australian Archipelago (Yagi et al., 2006). To better understand the fine-scale biogeographical patterns and processes of diversification of these subgenera, a comprehensive phylogenetic study is required.

Commonly named the peacock swallowtails, *Papilio* subgenus *Achillides* is particularly suitable to exploring diversification patterns because, first, it constitutes one of the most diverse clades of swallowtails in the Indo-Australian Archipelago. The subgenus *Achillides* is recovered as sister to an assemblage of the subgenus *Menelaides* (Southeast Asia) and part of subgenus *Princeps* (Africa and Southeast Asia) (Condamine et al., 2012). *Achillides* butterflies are widely distributed in this region, extending from Pakistan through Indo-China to Japan, and from the Malay Peninsula to Queensland (Australia) and New Caledonia (Fig. 1; see the Appendix S1). Nearly all species inhabit tropical or lowland rainforests, but some of them can be found in temperate areas or at high altitudes (e.g. *P. bianor*, *P. maackii*). In addition, many *Achillides* species have limited ranges, being restricted to long-isolated islands (e.g. *P. montrouzieri* in New Caledonia) or archipelagos (e.g. *P. neumoegeni* in Sumba, *P. pericles* in Timor) (see Fig. 1 for details).

Second, *Achillides* are popular among collectors, naturalists and researchers and so their taxonomy has been well documented (e.g. Harada, 1992; Izumi, 1993; Shimogori, 1997; Yoshimoto, 1998). Nevertheless, phylogenetic relationships within the subgenus remain unknown, and the number of species is variable among authors, ranging from 21 to 25 (see Shimogori, 1997; Bauer and Frankenbach, 1998; Häuser et al., 2005 for competing classifications). Traditionally *Achillides* has been split into four species groups, namely the *paris*, *palinurus*, *peranthus* and *ulysses* groups (Jordan, 1909; Munroe, 1961; Hancock, 1983). Although numerous taxonomic studies have focused on peacock swallowtails, the status of several species is still ambiguous, as is the status of more than 100 subspecies (Shimogori, 1997; Bauer and Frankenbach, 1998). For instance, *P. ulysses* (commonly known as the blue emperor) is mainly distributed in the Papua New Guinean region and contains at least 25 subspecies, most of which are isolated island endemics.

Third, *Achillides* species present a notable level of host plant specialization as they are known to feed exclusively on Rutaceae (Igarashi, 1984; Igarashi and Fukuda, 2000), a group of plants that is well distributed throughout the Indo-Australian Archipelago (Pfeil and Crisp, 2008). The distribution of phytophagous insects is dependent on those of their host plants (Becerra and Venable, 1999; Lomolino et al., 2010). Because any *Achillides* species is able to feed on a large array of Rutaceae (see Appendix S2), this suggests that they colonized an area previously occupied by Rutaceae species long before *Achillides* (see Pfeil and Crisp, 2008; Salvo et al., 2010 for an estimate of the origin of Rutaceae; Condamine et al., 2012). Thus, these kinds of biotic interactions arguably played a lesser role in the diversification of *Achillides* swallowtails.

Together, the above data suggest that the distribution pattern of *Achillides* diversity represents an excellent candidate for investigating biogeographical history and the causes that have promoted their diversification in the Indo-Australian Archipelago. Here, we aim to: (i) use dense taxon sampling to infer a robust phylogenetic framework for *Achillides* swallowtails; (ii) delimit several ambiguous taxa to re-examine Wallace's (1865) putative species clusters and their geographical distributions; (iii) investigate the historical biogeography of *Achillides* using a molecular relaxed-clock approach; and (*iv*) test the effect of palaeoclimate on diversification rates using birth–death likelihood analyses.

Material and methods

Taxon sampling

A total of 133 *Achillides* specimens were included in our study. Our sampling covers all species recognized by Shimogori (1997), Bauer and Frankenbach (1998) and Häuser et al. (2005) except three rare and endangered species (*Papilio buddha*, *P. chikae* and *P. elephenor*). Overall, this represents up to 22 described species recognized by different authors (Shimogori, 1997; Bauer and Frankenbach, 1998; Häuser et al., 2005). In addition, we attempted to include most of the described subspecies by covering 64% of the 106 known *Achillides* subspecies (Shimogori, 1997). The specimens were directly sampled in the field or were provided by collections of various institutes. One of the co-authors (A.M.C.) conducted taxonomic identifications of swallowtails, and all corresponding voucher materials were deposited in various research institutes (see Appendix S3). Permits were obtained by two co-authors (F.L.C. and A.M.C.) for all the species classified under the IUCN Red List of Threatened Species or in the CITES list. Because the outgroup selection is a crucial step in phylogenetics (Felsenstein, 2004), we relied on the most comprehensive and recent swallowtail phylogeny (Zakharov et al., 2004; Condamine et al., 2012) to

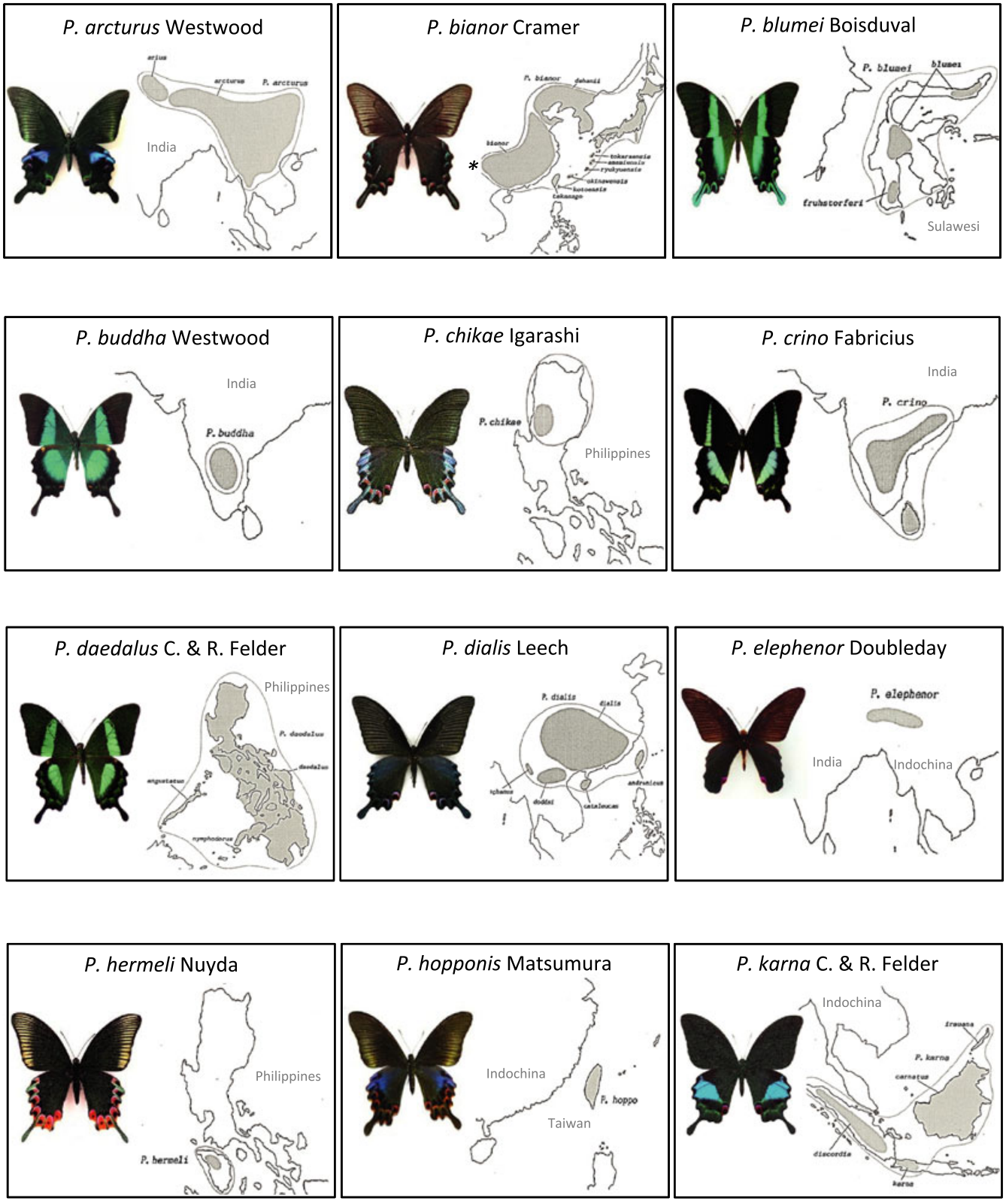


Fig. 1. Habitus of all species of the *Papilio* subgenus *Achillides* and their respective distributions (after Shimogori, 1997). For each species, the distribution of subspecies is also indicated. The asterisk indicates that the distribution of *P. bianor* now extends into the west through south-west China (*P. bianor gladiator*) to India (*P. bianor ganesa*) and Pakistan (*P. bianor polyctor*).

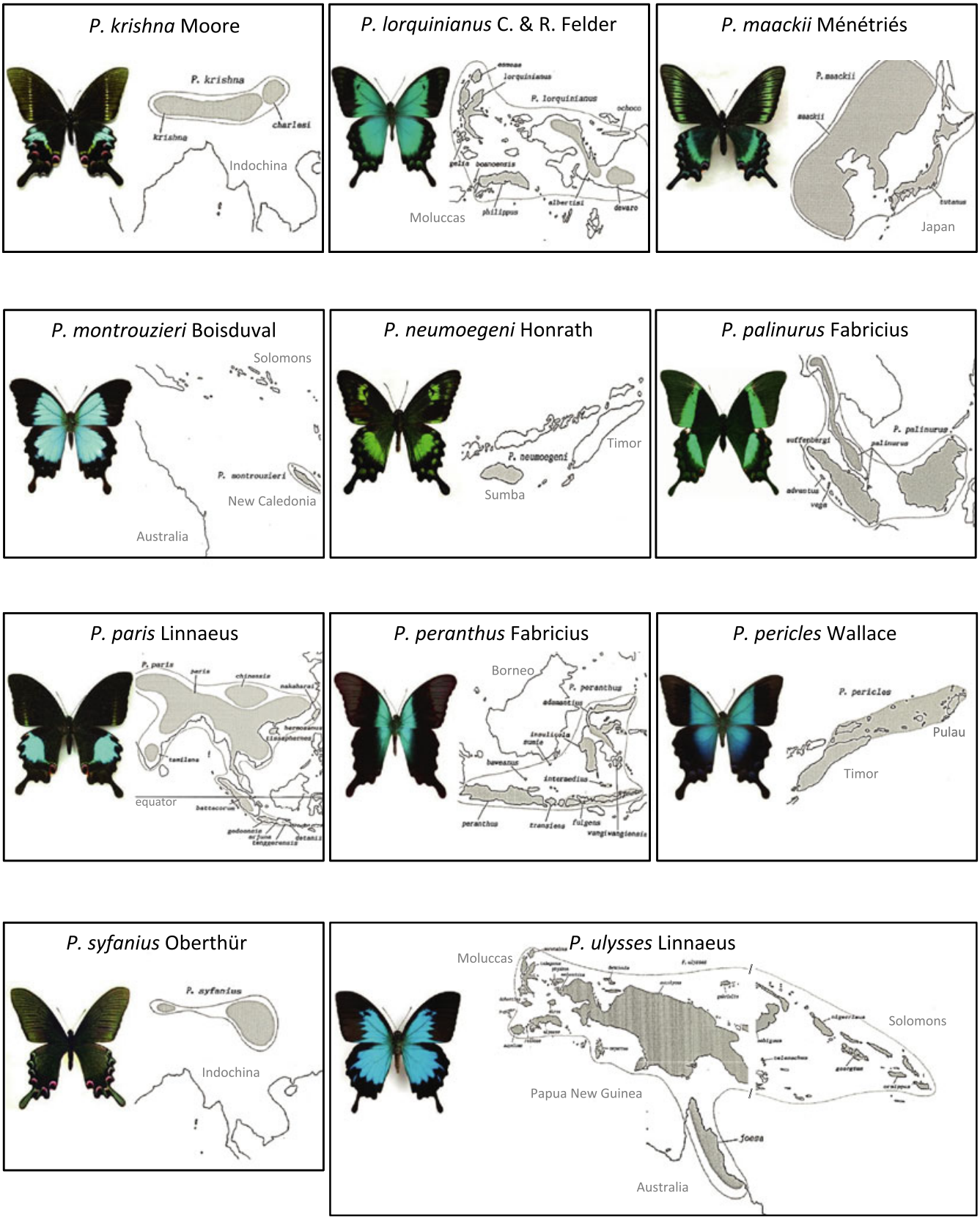


Fig. 1. (*Continued*).

choose four outgroups. These species (*Papilio demoleus*, *P. helenus*, *P. machaon* and *P. polytes*) were included as they represent subgenera that are close to the subgenus *Achillides* (Zakharov et al., 2004). Newly sampled taxa, sources of material and GenBank accession numbers for all the materials are given in Appendix S3 (new accession numbers are JQ982037–JQ982467).

Gene choice, molecular techniques and alignment

We followed previous work on swallowtail phylogeny that used DNA sequences (Caterino and Sperling, 1999; Reed and Sperling, 1999; Zakharov et al., 2004; Yagi et al., 2006) to expand the amount of analysed data available for global analyses. Here we used about 2.3 kb of three mitochondrial genes, namely cytochrome oxidase I (COI), ribosomal 16S RNA (16S) and NADH dehydrogenase 5 (ND5). We also used about 1.0 kb of the nuclear protein-coding gene elongation factor 1 alpha (EF-1α). The phylogenetic utility of these genes for Lepidoptera has been widely demonstrated (e.g. Cho et al., 1995; Zakharov et al., 2004; Simonsen et al., 2011), and a substantial dataset of lepidopteran sequences already exists for these genes. For newly sequenced species, total genomic DNA was extracted using the Qiagen® DNeasy tissue kit (Qiagen, Venlo, Netherlands). Polymerase chain reactions (PCR) were performed using the following programme: an initial 2 min denaturation at 94 °C, followed by 35 cycles of 1 min at 94 °C, 1 min at 45–52 °C (depending on primer combinations) and 1 min at 72 °C and then 7 min of final extension at 72 °C. Nucleotide sequences of the primers have been previously described and summarized (Simonsen et al., 2011). Positive PCR products were sent to Macrogen (Seoul, South Korea) and fragments were sequenced in both directions. Sequences were assembled into contiguous arrays using Geneious Pro 5.1 (available at: http://www.geneious.com/). The sequences were aligned with those retrieved from a previous study (Zakharov et al., 2004) using ClustalX 2.0 (Larkin et al., 2007) with default settings. Reading frames of COI, ND5 and EF-1α genes were checked under Mesquite 2.75 (available at: http://www.mesquiteproject.org).

Phylogenetic analyses

To estimate species relationships among *Achillides*, phylogenetic analyses were conducted using maximum parsimony (MP) and Bayesian inference (BI). Maximum parsimony analyses were carried out using TNT 1.1 (Goloboff et al., 2008). Initial heuristic searches were carried out using the Tree Bisection Reconnection (TBR) algorithm within the *Traditional Search* option in TNT, with random starting trees, 100 random-addition replicates and a *MaxTrees* value of 1000. More thorough analyses were further conducted using Random Sectorial Searches and Consensus-based Sectorial Searches (Goloboff, 1999), with the options for *Tree Ratchet*, *Tree Drifting* and *Tree Fusing* selected (Goloboff, 1999) within the *New Technology Search* option in TNT, with 100 random-addition replicates and with a *MaxTrees* value of 1000. For each analysis, 1000 non-parametric bootstrap replications were performed (standard sample with replacement).

Phylogenetic analyses were also performed with BI, which permits models of sequence evolution to be implemented (Holder and Lewis, 2003; Felsenstein, 2004; Nylander et al., 2004; Edwards, 2009). The dataset was combined and partitioned into five partitioning strategies (PS): (i) two partitions (one partition for the mitochondrial genes and one partition for the nuclear genes), (ii) four partitions (one partition per gene); (iii) four partitions (one partition per coding position plus one partition for 16S); (iv) seven partitions (one partition per coding position for the mitochondrial coding genes, one partition per coding position for the nuclear coding gene, one partition for 16S); and (v) ten partitions (one partition per coding position for each coding gene, one partition for the 16S). For each gene, the best-fit model of sequence evolution was selected with jModelTest (Posada, 2008) using both the corrected Akaike information criterion (AICc) and the Bayesian information criterion (BIC) (see Brown and Lemmon, 2007, for a discussion on the rationale for this setting).

Bayesian analyses were carried out with MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). The BI analyses settings were: two independent runs with eight Markov chain Monte Carlo (MCMC) procedures (one cold and seven incrementally heated) that ran for 10^7 generations, sampling the trees every 100th cycle and each MCMC started from a random tree. A conservative burn-in of 25% was applied after checking for stability on the log-likelihood curves and the split-frequencies of the runs. All sample trees prior to reaching these generations were discarded and remaining trees were used to generate a 50% majority rule consensus tree.

Bootstrap values (BV) in MP and posterior probabilities (PP) in BI estimated node supports for the trees. Support was considered to be strong if BV ≥ 70% in MP (Hillis and Bull, 1993; Felsenstein, 2004) and PP ≥ 0.95 in BI (Erixon et al., 2003; Felsenstein, 2004). Selection of the best-fit partitioning strategy in BI analyses was performed using Bayes Factors (BF) (Kass and Raftery, 1995; Brown and Lemmon, 2007). Estimates of harmonic mean of the likelihood values, obtained with the *sump* command in MrBayes, are used to approximate the BF. We considered that BF values > 10 significantly favour one model over another (Kass and Raftery, 1995).

Species delimitation procedure

The main objective of the species delimitation procedure is to consistently determine putative molecular species clusters by classifying the observed branching time intervals defined by the nodes in a clock-constrained phylogeny as either being the result of interspecific (diversification) or intraspecific (coalescent) processes of lineage branching (Pons et al., 2006). Although this approach can be used to clarify taxonomic ambiguities, here we focus on using this method to delimit species clusters to better assess the geographical and temporal diversification pattern and processes in the evolution of *Achillides*.

For comparison purposes, these analyses were performed with the combined dataset (i.e. the phylogeny obtained with all the genes and the best-fit partitioning strategy), with an extended COI-barcode marker only (Hebert et al., 2003) and with the EF-1α nuclear gene only. From a specific topology, we inferred an arbitrary ultrametric tree with the mean path length algorithm using PATHd8 (Britton et al., 2007), we then used the General Mixed Yule Coalescent (GMYC) to perform analysis of species delimitation (Pons et al., 2006). The method is implemented in script codes running under the R 2.13 package using functions from the *ape* library (Paradis et al., 2004) and the *splits* library (available at http://www.r-forge.r-project.org/projects/splits/). First, we visualized the branching time data by plotting the log of the number of lineages through time. A transition in branching rates should be visible as a sudden increase in slope of the plot towards the present. Second, we ran the threshold version of the GMYC model on the calibrated ultrametric tree of *Achillides*. Third, we compared the likelihood with that obtained assuming a single branching process for the tree. For the GMYC model, assuming all species have the same parameter values, the threshold model has five parameters (λ1, λ2, p1, p2 and T), whereas the null model has two (λ1 and p1); hence, there are three degrees of freedom for the comparison.

The delimitation of molecular species was used to build a "reduced dataset". This new dataset comprised a single sequence for each recovered species cluster. A sequence for a species cluster was obtained from the consensus of all the sequences constituting a species cluster using Geneious 5.1 and default settings. This reduced dataset was used to perform MP and BI analyses with the corresponding aforementioned settings. This topology was used for estimates of divergence times and analyses of geographical range evolution.

Bayesian estimates of divergence times

The applicability of a molecular clock was preliminarily investigated for each node using PATHd8 (Britton et al., 2007). As the hypothesis of a molecular clock was not statistically supported for our dataset, methods of dating that account for rate variation across lineages were used (Thorne et al., 1998; Sanderson, 2002; Drummond et al., 2006). Here we implement Bayesian relaxed-clock (BRC) analyses, which use MCMC procedures to approximate the posterior distribution of rates and divergence times and simultaneously infer their credibility intervals. These molecular relaxed-clock methods have also introduced flexible techniques for incorporating calibrations leading to a discussion about approaches to calibrating estimates of divergence times (Yang and Rannala, 2006; Hug and Roger, 2007; Ho and Phillips, 2009). In our study, the BRC analyses were carried out with the program BEAST 1.6.2 (Drummond and Rambaut, 2007), which uses the log-normal model of Thorne et al. (1998), which is uncorrelated in BEAST. The xml-file for the BEAST analysis was created under the interface BEAUti (included in the BEAST package) with the following non-default settings and priors: the *Site Model* was set based on the model used in the original Bayesian analyses, the *Clock Model* was set to a relaxed clock with uncorrelated rates, the *Tree Model* was set to a Yule process of speciation, and the *MCMC parameters* were fixed to 2.5×10^7 or 5×10^7 generations with sampling every 1000 generations and the first 25% discarded as burn-in. The remaining parameters were left as default. Several independent analyses have been performed to check the likelihood scores under the LogCombiner and Tracer programs (included in the BEAST package) using the effective sample size criterion (Drummond et al., 2006).

In BRC analyses, a fundamental step is the choice of calibration points to adjust the molecular clock (Ho and Phillips, 2009). The selection and parameterization of calibration points is sensitive especially when considering further biogeographical analyses in the study (problem of circularity). Geological calibration, in particular, represents maximum ages for nodes (Ho and Phillips, 2009), but using this calibration is very dangerous because taxa might well have begun to diverge before the geological separation and also because geological reconstructions are not immutable (Heads, 2005; Sanmartín et al., 2008). For swallowtails, the fossil record is scarce and the three known unambiguous papilionid fossils do not belong to the genus *Papilio* (Condamine et al., 2012). We chose to use previous results on divergence times of the entire family Papilionidae based on the most comprehensive taxon sampling and molecular dataset, including all the available fossils (Condamine et al., 2012). The age and divergence times of Papilionidae were inferred using fossils as minimum age constraints bounded with maximum age corresponding to the origin of angiosperms (estimated from molecular dating analyses, *sensu* Bell et al., 2010). Thus, we followed the recommendations of several authors

(e.g. Yang and Rannala, 2006; Hug and Roger, 2007; Ho and Phillips, 2009) who advocated the use of a uniform distribution or normal distribution for secondary calibrations. As we recovered a similar phylogenetic pattern, we choose to use previous age estimates for the nodes that correspond to similar common ancestors. Five calibration points were retrieved and were set to uniform prior. The root was constrained to be within 19.17–28.61 Ma, the node between *P. demoleus* and *Achillides* was set to 16.74–25.24 Ma, that of *P. demoleus* and *P. helenus* constrained to 15.87–24.02 Ma, that of *P. helenus* and *P. polytes* set to 11.6–17.88 Ma, and finally the crown of *Achillides* was constrained to be within 13.09–21.85 Ma (for details see Condamine et al., 2012).

Historical biogeography

Ancestral areas and geographical speciation scenarios for *Achillides* were inferred using the dispersal–extinction–cladogenesis (DEC) model (Ree et al., 2005; Ree and Smith, 2008) of range evolution under the program Lagrange (available at: http://www.code.google.com/p/lagrange/). We chose to use this method because the DEC model describes ancestor–descendant transitions between geographical ranges by anagenetic and cladogenetic processes such as range expansion (dispersal), range contraction (local extinction) and range subdivision/inheritance (vicariance). When the ancestral range includes two or more areas, DEC requires that between-area vicariance events separate a single area from the remainder of the ancestral range. Furthermore, unlike DIVA, DEC allows speciation within one area of a wide ancestral range, which results in one descendant occupying the area of speciation whereas the other inherits the entire range (peripatric speciation). The method specifies the likelihood of species-range data arrayed at the tips of a phylogenetic tree as a function of rates of dispersal and local extinction (Ree et al., 2005; Ree and Smith, 2008), hereafter called the ML-DEC method.

The distribution of *Achillides* extends across the Oriental and Australasian regions. We divided these two biomes into smaller biogeographical identities to get more resolution in the inference of the ancestral area for the root. For that we defined our areas using paleogeography arguments with tectonic reconstructions (e.g. Scotese, 2001; Hall, 2002; Metcalfe, 2005; Pelletier, 2006), using biodiversity hotspot information (Myers et al., 2000; Mittermeier et al., 2004), and the present-day distribution of *Achillides* swallowtails (Shimogori, 1997). The model comprised 11 component areas delimited and listed in Appendix S1 (species richness is indicated for each area).

Species ranges were defined by presence–absence data by excluding marginal distribution or human introduction (Hines, 2008; Nylander et al., 2008). The 11 areas yield a set of $2^{11} = 2048$ theoretically possible geographical ranges (area subsets), many of which were excluded from consideration based on the biological implausibility of their spatial configurations (e.g. wide disjunction between Japan and New Caledonia). The goal is to define a palaeogeological model taking into account the geological history of the Earth with biologically plausible ranges (i.e. adjacent areas). We also discarded ranges larger than three areas in size that were not subsets of observed species ranges. The motivation for this step was to create the best-fit biogeographical model in relation to the known geological history of Earth. One benefit of this approach is to reduce the dimensions of the model's transition matrix, thus increasing its computational feasibility.

Following the principles described by Ree and Smith (2008), we constructed temporal constraints on rates of dispersal between areas based on palaeogeographical reconstructions of area position through time (e.g. Scotese, 2001; Hall, 2002; Metcalfe, 2005; Pelletier, 2006). These constraints were implemented as a series of four time slices. All time slices together spanned the past 20 Ma, with each slice being 5 Ma in duration. For each time slice, we constructed a matrix of scaling factors (between zero and 0.5) for the dispersal rate between areas according to their geographical position, interpreting greater distances and/or the extent of geographical barriers (e.g. sea straits, mountain chains) as being inversely proportional to the expected rate.

An obvious advantage of the ML-DEC method in biogeography is the statistical framework permitting local optimization for the root (Ree and Smith, 2008). We performed specific tests in which the root was constrained to be either a single area, or the union of two and three areas. A 2-log likelihood unit's threshold was used to choose which area was supported (Ree et al., 2005; Ree and Smith, 2008). This likelihood framework also allows specific geographical scenarios along the chronogram.

Diversification rates

To investigate the tempo and mode of species diversification rates over time, we followed a step-by-step procedure under the R 2.13 package with the *ape* (Paradis et al., 2004) and *laser* libraries (Rabosky, 2006b). Accumulation of lineages through time was graphically visualized using lineages-through-time (LTT) plots for all species. Second, we estimated the overall diversification rate (speciation minus extinction) under a simple birth–death model (Magallón and Sanderson, 2001), with net diversification rates being calculated for three relative extinction rates ($\varepsilon = 0/0.5/0.9$). We then tested the null hypothesis that per-lineage speciation and extinction rates have remained constant through time with the γ-statistics

under the pure birth (Yule) process (constant speciation rate; Pybus and Harvey, 2000). Fourth, we tested the hypothesis of rate shifts using a likelihood-based method to compare models with a constant diversification rate with those with one or more rate shifts (Rabosky, 2006a). Single-rate and two-rate models were compared by AIC and likelihood ratio tests (Rabosky, 2006a). For all analyses the tests were also conducted using simulated incomplete phylogenies with a Monte Carlo constant-rates test to account for the effects of taxon sampling on our conclusions, namely the lack of three species (*P. buddha*, *P. chikae* and *P. elephenor*). Finally, we tested the hypothesis of rate shifts occurring at major decreases or increases on the LTT plot by using a likelihood-based method to compare models with a constant diversification rate with those with one or more rate shifts (Rabosky, 2006a; Rabosky and Lovette, 2008). We did not include extinction rate in this model because estimates under a constant rate birth–death model tended toward 0, and it is difficult to obtain meaningful estimates of separate extinction and speciation parameters under discrete-shift models similar to those considered here (confidence intervals on extinction rates are very large; Rabosky, 2006a; Rabosky and Lovette, 2008).

Table 1
Selection of substitution model of sequence evolution using both the Bayesian information criterion (BIC) and the corrected Akaike information criterion (AICc)

Partition	BIC	AICc
All combined genes	GTR + G	GTR + G
16S rRNA	HKY + G	HKY + G
CO1	GTR + G + I	GTR + G + I
ND5	TrN + G	TrN + G
EF-1α	TrNef + G	TrNef + G
All coding genes (1st position)	TrN + G	GTR + G
All coding genes (2nd position)	HKY + G	TPM1uf + G
All coding genes (3rd position)	TPM1uf + G	GTR + G
Mitochondrial coding genes (1st position)	TrN + G	TrN + G
Mitochondrial coding genes (2nd position)	HKY + G	HKY + G
Mitochondrial coding genes (3rd position)	TPM1uf + G	TPM1uf + G
CO1 (1st position)	TrN + G	TrN + G
CO1 (2nd position)	HKY + G	F81 + G
CO1 (3rd position)	GTR + G	TPM1uf + G
ND5 (1st position)	HKY + G	HKY + G
ND5 (2nd position)	F81	JC
ND5 (3rd position)	TrN + G	TrN + G
EF1a (1st position)	F81	F81
EF1a (2nd position)	JC	JC
EF1a (3rd position)	HKY + G	K80 + G

Results

Phylogenetic relationships

The molecular matrix comprises 3259 nucleotides for 137 individuals (133 *Achillides* and four outgroups; see Appendix S3). We sequenced 67 subspecies among the 106 that are recognized in the most comprehensive work of Shimogori (1997). Under MP, Traditional Search analyses using the TBR algorithm in TNT yield 60 equally parsimonious trees (length = 2227). Only 13 equally parsimonious but not shorter trees (length = 2227) were recovered using the New Technology Search and the various algorithms in TNT. Both MP analyses recovered a similar phylogenetic pattern. The selection of the substitution model of sequence evolution is described for each partition in Table 1. In the BI analyses, the partitioning strategy with ten partitions was selected as the best-fit strategy through the BF comparisons (Table 2). The different PS yield similar topologies that differ only by the terminal branching of some taxa, and the low average split frequencies values revealed good convergence between the two runs (Table 2).

Combined and partitioned BI and MP analyses found similar phylogenetic relationships (Fig. 2). Overall, for MP and BI analyses, nodes have moderate to strong support values (BV and PP). Most importantly, almost all internal nodes and all the species nodes are supported by BV > 70% and PP = 1.0 (low supports are only recovered at the intraspecific level, Fig. 2). The subgenus *Achillides* is recovered as monophyletic with strong support (BV > 70% and PP = 1.0). Within the subgenus, we recovered a phylogenetic pattern with four clades labelled C1–C4 in Fig. 2. First, a two-species clade (*P. montrouzieri* and *P. ulysses*) was found in a sister position to the remaining *Achillides* species with high support (BV > 70% and PP = 1.0; C1 in Fig. 2). The species *P. crino*, labelled C2, was recovered in a sister position to a large clade (comprising C3 and C4) in all methods used, as opposed to its traditional placement within the *palinurus* group, although support for that position was moderate to strong (BV = 63% and PP = 1.0). The third and fourth clades were always recovered as sister groups in all MP and BI analyses (BV = 63% and PP = 0.92). Clade C3 (BV = 74% and PP = 1.0) comprises *P. blumei*, *P. daedalus*, *P. lorquinianus*, *P. neumoegeni*, *P. palinurus*, *P. peranthus* and *P. pericles*. Clade C4 (BV = 90% and PP = 1) encompasses the remaining species (*P. arcturus*, *P. bianor*, *P. dehaanii*, *P. dialis*, *P. hermeli*, *P. hopponis*, *P. karna*, *P. krishna*, *P. maackii*, *P. paris* and *P. syfanius*). Only two branching differences were recovered between MP and BI analyses (indicated by NA on Fig. 2). In MP, *P. krishna* is sister to *P. maackii* + *P. syfanius* (BV < 50%), whereas it is sister to *P. arcturus* + *P. hermeli* in

Table 2
Scores obtained by Bayesian Inference with different partitioning strategies and tests on the monophyly of some species

	Harmonic mean (BIC)	Harmonic mean (AICc)	No. of parameters	BF
Partitioning strategies (PS)				
PS 1 (two partitions)	17 457.15	17 457.15	25	**
PS 2 (four partitions)	17 104.00	17 104.00	49	**
PS 3 (four partitions)	16 847.22	16 841.34	49	**
PS 4 (seven partitions)	16 781.86	16 928.30	85	Best
PS 5 (ten partitions)	16 782.89	17 113.78	121	n.s.
Tests on the monophyly of some species				
P. paris monophyletic	16 788.24	16 941.82	85	*
P. maackii and *P. syfanius* monophyletic	16 804.88	16 954.13	85	**

The first part of the table comprises the results of the different Bayesian analyses using several partitioning strategies (PS 1–5). Harmonic means and number of parameters are used to estimate Bayes Factors (BF) to select the best-fit partitioning strategy for the dataset. The second part of the table corresponds to tests on the monophyly of some species recovered as paraphyletic with the best-fit PS. We constrained *P. paris* and *P. maackii* + *P. syfanius* to be monophyletic. As for the first part, the same scores are summarized for the three constrained analyses. To estimate BF for each analysis, we used the following formula: $BF = 2 \times (\ln L1 - \ln L0) + (P1 - P0) \times \ln(0.01)$ (Kass and Raftery, 1995).

n.s., non-significant; *, significant; **, very significant.

BI (PP = 0.47). *Papilio hermeli* is sister to the *bianor* group in MP (BV < 50%), but sister to *P. dialis* in BI (PP = 0.84).

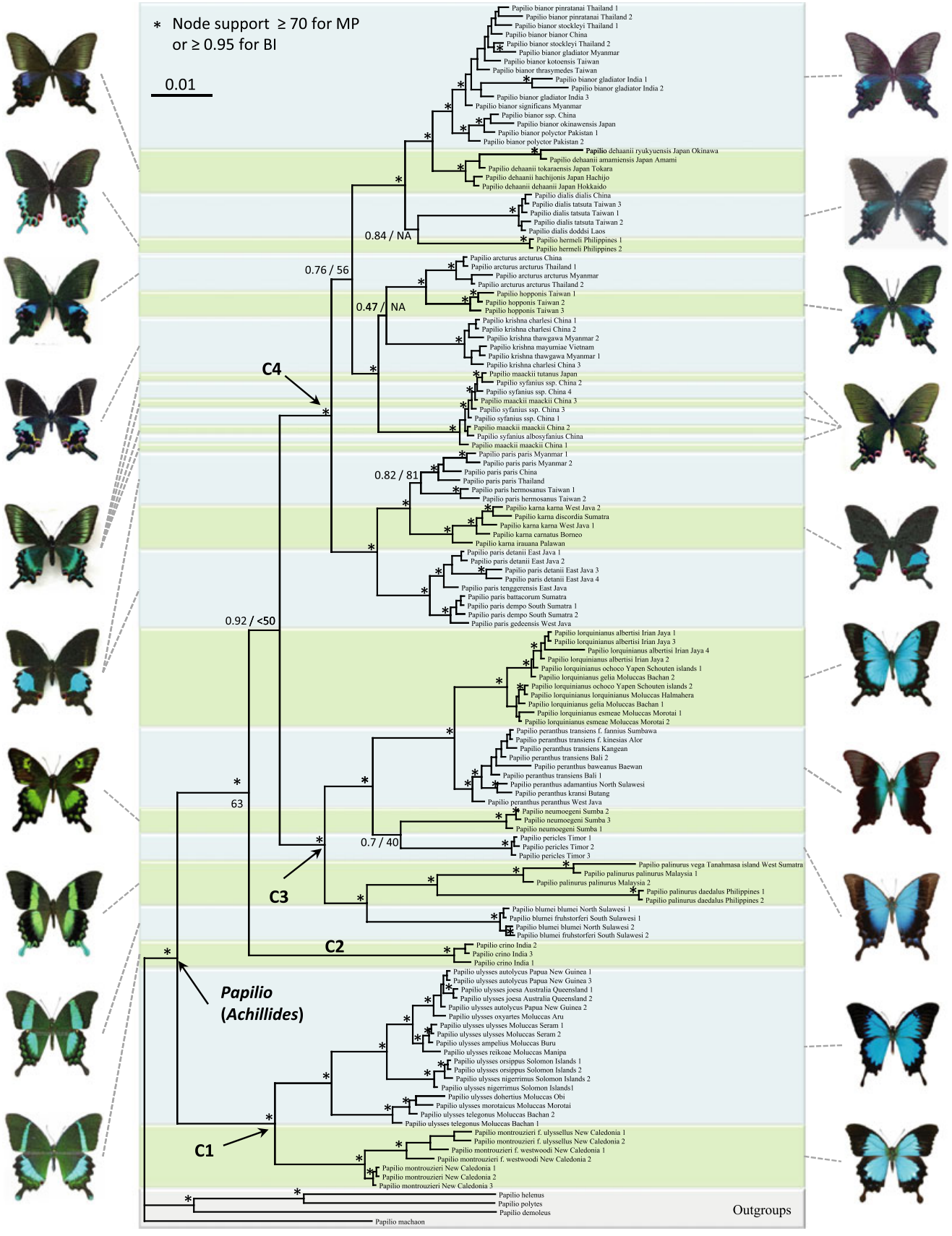
In addition, three species were recovered as paraphyletic in single gene analyses (especially with the nuclear gene EF-1α) and combined partitioned analyses as well, namely *P. paris*, *P. maackii* and *P. syfanius* (Fig. 2). *Papilio paris* is split into two strongly supported clades (BV > 70% and PP = 1.0), separated by a clade that encompasses all sampled individuals of *P. karna*. These two groups are geographically well structured: the first occurs in the Asian mainland whereas the second is distributed in Sundaland. The other case of paraphyly corresponds to the species *P. maackii* and *P. syfanius*, which are mixed in a single well-supported clade in all phylogenetic analyses (BV > 70% and PP = 1). Supplementary analyses were performed for each case of species paraphyly (Table 2): a first analysis was conducted with all sampled individuals of *P. paris* constrained to monophyly, and in a second analysis, members of *P. maackii* and *P. syfanius* were constrained to reciprocal monophyly. For each constrained analysis, BFs were assessed through the estimation of harmonic means. These analyses reveal that the paraphylies are confirmed because the unconstrained topology is significantly better than the constrained topologies enforcing the monophyly of *P. paris* or *P. maackii* + *P. syfanius* (Table 2).

Species delimitation

The existence of distinct phylogenetic lineages was corroborated by the analysis of the branching rate pattern. An LTT plot based on the ultrametric tree demonstrated a sudden increase in branching rate towards the present, corresponding to the switch from interspecific to intraspecific branching events (Fig. 3). To determine the position of the switch, the method of Pons et al. (2006) was applied to the ultrametric tree. The GMYC model was preferred over the null model of uniform branching rates (logL = 670.70, compared with the null model logL = 662.60; $2\Delta L = 16.196$, χ^2 test, d.f. = 3, $P < 0.00103$). The model fitted the switch in the branching pattern occurring at −0.157053 (i.e. T of the ML solution, the time separating the ingroup root from the present was arbitrarily assigned to 1), leading to an estimate of 26 putative species (Fig. 3). Overall, we recovered four putative species in clade C1, eight in clade C3 and 14 in clade C4 (Fig. 3). Note that four species (*P. bianor*, *P. dehaanii*, *P. paris* and *P. ulysses*) appear to correspond to several putative species (two, two, three and three, respectively) following this approach (therefore they are labelled with distinct numbers in Fig. 3) and two species are combined into one putative species [*P. maackii* and *P. syfanius*, hereafter *P. maackii sensu lato* (s.l.)]. The confidence interval for the threshold gave a range of 18–31 putative species with likelihood scores ranging from 668.85 to 668.89 respectively (i.e. estimates falling within 2 log-likelihood units of the ML solution).

By comparison, this approach was also applied to the topology obtained with the COI gene alone (the extended-barcode marker; see Appendix S4). The GMYC model was also preferred over the null model (logL = 313.44, compared with logL = 320.88; $2\Delta L = 14.878$, χ^2 test, d.f. = 3, $P < 0.0019$). By contrast, the

Fig. 2. Phylogenetic relationships of the *Papilio* subgenus *Achillides* obtained with the best-fit partitioning strategy (seven partitions) under Bayesian inference. Posterior probabilities (PP) and maximum parsimony bootstrap values (BV) are shown by nodes (an asterisk indicates PP ≥ 0.95 or BV ≥ 70%). NA indicates that MP analyses did not recover this topology. When no values are shown (especially at the intraspecific level) the BV and PP were ≤ 70% and 0.95, respectively. For deeper nodes, values are indicated. At the tips, the species and subspecies name is given, plus locality. A blue or green horizontal band delimits each species. *Achillides* and four major clades (C1–C4) are labelled.


* Node support ≥ 70 for MP
or ≥ 0.95 for BI
0.01
Papilio bianor pinratanai Thailand 1
Papilio bianor pinratanai Thailand 2
Papilio bianor stockleyi Thailand 1
Papilio bianor bianor China
Papilio bianor stockleyi Thailand 2
Papilio bianor gladiator Myanmar
Papilio bianor kotoensis Taiwan
Papilio bianor thrasymedes Taiwan
Papilio bianor gladiator India 1
Papilio bianor gladiator India 2
Papilio bianor gladiator India 3
Papilio bianor significans Myanmar
Papilio bianor ssp. China
Papilio bianor okinawensis Japan
Papilio bianor polyctor Pakistan 1
Papilio bianor polyctor Pakistan 2
Papilio dehaanii ryukyuensis Japan Okinawa
Papilio dehaanii amamiensis Japan Amami
Papilio dehaanii tokaraensis Japan Tokara
Papilio dehaanii hachijonis Japan Hachijo
Papilio dehaanii dehaanii Japan Hokkaido
Papilio dialis dialis China
Papilio dialis tatsuta Taiwan 3
Papilio dialis tatsuta Taiwan 1
Papilio dialis tatsuta Taiwan 2
Papilio dialis doddsi Laos
Papilio hermeli Philippines 1
Papilio hermeli Philippines 2
0.84 / NA
Papilio arcturus arcturus China
Papilio arcturus arcturus Thailand 1
Papilio arcturus arcturus Myanmar
Papilio arcturus arcturus Thailand 2
Papilio hopponis Taiwan 1
Papilio hopponis Taiwan 2
Papilio hopponis Taiwan 3
0.76 / 56
0.47 / NA
Papilio krishna charlesi China 1
Papilio krishna charlesi China 2
Papilio krishna thawgawa Myanmar 2
Papilio krishna mayumiae Vietnam
Papilio krishna thawgawa Myanmar 1
Papilio krishna charlesi China 3
C4
Papilio maackii tutanus Japan
Papilio syfanius ssp. China 2
Papilio syfanius ssp. China 4
Papilio maackii maackii China 3
Papilio syfanius ssp. China 3
Papilio syfanius ssp. China 1
Papilio maackii maackii China 2
Papilio syfanius albosyfanius China
Papilio maackii maackii China 1
Papilio paris paris Myanmar 1
Papilio paris paris Myanmar 2
Papilio paris paris China
Papilio paris paris Thailand
0.82 / 81
Papilio paris hermosanus Taiwan 1
Papilio paris hermosanus Taiwan 2
Papilio karna karna West Java 2
Papilio karna discordia Sumatra
Papilio karna karna West Java 1
Papilio karna carnatus Borneo
Papilio karna irauana Palawan
Papilio paris detanii East Java 1
Papilio paris detanii East Java 2
Papilio paris detanii East Java 3
Papilio paris detanii East Java 4
Papilio paris tenggerensis East Java
Papilio paris battacorum Sumatra
Papilio paris dempo South Sumatra 1
Papilio paris dempo South Sumatra 2
Papilio paris gedeensis West Java
0.92 / <50
Papilio lorquinianus albertisi Irian Jaya 1
Papilio lorquinianus albertisi Irian Jaya 3
Papilio lorquinianus albertisi Irian Jaya 4
Papilio lorquinianus albertisi Irian Jaya 2
Papilio lorquinianus ochoco Yapen Schouten islands 1
Papilio lorquinianus gelia Moluccas Bachan 2
Papilio lorquinianus ochoco Yapen Schouten islands 2
Papilio lorquinianus lorquinianus Moluccas Halmahera
Papilio lorquinianus gelia Moluccas Bachan 1
Papilio lorquinianus esmeae Moluccas Morotai 1
Papilio lorquinianus esmeae Moluccas Morotai 2
Papilio peranthus transiens f. fannius Sumbawa
Papilio peranthus transiens f. kinesias Alor
Papilio peranthus transiens Kangean
Papilio peranthus transiens Bali 2
Papilio peranthus baweanus Baewan
Papilio peranthus transiens Bali 1
Papilio peranthus adamantius North Sulawesi
Papilio peranthus kransi Butang
Papilio peranthus peranthus West Java
63
Papilio neumoegeni Sumba 2
Papilio neumoegeni Sumba 3
Papilio neumoegeni Sumba 1
0.7 / 40
Papilio pericles Timor 1
Papilio pericles Timor 2
Papilio pericles Timor 3
Papilio palinurus vega Tanahmasa island West Sumatra
Papilio palinurus palinurus Malaysia 1
Papilio palinurus palinurus Malaysia 2
Papilio palinurus daedalus Philippines 1
Papilio palinurus daedalus Philippines 2
C3
Papilio blumei blumei North Sulawesi 1
Papilio blumei fruhstorferi South Sulawesi 1
Papilio blumei blumei North Sulawesi 2
Papilio blumei fruhstorferi South Sulawesi 2
C2
Papilio crino India 2
Papilio crino India 3
Papilio crino India 1
Papilio (Achillides)
Papilio ulysses autolycus Papua New Guinea 1
Papilio ulysses autolycus Papua New Guinea 3
Papilio ulysses joesa Australia Queensland 1
Papilio ulysses joesa Australia Queensland 2
Papilio ulysses autolycus Papua New Guinea 2
Papilio ulysses oxyartes Moluccas Aru
Papilio ulysses ulysses Moluccas Seram 1
Papilio ulysses ulysses Moluccas Seram 2
Papilio ulysses ampelius Moluccas Buru
Papilio ulysses reikoae Moluccas Manipa
Papilio ulysses orsippus Solomon Islands 1
Papilio ulysses orsippus Solomon Islands 2
Papilio ulysses nigerrimus Solomon Islands 2
Papilio ulysses nigerrimus Solomon Islands1
Papilio ulysses dohertius Moluccas Obi
Papilio ulysses morotaicus Moluccas Morotai
Papilio ulysses telegonus Moluccas Bachan 2
Papilio ulysses telegonus Moluccas Bachan 1
C1
Papilio montrouzieri f. ulyssellus New Caledonia 1
Papilio montrouzieri f. ulyssellus New Caledonia 2
Papilio montrouzieri f. westwoodi New Caledonia 1
Papilio montrouzieri f. westwoodi New Caledonia 2
Papilio montrouzieri New Caledonia 1
Papilio montrouzieri New Caledonia 2
Papilio montrouzieri New Caledonia 3
Papilio helenus
Papilio polytes
Papilio demoleus
Papilio machaon
Outgroups

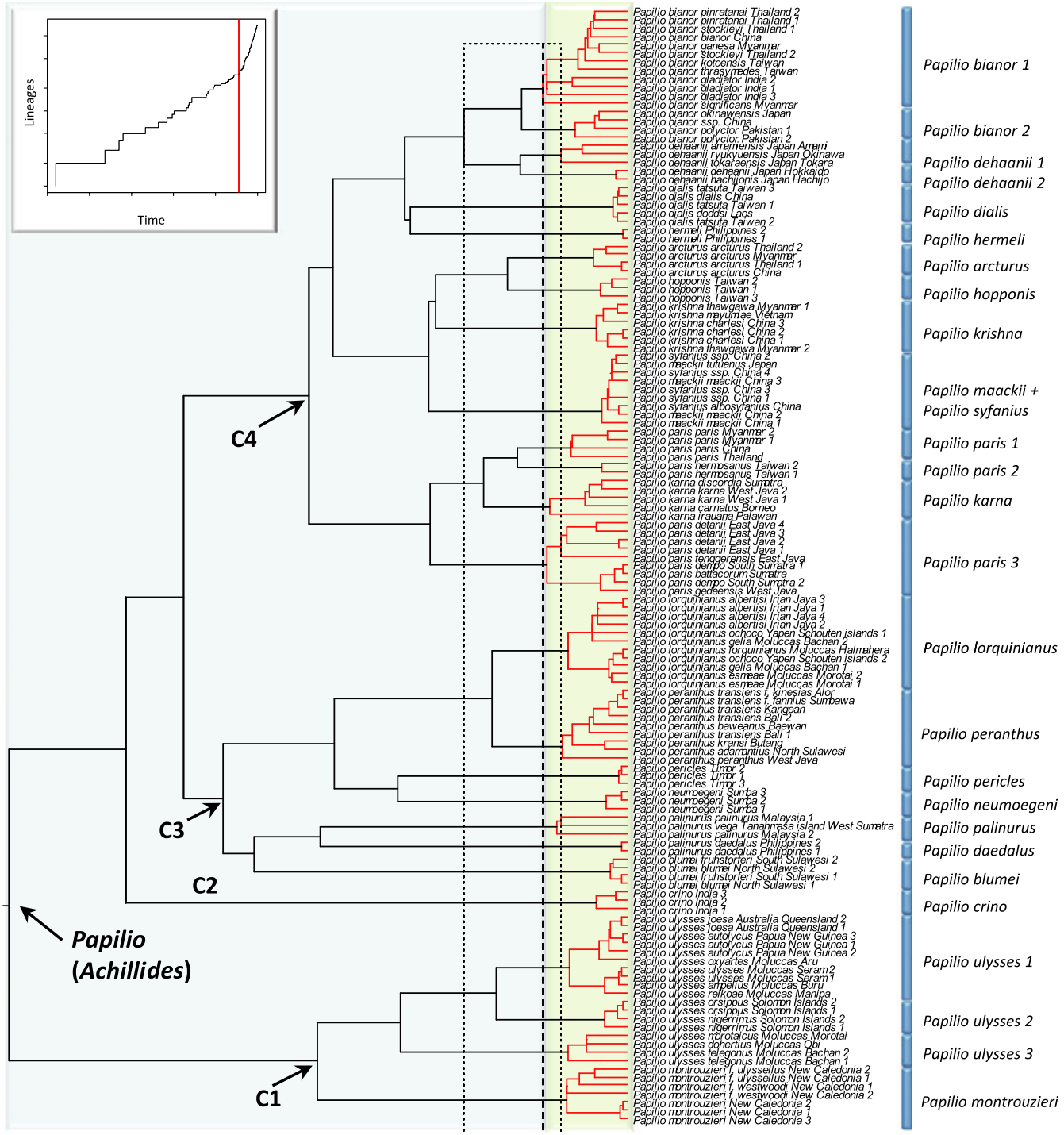


Fig. 3. Ultrametric tree of the *Papilio* subgenus *Achillides* obtained with PATHd8, showing clusters of specimens recognized as putative species. Genetic clusters recognized as a putative species are highlighted in red and separated by longer black branches. Blue and green background shadings represent the threshold (vertical black dashed line) between putative interspecific and intraspecific branching, respectively. The top corner graph shows the lineages-through-time plot based on the ultrametric tree. The sudden increase in branching rate, indicated by a red line, corresponds to the shift from interspecific to intraspecific lineage branching. The vertical bars group all sequences within each significant cluster, labelled by a temporary species name. The four major clades are labelled (C1–C4).

analysis recovered 33 putative species, with a confidence interval ranging from 19 to 44 estimated species (likelihood scores of 315.20 and 315.24, respectively). For details in the difference of species delimitation results between the multi-markers approach and the barcode approach, see Appendix S4. We conservatively chose to keep the number of species obtained with the multiple marker approach for the set of taxa used for the biogeographical analyses.

Estimates of divergence times and historical biogeography

The results of divergence time analysis under a BRC method are represented in Fig. 4. The complete list of median ages and their 95% highest posterior density (HPD) for each node are given in Table 3. Using Tracer, we observed through effective sample size (threshold fixed to 200) that all runs of the two analyses (2.5×10^7 and 5×10^7 generations) provide similar ages for *Achil-*

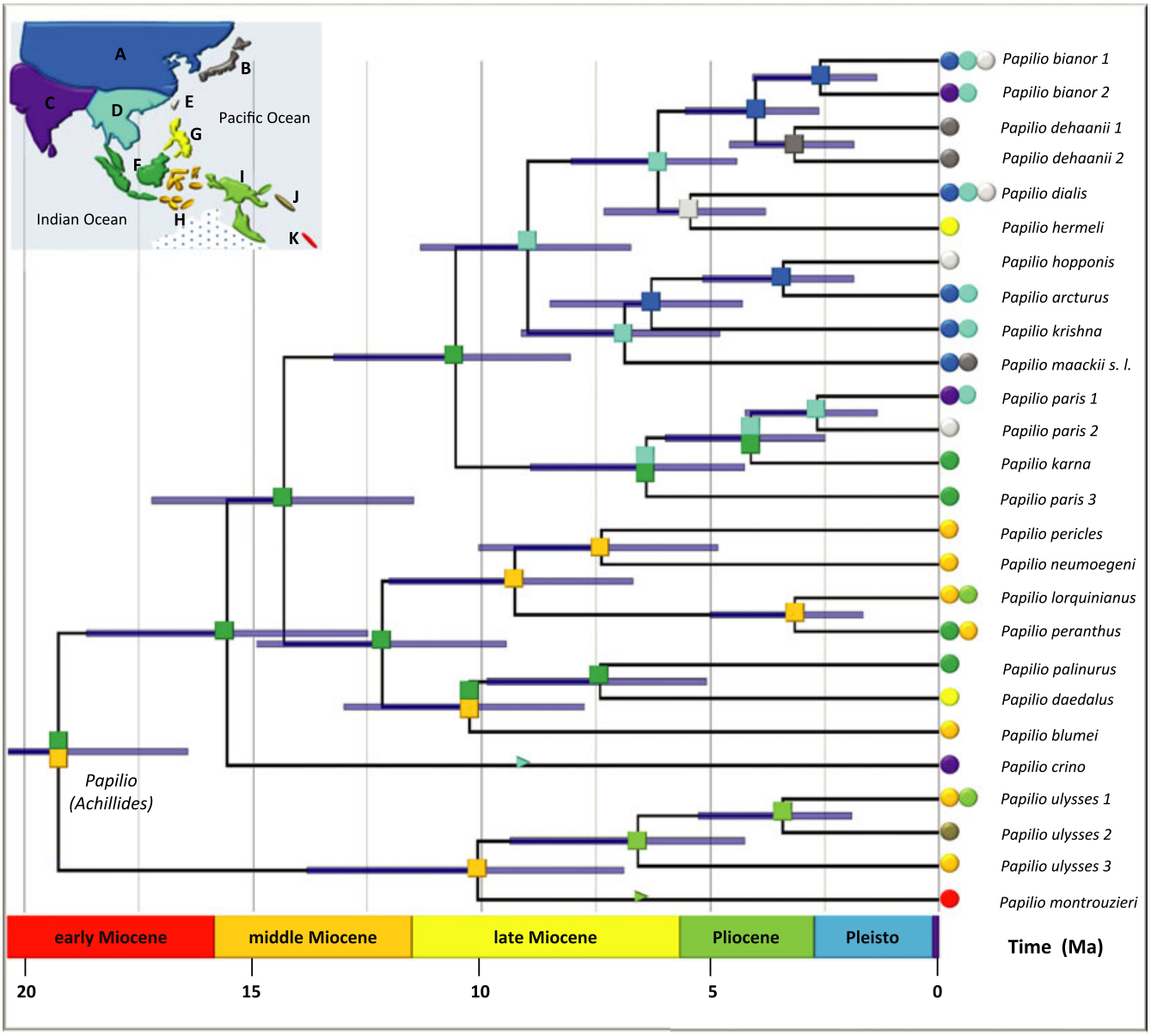


Fig. 4. Historical biogeography of the *Papilio* subgenus *Achillides*. The Bayesian chronogram shows phylogenetic relationships of *Achillides* (outgroups removed), with ages given as Ma and the 95% highest posterior density shown for each node. A 5-Myr timescale is placed at the bottom of the chronogram spanning the epochs since 20 Ma. The top left corner map represents the Indo-Australian Archipelago delimited into 11 areas. For each node, the biogeographical reconstruction is indicated by a coloured square for the ML-DEC (representing the most likely ancestral area). Colours of squares correspond to the coloured area on the map. Coloured triangles on branches indicate dispersal (optimizations of the ML-DEC analyses) from the ancestral region to the present-day region. The present-day distribution of each species is informed at the tips by coloured circles referring to the coloured areas on each map.

Table 3
Combined results of phylogeny, estimates of divergence times and biogeography of *Papilio* subgenus *Achillides*

		Phylogeny		BRC chronogram		ML-DEC reconstructions		
Taxonomy	Nodes	BV	PP	Age	95% CI	Range(s)	ln L	RP
Root*	1	NA	NA	22.86	19.17–26.8	NA	NA	NA
Papilio demoleus–Papilio (*Achillides*)*	2	100	NA	21.27	18.14–24.45	NA	NA	NA
*Papilio demoleus–Papilio helenus**	3	100	0.97	18.74	15.9–21.7	NA	NA	NA
*Papilio polytes–Papilio helenus**	4	100	1	13.12	11.6–15.87	NA	NA	NA
Papilio (*Achillides*)*	5	90	1	19.25	16.42–21.85	FH	89.54	Best
Clade 1	6	100	1	10.09	6.89–13.81	H	90.16	0.537
Papilio ulysses 1–*Papilio ulysses* 3	7	99	1	6.59	4.25–9.38	HI	90.74	0.302
Papilio ulysses 1–*Papilio ulysses* 2	8	99	1	3.42	1.91–5.26	I	90.79	0.285
Clade 2 (*P. crino*)	9	NA	1	15.57	12.49–18.64	F	90.16	0.537
Clade 3–Clade 4	10	61	1	14.32	11.48–17.21	F	90.49	0.387
Clade 3	11	75	0.8	12.17	9.47–14.94	F	90.17	0.532
Papilio blumei–Papilio daedalus	12	87	1	10.27	7.75–13	FH	90.92	0.252
Papilio daedalus–Papili palinurus	13	99	1	7.41	5.08–9.89	F	90.74	0.299
Papilio pericles–Papilio lorquinianus	14	89	1	9.28	6.69–12.04	H	90.92	0.252
Papilio pericles–Papilio neumoegeni	15	6	1	7.38	4.84–10.07	H	90.11	0.563
Papilio peranthus–Papilio lorquinianus	16	100	1	3.16	1.66–5.01	H	90.11	0.563
Clade 4	17	90	1	10.57	8.05–13.24	F	90.17	0.532
Papilio paris 1–*Papilio paris* 3	18	95	1	6.4	4.25–8.94	DF	90.26	0.487
Papilio karna–Papilio paris 2	19	85	1	4.12	2.5–5.99	DF	90.37	0.433
Papilio paris 1–*Papilio paris* 2	20	43	0.96	2.67	1.36–4.24	D	90.34	0.448
Papilio maackii s.l. –Papilio bianor 1	21	60	0.9	8.99	6.74–11.34	D	90.26	0.487
Papilio maackii s.l. –Papilio arcturus	22	NA	1	6.87	4.8–9.14	D	90.87	0.263
Papilio krishna–Papilio arcturus	23	72	0.96	6.29	4.31–8.5	A	91.54	0.135
Papilio hopponis–Papilio arcturus	24	98	1	3.41	1.87–5.17	D	91.13	0.204
Papilio hermeli–Papilio bianor 1	25	99	1	6.15	4.42–8.04	D	90.87	0.263
Papilio dialis–Papilio hermeli	26	8	0.94	5.44	3.79–7.33	DE	91.38	0.159
Papilio bianor 1–*Papilio dehaanii* 1	27	77	1	4.02	2.62–5.55	A	91.38	0.159
Papilio bianor 1–*Papilio bianor* 2	28	52	0.54	2.6	1.37–4.08	AB	90.88	0.262
Papilio dehaanii 1–*Papilio dehaanii* 2	29	4	0.99	3.16	1.87–4.58	B	90.88	0.262

The table refers to Fig. 4, organized by node number. For each node, we indicated (i) the phylogenetic support with the bootstrap values (BV) for maximum parsimony and posterior probability (PP) obtained in Bayesian analysis, (ii) the estimated median age with the 95% highest posterior density for the chronogram obtained using hard bounds under the Bayesian relaxed-clock method, and (iii) the ancestral range reconstruction obtained by ML-DEC analyses with the likelihood score (ln L) and the relative probability (RP). For geographical analysis, only the first split at a node is given and can be non-significant. Asterisks indicated where a constraint was used to calibrate the molecular clock. NA, not available because of the rooting process or because outgroups were removed for geographical range evolution analyses.

lides. Overall, an early Miocene origin of the subgenus *Achillides* was recovered (*ca*. 19.25 Ma, HPD 16.4–21.8). Clade C1 appeared at 10 Ma (HPD 6.9–13.8). Clade 2 (i.e. the divergence of *P. crino*) appeared at 15.6 (HPD 12.5–18.6). The split between clade C3 and C4 originated around 14.3 Ma (HPD 11.5–17.2). Finally, clade C3 and clade C4 appeared at 12.2 Ma (HPD 9.5–14.9 Ma) and 10.6 Ma (HPD 8.1–13.2 Ma), respectively. For more details on each node (date and 95% HPD), see Table 3.

The results of ML-DEC analyses of geographical range evolution are presented in Fig. 4. Ancestral areas and likelihood scores (L) with their relative probabilities (RP) are given for each node in Table 3. For the root, local optimizations under the ML-DEC method (with either two or three *maxareas*) recovered an optimal range area for the Sundaic region ($L = -91$ for two *maxareas*), which was nonetheless not significantly different from those of the Wallacea region ($L = -91.05$ for two *maxareas*). Other ranges were not statistically supported (more than 2 log-likelihood units of difference) by alternative local optimizations (e.g. for the Indo-Chinese region, $L = -94.71$ for two *maxareas*). Together, this evidence indicates that the ancestral area for the common ancestor of *Achillides* is more likely to be in the Sundaic and/or Wallacea regions. To improve the precision of this inference, supplementary ML-DEC analyses were conducted with the root enforced to be a combination of two or three areas. The combination of the Sundaland and Wallacea regions was significantly supported over the remaining combinations ($L = -89.54$ for two *maxareas*) even when three areas were constrained at the root (best score for Sundaland + Wallacea + Sahul, $L = -92.26$).

The results of both analyses of ancestral area evolution unambiguously support a dynamic dispersal pattern (rate of dispersal = 0.2481 for the optimization under the best ML-DEC model) and few vicariance events were reconstructed. Note that the DEC analyses, with-

out constraints on the root, recovered the same ancestral areas for the remaining nodes.

Diversification rates

Temporal changes in diversification rates are presented using an LTT plot in Appendix S5. Under a birth–death model and using three extinction rate values (ε = 0/0.5/0.9), net diversification rates were estimated at 0.135/0.122/0.064 per Ma for our dataset and at 0.137/0.124/0.065 per Ma taking into account the three missing species. The one-tailed γ-test rejected the null hypothesis of rate constancy (critical-γ = −2.7212, P = 0.0032). The likelihood-based method of diversification analysis provided strong support for a rate-variable model over the constant-rate model (AIC_{RC} = 37.84, AIC_{RV} = 27.99). Following Rabosky (2006a), we have confidence that the rate-variable model best approximates the data because the AIC difference between rate-constant and rate-variable is >4 (in our case ΔAIC_{RC} = 9.85). Both γ-statistic and likelihood analyses confirmed that diversification rates have varied through time. Among the rate-variable models (DDL, DDX, Yule-2-rate, Yule-3-rate and Yule-4-rate), we found that the best-fit model is the Yule-3-rate model. This result thus suggests that two rate shifts have occurred during the evolutionary history of *Achillides*.

When exploring the tempo of diversification and especially the impact of climate events that occurred during the diversification of the group, non-significant effects were found between the origin of *Achillides* and 13 Ma (P = 0.915, LR = 0.011) corresponding to the early divergence among the four clades (Appendix S5). However, significant effects on diversification rates were found at 10.3 Ma (P = 0.024, LR = 7.423; between 15 and 9 Ma), at 7.4 Ma (P = 0.014, LR = 8.572; between 9 and 6 Ma) and at 3.5 Ma (P = 0.017, LR = 11.899; between 5 and 2.5 Ma). Diversification rates were inferred to be higher after these events (under the ML-based method fitting pure birth model to branching times).

Discussion

Phylogenetic relationships and species delimitation within Achillides

The results of the phylogenetic analyses are well supported and consistent regardless of the inference method used (BI or MP), as has commonly been found in other studies (Rindal and Brower, 2011). This strong phylogenetic signal is probably attributable to dense taxonomic sampling and to the size and phylogenetic usefulness of the selected molecular markers (Sperling, 2003). Our results clearly supported the monophyly of *Achillides*, in agreement with previous studies (Zakharov et al., 2004; Yagi et al., 2006). Compared with the work of Yagi et al. (2006), which constitutes the first molecular phylogenetic study of *Achillides*, we found a similar topology even though their study was only based on the ND5 molecular marker and comprised either 11 or 14 species, depending on competing classifications (Shimogori, 1997; Bauer and Frankenbach, 1998; Häuser et al., 2005). The phylogenetic hypotheses recovered in the two studies are in general agreement, with the species belonging to *lorquinianus* and *palinurus* groups appearing as basal to the remaining species sampled by Yagi et al. (2006). Within our clade 4, two major subclades are also comparable in both studies, including *P. bianor*, *P. dehaanii*, *P. dialis*, *P. hermeli* and *P. polyctor* (Yagi et al.'s clade 1), which is sister to a clade comprising *P. arcturus*, *P. hopponis*, *P. karna*, *P. krishna*, *P. maackii*, *P. paris* and *P. syfanius* (Yagi et al.'s clade 2) (Fig. 2).

Apart from these similarities, our *Achillides* species tree presents significant new insights. In particular, clade 1, containing two lineages (*P. montrouzieri* and *P. ulysses*), was recovered in a sister position to all the remaining species. Interestingly, this placement has major implications with regard to the biogeographical pattern (see below). All species are clearly assigned to four well-resolved clades that exhibit distinctive geographical distributions, with the exception of *P. crino*. The latter species is found in an enigmatic and moderately supported position as sister to the C3–C4 clade (Fig. 2), in contrast to its traditional placement within the *palinurus* group (Jordan, 1909). We anticipate that the addition of *P. buddha*, also traditionally placed in the *palinurus* group and occurring in India as well, may bring more resolution on the issue of the phylogenetic position of *P. crino*.

Biogeographic history of Achillides*: roles of Wallacean geological history and palaeoclimate*

Molecular dating analyses suggest the subgenus *Achillides* diverged from its sister group (subgenera *Menelaides* and *Princeps*) around 21 Ma (HPD 18.1–24.5) and diversified around 19.2 Ma (HPD 16.4–21.8) in the early Miocene. This age estimate is also consistent with *Achillides* belonging to a relatively derived group within *Papilio*, a genus that mostly diversified during the middle Cenozoic (Zakharov et al., 2004; Condamine et al., 2012). In other groups of Lepidoptera, such as nymphalid and pierid butterflies, the origin and diversification of major lineages also mostly happened during the mid Cenozoic (e.g. Braby et al., 2006; Peña and Wahlberg, 2008; Wahlberg et al., 2009).

Broadly, the distribution of *Achillides* taxa can be grouped by clade, with clade 1 being endemic to Wallacea and the Australasian region, clade 2 (*P. crino*)

endemic to India, clade 3 ranging across Sundaland and Wallacea, and clade 4 being found in Sundaland and the Indo-Chinese region (Fig. 4). Speciation events are notably older for Sundaic and Wallacean taxa in comparison with species from remaining regions, and long branches imply extended isolation of Sundaic and Wallacean taxa or extinction events. Interestingly, such a pattern has been documented for other groups of butterflies co-occurring in the region (Müller and Beheregaray, 2010; Müller et al., 2010). Yet the temporal importance of Wallacea for the evolution of butterflies has remained enigmatic. Vane-Wright and de Jong (2003) showed that some endemic Sulawesi butterflies represented ancient lineages, a finding supported in other groups (de Boer, 1995).

Geological events as the motor of early geographical diversification

From a biogeographical viewpoint, our molecular dating indicates that the diversification of *Achillides* took place during major tectonic events (Hall, 2002; Metcalfe, 2005). When *Achillides* arose at 20 Ma, Southeast Asia underwent a period of intense tectonic activity with the formation of numerous biogeographical barriers (islands, mountain ranges and deep-water basins; Sanmartín and Ronquist, 2004). Our biogeographical analyses inferred that Sundaland + Wallacea constituted the most likely ancestral area of origin of the subgenus (Figs 4 and 5A), in agreement with a biogeographical analysis of Papilionidae that indicated that the subgenera *Achillides*, *Menelaides* and *Princeps* appeared in Southeast Asia (Condamine et al., 2012). According to several authors (e.g. Hall, 1998, 2002; Metcalfe, 2005), Sundaland appeared as a large peninsula, and geological evidence indicated that Wallacea was partially present 20 Ma (Fig. 5a; Hall, 2002). Shallow seas and palaeo-islands could have connected the two biogeographical entities, a scenario favouring allopatric diversification (Müller et al., 2010).

From this ancestral area, the common ancestor of *Achillides* evolved into three lineages (clade 1, clade 2 and clade 3 + clade 4). The diversification of the predominantly Wallacean clade (clade 1) and of the remaining *Achillides* in Sundaland (Figs 4 and 5b) is associated with a primarily vicariant event probably attributable to geological processes and not to climatic events (Appendix S5). A possible explanation of this vicariant event is the geological rearrangement of Wallacea, and especially for Sulawesi Island, which has changed substantially over the last 20 Myr (Hall, 2002; Metcalfe, 2005). Sulawesi has a unique geological history, and the islands that comprise it probably represent terranes that were sliced from the northern margin of New Guinea at different periods until its formation was completed 5 Ma (Wilson and Moss, 1999; Lohman et al., 2011). These events probably favoured the role of geology in the allopatric diversification of *Achillides*: the first lineage became isolated from the second with the opening of the Makassar Strait (comparable to Wallace's line) between Sundaland and Wallacea (Fig. 5b; Hall, 1998, 2002). The Makassar Strait and much of East Borneo and West Sulawesi was a wide, and locally deep, marine region forming a substantial barrier between Sundaland and Sulawesi (Fig. 5b; Hall, 2009). Such a barrier has also been evidenced in the distribution of the *Cethosia* (Lepidoptera, Nymphalidae) clades (Müller and Beheregaray, 2010). The initial genesis of Wallacea coincides temporally and closely with the earliest splitting in the two aforementioned *Achillides* clades with basal nodes at *ca*. 10 and 15 Ma, respectively. Within clade 3, the common ancestor extended its geographical range from Sundaland to Wallacea and two lineages separated from the common ancestor around 12.2 Ma (Fig. 5c). These two lineages arose with the anticlockwise rotation of Borneo, which provoked intense volcanism in the western part of the Wallacean region, probably facilitating dispersal to Wallacea through the formation of oceanic islands (Hall, 1998, 2002; Lohman et al., 2011).

Within the last 10 Myr, clade 1, which was first established in Wallacea (Fig. 4), extended its distribution southward to the Sahul and diversified (Fig. 5d). As the Australian plate drifted northward to collide with the Eurasian plate, new land bridges were created between Wallacea and Papua New Guinea through Halmahera and the Bird's Head (Wilson and Moss, 1999; Hall, 2002; Metcalfe, 2005). This pattern favours the role of dispersal events to explain the present distribution of clade 1. Trans-oceanic dispersal of the common ancestor can probably be accounted for by their flying abilities, as they have been able to colonize almost all tropical islands or archipelagos in the Australasian region (Fig. 1; Wallace, 1865). Indeed the four species of the *ulysses* group have colonized all Papua New Guinean islands, the Solomon Islands, and even New Caledonia. The colonization of New Caledonia may represent long-distance dispersal of the common ancestor of clade 1 from surrounding regions. The dispersal to New Caledonia would have been realized either from the Solomon Islands via the Vanuatu archipelago, implying extinction of an ancestral population in these islands, or directly from Australia/Papua New Guinea across intervening islands. Biogeographical analyses recovered an ancestral dispersal from Sahul ($L = -90.74$) that was more likely than from the Solomon Islands ($L = -93.94$).

Other notable events include the separation of the lineage that leads to the Indo-Chinese *P. dialis* and the Philippine *P. hermeli* in the late Miocene around 5.4 Ma, which cannot be attributed to climatic or dispersal events. Phylogenetic and dating results imply

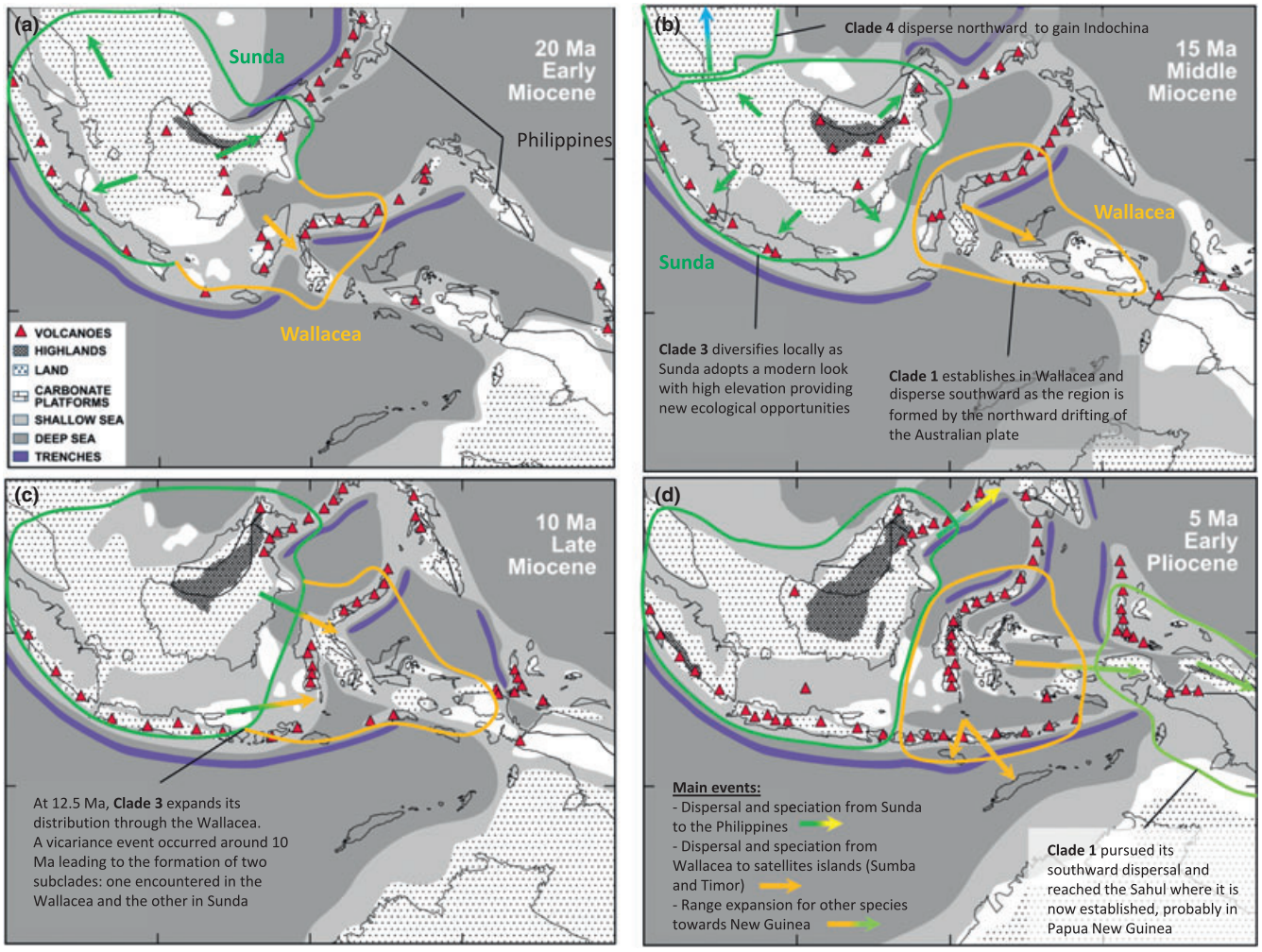


Fig. 5. Biogeographical scenarios for clades 1 and 3 of the *Papilio* subgenus *Achillides* illustrated by maps. (a) Origin and putative ancestral distribution of *Achillides*. *Achillides* appeared in the Sundaic–Wallacean region around 20 Ma. (b) The initial geographical divergence among the three main clades of *Achillides* near 15 Ma. (c) Range expansion of clade 3 followed by a vicariance event among the Sundaic lineage and Wallacean lineage around 10 Ma. (d) Local diversifications of clades 1 and 2 by dispersal events to the Philippines, Sumba and Timor and to Papua New Guinea near 5 Ma. Successive dispersal and range expansions are labelled with coloured arrows. Palaeogeographic maps were redrawn from Hall (1998, 2002).

past geological connections between the two regions. A connection between Taiwan and the Philippines (Luzon) through an arc–trench system has been proposed (Honza and Fujioka, 2004). It has been generally accepted that the formation of Taiwan resulted from the collision of the Luzon Arc with the Eurasian continental margin or with the former Ryukyu subduction zone (Honza and Fujioka, 2004). The Luzon Arc formed during the early to middle Miocene and probably connected Taiwan with Luzon from the late Miocene to late Pliocene (*ca.* 10–2 Ma, Honza and Fujioka, 2004), which facilitated the spread of organisms. As Taiwan was connected to southern China at various stages during the Pliocene and Pleistocene, based on bathymetric studies (Honza and Fujioka, 2004), it is plausible that the ancestor of *P. dialis* and *P. hermeli* was spread from Taiwan to the Philippines by this arc connecting Luzon and Taiwan, and that a vicariance event separated the two current lineages. This scenario is corroborated by the presence of *P. dialis* in Taiwan, as well as in Indo-China (Fig. 1). A similar palaeo-connection is likely to have resulted in the distributions of the closely related *Papilio* swallowtail species *P. xuthus* (Taiwan, Japan and China) and *P. benguetanus* (northern Luzon), as well as *Cethosia* nymphalids (Müller and Beheregaray, 2010).

Climate change as a major driving force in late diversification

Although it is difficult to link significant effects of past climate changes to the diversification rates of organisms (Currie et al., 2004; Winkler et al., 2009), in this study we uncovered evidence for three shifts in diversification

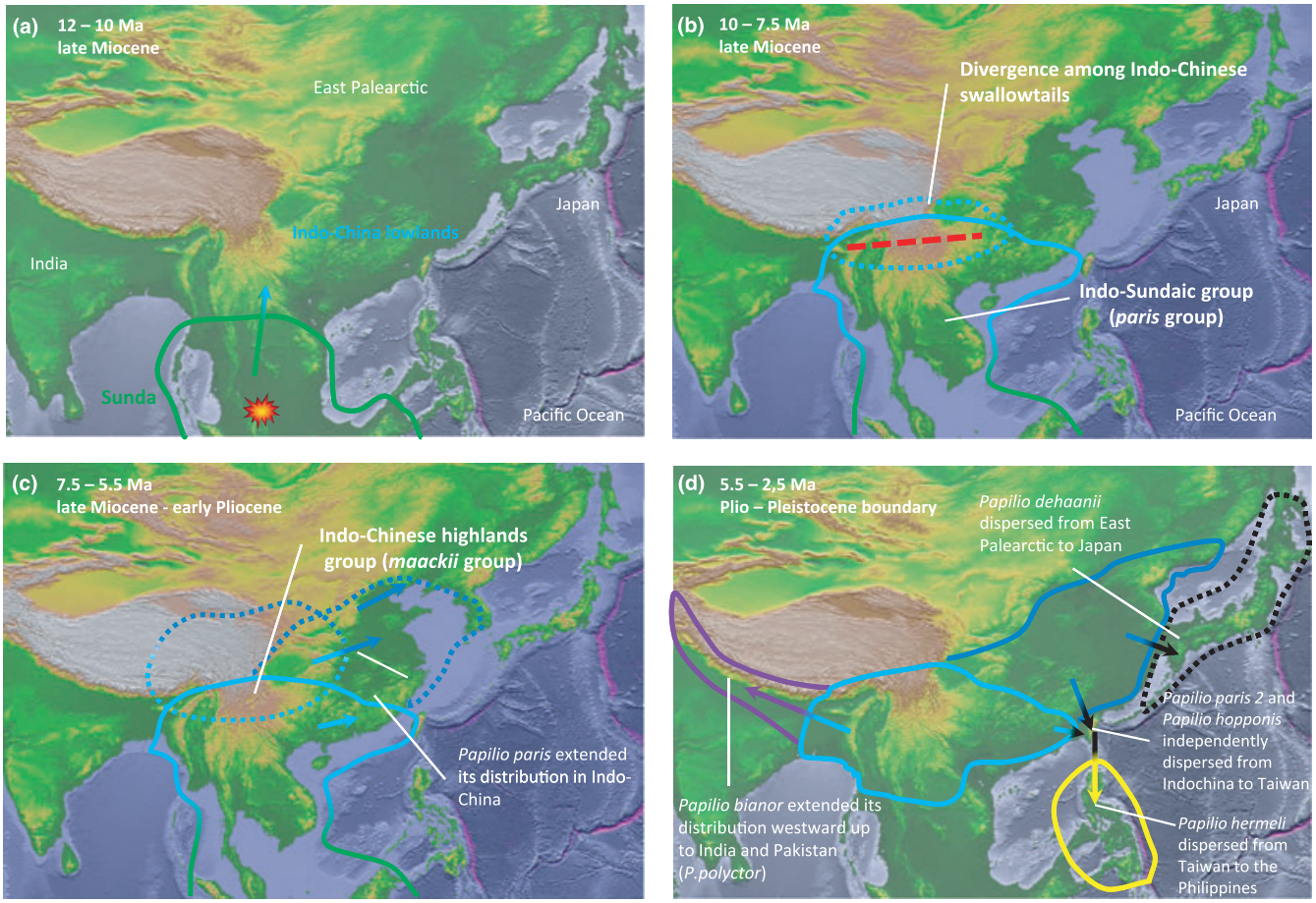


Fig. 6. Biogeographic scenarios for clade 4 of the *Papilio* subgenus *Achillides* illustrated by maps of the Asian mainland. As the palaeogeographical situation at 10 Ma was very similar to that of today, the same map is used, except for the sea-level decreases for maps A and C. (a) Origin of clade 4 and initial diversification in Sunda when sea level was low between 12 and 10 Ma. (b) Northward dispersal and early divergence among Indo-Chinese and Sundaic species. The diversification of clade 4 continued and three lineages arose: a Sundaic group, an Indo-Chinese group and a Chinese group between 10 and 7.5 Ma. (c) The Chinese group extended its distribution northward, whereas the Indo-Chinese lineage expanded westward to India (parapatric speciation). (d) Late range expansions and dispersals to islands when sea level dropped. During the Plio-Pleistocene glaciations, several island dispersals are identified to Japan (*P. dehaanii* 1 and 2), to Taiwan (*P. hopponis* and *P. paris* 2 from clade 3) and to the Philippines (*P. hermeli*). Lines indicate putative ancestral areas of distribution of clades (dashed lines for subclades), while arrows identify dispersals.

rates that may be correlated with specific climatic events. These results are of interest for: (i) understanding diversification processes, especially in the Indo-Australian Archipelago where geological vicariance is often postulated; and (ii) understanding the effects of global climate change as studies showing such climate effects are scarce (Winkler et al., 2009; Condamine et al., 2012). According to the LTT plot and birth–death likelihood tests, we found that diversification rates of *Achillides* shifted at 10.3 and at 7.4 Ma in the late Miocene, and also at 3.5 Ma at the Pliocene–Pleistocene boundary (Figs 6 and 7).

We recovered higher diversification rates just after these three periods and interpreted them as a response to global climate dynamics (Zachos et al., 2001), rather than to geological events, given that the area was tectonically stable during these periods (Wilson and Moss, 1999; Hall, 2002; Metcalfe, 2005). In the Cenozoic, Earth's climate underwent significant and complex evolution, including gradual trends of warming and cooling (Zachos et al., 2001). The most prominent response to climate change in the past has probably been sea-level fluctuations (Voris, 2000; Miller et al., 2005), which would have seen the islands of the Indo-Australian Archipelago repeatedly connected and isolated. This probably led to dispersal along the Sundaic shelf during periods of low sea level, and promoted subsequent diversification through isolation during periods of higher sea level (Figs 6 and 7; Woodruff and Turner, 2009).

Around 10.3 Ma, the shift in diversification rate is attributed to a cooling event (Fig. 6). It is well known

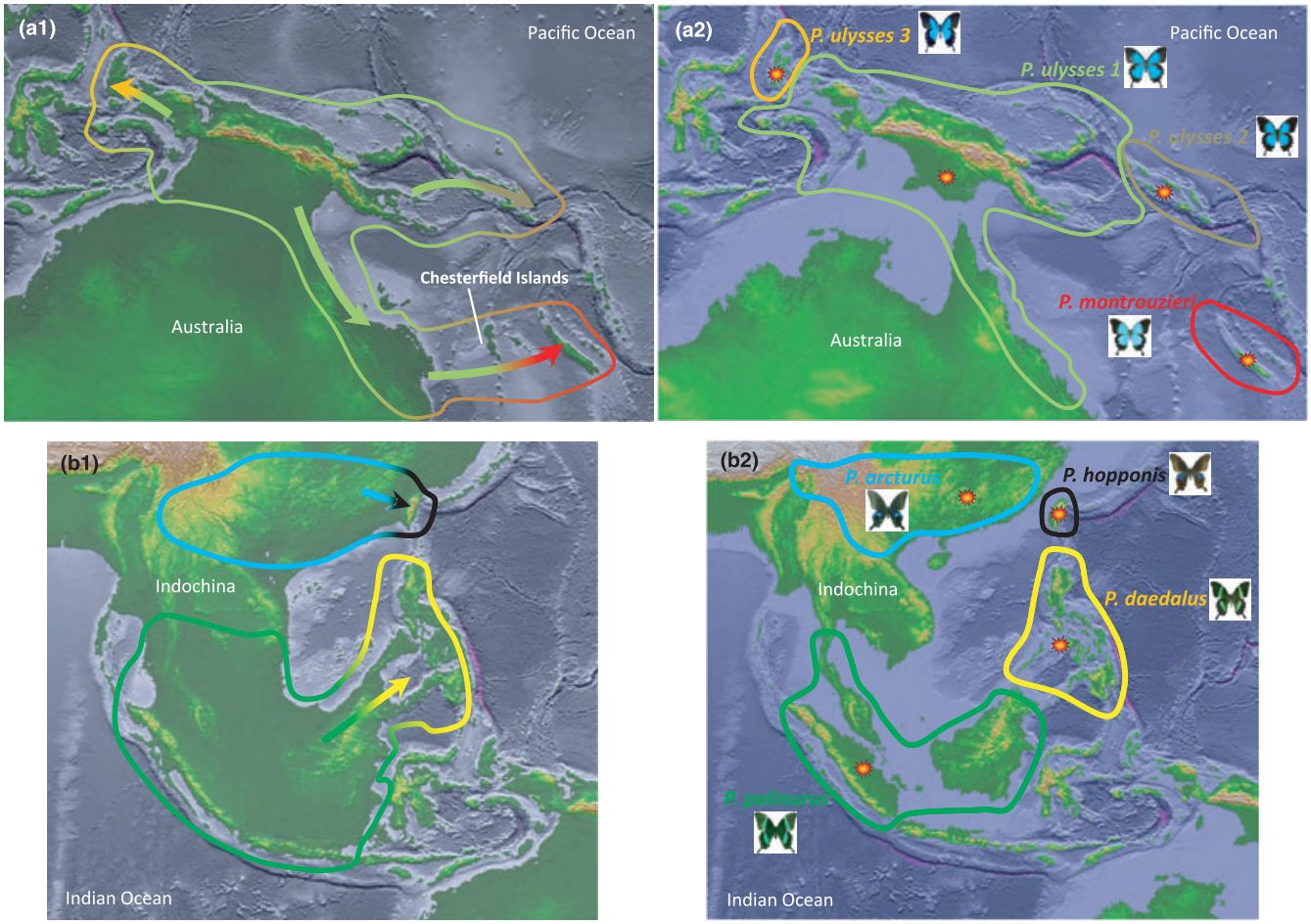


Fig. 7. Schematic maps showing the effects on diversification of the *Papilio* subgenus *Achillides* during the Plio-Pleistocene glaciations. (a) The maps represent the diversification of clade 1 (*ulysses* group). Ten million years ago, a period when sea level decreased due to a cooling event (a1), the common ancestor was probably distributed throughout the Sahulian region. Range expansion has been facilitated because land bridges existed and distribution areas were thus large. When sea level increased (a2), the populations became isolated and evolved by allopatry. (b) The maps represent the diversification of two pairs of sister species (*P. arcturus*/*P. hopponis* and *P. daedalus*/*P. palinurus*) in Indochina and Sunda. Similarly to the *ulysses* group, when sea level decreased during the cooling period (b1), dispersals were facilitated and distribution areas were large. When sea level increased (b2), populations became isolated and evolved by vicariance.

that cooling periods favoured a decrease of global sea level by accumulating water on ice-sheets at the poles, and the inverse is true during warmer times (Zachos et al., 2001; Miller et al., 2005). This fall in sea level probably facilitated range expansion among areas normally isolated, especially in the Indo-Australian Archipelago, and then favoured vicariance events (Hall, 2002; Woodruff and Turner, 2009). This climate change corresponds to the appearance of clade 1 and 3 (Fig. 6). Clade 1 diversified early from the common ancestor by geological vicariance (Fig. 4), but climate change also appeared to influence the diversification within clade 1, which involved a split between *P. montrouzieri* and the three putative *P. ulysses* species. Above, we assumed that *P. montrouzieri* appeared by long-distance dispersal from Sahul, and these results suggest that such an exceptional event (more than 1000 km of oceanic barrier) was probably facilitated by sea level decrease during the cooling event (Fig. 7a). North-eastern Australia and New Caledonia are surrounded by shallow seas and by lagoons with coral reefs. When sea level dropped, these extended areas of lagoons probably became land and previously emerged archipelagos, such as the Chesterfield Islands, have been recognized as occurring between Australia and New Caledonia on the Lord Howe ridge (Cluzel et al., 2001; Pelletier, 2006). Thus, the distance between the two biogeographical areas was reduced and dispersal could have been easier for flying insects along stepping-stones (Fig. 7a). We thus suggest that both geological and climatic events have contributed to the diversification of clade 1, which demonstrates that combined analyses are suitable for

understanding fine-grained biogeographical processes. Conversely, for clade 3, climate seems to be an important factor for its diversification. The split between clade 3 and clade 4 is probably not due to geological vicariance (Fig. 4). Climatic influence has been postulated for mammal diversifications in the Malay Peninsula (Woodruff and Turner, 2009), and it has been proposed that the peninsula was partially isolated several times by sea level rise (Lohman et al., 2011). This assumption is of notable importance because the divergence between clade 3 and clade 4 could have been the result of a climatic vicariance (Fig. 5b). Clade 3 occupied the southern part of Sundaland (Greater Sundas) whereas clade 4 inhabited the northern part (Thailand) (Figs 5b and 7a). Species of clade 4 have probably been able to adapt to more subtropical or temperate environments and further colonized the most northern areas towards China and Japan (Fig. 6). Finally in Wallacea, the cooling event provoked a decrease in sea level that has permitted dispersal between western Sulawesi and Sundaland. The ancestor of the Sundaic lineage, within clade 3, probably extended its range to Sulawesi and evolved into *P. blumei*. However, the Makassar Strait is a deep-water trench, so Borneo and Sulawesi cannot have been linked when sea level dropped (Wilson and Moss, 1999), corroborating that this biogeographical event is the result of an inter-island dispersal combined with climatic vicariance event.

Near 7.4 Ma, a sudden cooling event is recorded (Zachos et al., 2001) and corresponds to a new shift in diversification rates within *Achillides* (Appendix S5). The sister relationships of *P. neumoegeni* and *P. pericles*, in the Wallacean lineage of clade 3, may be attributed to this climatic event. We hypothesize that the Wallacean common ancestor dispersed from Sulawesi to the Lesser Sundas when sea level dropped. This ancestor then reached the southernmost islands, which have always been isolated, and when sea level increased, the populations evolved into two endemic species on the Lesser Sundas: *P. neumoegeni* on Sumba and *P. pericles* in Timor. Other species also resulted from the sea level decrease: *P. palinurus* extended its distribution from the Greater Sundas to the Philippines through temporal connections with Palawan, and evolved into *P. daedalus* (Fig. 7b). This scenario is corroborated by our divergence time estimates, which are congruent with the formation of the Philippines (Honza and Fujioka, 2004).

For *Achillides*, the Pliocene–Pleistocene boundary marked a period of peak diversification with five synchronous speciation events around 3.5 Ma affecting all biogeographical regions (Fig. 4; Appendix S5). It suggests that the impact of this late climatic phenomenon was widespread in the Indo-Australian Archipelago, in agreement with Lohman et al. (2011). The early Pliocene (5.5 Ma) is marked by a subtle warming event until 3.2 Ma, when benthic $\delta^{18}O$ increased again, reflecting the onset of northern hemisphere glaciations (Appendix S5; Zachos et al., 2001). Continuous sea level fluctuations have permitted the common ancestor of *P. ulysses 1–2* to disperse to the Solomon Islands and give rise to *P. ulysses 2* (Fig. 7a), as well as the dispersal of *P. arcturus* in Taiwan to evolve into *P. hopponis* (Fig. 7b). Moreover, the Pliocene–Pleistocene sea level rises and drops also appear to have shaped the subgenus at the subspecific level. The various subspecies of *P. lorquinianus*, *P. paris*, *P. peranthus* and *P. ulysses* (Fig. 1) in Sundaland and Wallacea probably diverged from an ancestor that spanned the entire island group when they were connected during a period of low sea level (Fig. 7). Woodruff and Turner (2009) showed that many rapid sea level rises in the last 5 Myr have probably resulted in faunal compression and expansion, corroborated here by the extant distribution pattern of *Achillides*, being mainly shaped by climatic vicariance.

Conclusions and insights for the biogeography of the Indo-Australian Archipelago

The historical biogeography of the genus *Papilio* represents a significant challenge. *Papilio* is a highly complex and diverse taxonomic unit with at least 200 species and many more subspecies (Häuser et al., 2005). Biogeographical analysis of the subgenus *Achillides* has provided insight into the history of the region and the processes that have shaped the evolution of these swallowtails. We showed that both geological and climatic events have probably influenced the geographical diversification of a species-rich insect group. This study has also produced valuable results as it is the first time that several significant effects of climate changes have been proposed for butterflies of the Indo-Australian Archipelago. The sea level fluctuations have probably been a major factor in their late geographical diversification. However, we remain cautious about our results and interpretations, especially in a region such as the Indo-Australian Archipelago. New Caledonia serves to remind us of fundamental geological and biogeographical problems in the region (Hall, 1998). Thus, investigation of the biogeographical history of organisms in this region requires good knowledge of palaeogeographical and palaeoclimatic history.

Beyond purely process-oriented scientific purposes, another goal of biodiversity surveys is to obtain a significant, broad representation of evolutionary history for extant organisms (Sechrest et al., 2002). Phylogenetics and a model-based historical biogeographical approach are necessary to assess the amount of evolutionary novelty represented in different lineages. Our study has thus revealed that biodiversity hotspots are not solely areas of high specific richness and high levels of ende-

mism, but are also important reservoirs of a unique and threatened evolutionary history (Sechrest et al., 2002). This particularly applies to the Wallacea hotspot in which the *Achillides* originated and further diversified. Within this framework, conservation efforts may be directed to ameliorate the current anthropogenic threats confronting these unique swallowtails while preserving a representation of their evolutionary history (Sechrest et al., 2002; Mittermeier et al., 2004).

Acknowledgements

We thank all the collectors and providers of numerous specimens such as Laurent Soldati, Hervé Jourdan, Kotaro Saito, Yutaka Inayoshi and Kiyoshi Miura. We thank Anne-Laure Clamens for technical assistance with the molecular work, and Frédérique Serqueira from the SFR 119 ''Montpellier Environnement Biodiversité'' for help in the sequencing of some specimens. We are also grateful to Alexandre Dehne Garcia for help with computer analysis, as well as Marie Pagès for advice on species delimitation analyses. We thank the Environment Management of Province Sud (DENV) in New Caledonia for providing us with collecting permits and field assistance. Partial funding for this study was obtained through the programme ''ANR Biodiversité'' of the French National Agency for Research (Project ANR BIONEOCAL) and by a grant from the graduate school of Montpellier II University (SIBAGHE). F.A.H.S. acknowledges funding from an NSERC Discovery Grant. F.L.C. conceived the study, participated in its design and coordination, collected samples, analysed data and drafted and revised the manuscript. E.F.A.T. carried out the molecular genetic studies, participated in the sequence analysis and revised the manuscript. A.M.C. collected and identified swallowtail specimens and revised the manuscript. G.G. carried out the molecular genetic studies and sequencing. F.A.H.S. provided additional samples and revised the manuscript. G.J.K. funded the study, analysed the data and revised the manuscript. All authors read and approved the final manuscript.

References

Bauer, E., Frankenbach, T., 1998. Butterflies of the World. Pt.1. Papilionidae I: *Papilio* subgenus *Achillides*, *Bhutanitis*, *Teinopalpus*. Goecke und Evers, Keltern.

Becerra, J.X., Venable, D.L., 1999. Macroevolution of insect-plant associations: the relevance of host biogeography to host affiliation. Proc. Natl. Acad. Sci. USA 96, 12625–12631.

Bell, C.D., Soltis, D.E., Soltis, P.S., 2010. The age and diversification of the angiosperms re-revisted. Am. J. Bot. 97, 1296–1303.

Bininda-Emonds, O.R.P., Cardillo, M., Jones, K.E., MacPhee, R.D.E., Beck, R.M.D., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L., Purvis, A., 2007. The delayed rise of present-day mammals. Nature 446, 507–512.

de Boer, A.J., 1995. Islands and cicadas adrift in the West-Pacific, biogeographic patterns related to plate tectonics. Tijd. v. Ent. 138, 169–244.

Braby, M.F., Vila, R., Pierce, N.E., 2006. Molecular phylogeny and systematics of the Pieridae (Lepidoptera: Papilionoidea): higher classification and biogeography. Zool. J. Linn. Soc. 147, 239–275.

Britton, T., Anderson, C.L., Jacquet, D., Lundqvist, S., Bremer, K., 2007. Estimating divergence times in large phylogenetic trees. Syst. Biol. 56, 741–752.

Brown, J.M., Lemmon, A.R., 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. Syst. Biol. 56, 643–655.

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 gene for low level phylogenetics: elongation factor-1 alpha recovers morphology-based tree for heliothine moths. Mol. Biol. Evol. 12, 650–656.

Cluzel, D., Aitchison, J.C., Picard, C., 2001. Tectonic accretion and underplating of mafic terranes in the Late Eocene intraoceanic fore-arc of New Caledonia (Southwest Pacific): geodynamic implications. Tectonophysics 340, 23–59.

Condamine, F.L., Sperling, F.A.H., Wahlberg, N., Rasplus, J.-Y., Kergoat, G.J., 2012. What causes latitudinal gradients in species diversity? Evolutionary processes and ecological constraints on swallowtail biodiversity. Ecol. Lett. 15, 267–277.

Currie, D.J., Mittelbach, G.G., Cornell, H.V., Field, R., Guégan, J.-F., Hawkins, B.A., Kaufman, D.M., Kerr, J.T., Oberdorff, T., O''Brien, E., Turner, J.R.G., 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecol. Lett. 7, 1121–1134.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Edwards, S.V., 2009. Is a new and general theory of molecular systematics emerging? Evolution 63, 1–19.

Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52, 665–673.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer, Sunderland, MA.

Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15, 415–428.

Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Grandcolas, P., Murienne, J., Robillard, T., Desutter-Grandcolas, L., Jourdan, H., Guilbert, E., Deharveng, L., 2008. New Caledonia: a very old Darwinian island? Philos. Trans. R. Soc. B 363, 3309–3317.

Hall, R., 1998. The plate tectonics of Cenozoic SE Asia and the distribution of land and sea. In: Hall, R., Holloway, J.D. (Eds.), Biogeography and Geological Evolution of SE Asia. Backhuys Publishers, Leiden, pp. 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.

Hancock, D.L., 1983. Classification of the Papilionidae (Lepidoptera): a phylogenetic approach. Smithersia 2, 1–48.

Harada, M., 1992. The status of *Achillides bianor* (Cramer) in Sichuan, China. Butterflies 2, 19–25.

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). Available at: http://www.insects-online.de/frames/papilio.htm [accessed January 2012].
Heads, M., 2005. Dating nodes on molecular phylogenies: a critique of molecular biogeography. Cladistics 21, 62–78.
Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. Proc. Biol. Sci. 270, 313–321.
Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.
Hines, H.M., 2008. Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). Syst. Biol. 57, 58–75.
Ho, S.Y.W., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58, 367–380.
Holder, M., Lewis, P.O., 2003. Phylogeny estimation: traditional and Bayesian approaches. Nat. Rev. Genet. 4, 275–284.
Honza, E., Fujioka, K., 2004. Formation of arcs and backarc basins inferred from the tectonic evolution of Southeast Asia since the Late Cretaceous. Tectonophysics 384, 23–53.
Hug, L.A., Roger, A.J., 2007. The impact of fossils and taxon sampling on ancient molecular dating analyses. Mol. Biol. Evol. 24, 1889–1897.
Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., St. John, O., Wild, R., Hammond, P.M., Ahrens, D., Balke, M., Caterino, M.S., Gómez-Zurita, J., Ribera, I., Barraclough, T.G., Bocakova, M., Bocak, L., Vogler, A.P., 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. Science 318, 1913–1916.
Igarashi, S., 1984. The classification of the Papilionidae mainly based on the morphology of their immature stages. Trans. Lepido. Soc. Japan (Tyô To Ga) 34, 41–96.
Igarashi, S., Fukuda, H., 2000. The Life Histories of Asian Butterflies, Vol. 2. Tokai University Press, Tokyo.
Izumi, K., 1993. Geographic variation of wing shape in *Papilio bianor* (Insecta; Lepidoptera; Papilionidae). Bull. Biogeogr. Soc. Japan 48, 27–31.
Jordan, H.E.K., 1909. Fauna indoaustralica. In: Seitz, A. (Ed.), Macrolepidoptera of the World, Vol. 9. Kernen, Stuttgart, pp. 25–109.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.
Klaus, S., Schubart, C.D., Streit, B., Pfenninger, M., 2010. When Indian crabs were not yet Asian – biogeographic evidence for Eocene proximity of India and Southeast Asia. BMC Evol. Biol. 10, 287.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X, version 2.0. Bioinformatics 23, 2947–2948.
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. Ann. Rev. Ecol. Evol. Syst. 42, 205–226.
Lomolino, M.V., Riddle, B.R., Whittaker, R.J., Brown, J.H., 2010. Biogeography, 4th edn. Sinauer, Sunderland, MA.
Magallón, S., Sanderson, M.J., 2001. Absolute diversification rates in angiosperm clades. Evolution 55, 1762–1780.
Metcalfe, I., 2005. Asia: south-east. In: Selley, R.C., Cocks, L.R.M., Plimer, I.R. (Eds.), Encyclopedia of Geology, Vol. 1. Elsevier, Oxford, pp. 169–198.
Miller, K.G., Kominz, M.A., Browning, J.V., Wright, J.D., Mountain, G.S., Katz, M.E., Sugarman, P.J., Cramer, B.S., Christie-Blick, N., Pekar, S.F., 2005. The Phanerozoic record of global sea-level change. Science 310, 1293–1298.
Mittermeier, R.A., Robles Gil, P., Hoffmann, M., Pilgrim, J., Brooks, T., Mittermeier, C.G., Lamoreux, J.L., da Fonseca, G.A.B., 2004. Hotspots revisited: Earth's Biologically Richest and Most Threatened Ecoregions. CEMEX, Mexico City, Mexico.
Müller, C.J., Beheregaray, L.B., 2010. Palaeo island-affinities revisited – Biogeography and systematics of the Indo-Pacific genus *Cethosia* Fabricius (Lepidoptera: Nymphalidae). Mol. Phylogenet. Evol. 57, 314–326.
Müller, C.J., Wahlberg, N., Beheregaray, L.B., 2010. "After Africa": the evolutionary history and systematics of the genus *Charaxes* Ochsenheimer (Lepidoptera: Nymphalidae) in the Indo-Pacific region. Biol. J. Linn. Soc. 100, 457–481.
Munroe, E., 1961. The classification of the Papilionidae (Lepidoptera). Can. Entomol. Suppl. 17, 1–51.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Nylander, J.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.
Nylander, J.A.A., Olsson, U., Alström, P., Sanmartín, I., 2008. Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to dispersal-vicariance analysis of the Trushes (Aves: *Turdus*). Syst. Biol. 57, 257–268.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290.
Pelletier, B., 2006. Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. In: Payri, C.E., Richer de Forges, B. (Eds.), Compendium of Marine Species from New Caledonia Documents Scientifiques et Techniques, II4. Institut de Recherche pour le Développement, Nouméa, pp. 17–30.
Peña, C., Wahlberg, N., 2008. Prehistorical climate change increased diversification of a group of butterflies. Biol. Lett. 4, 247–278.
Pfeil, B.E., Crisp, M.D., 2008. The age and biogeography of *Citrus* and the orange subfamily (Rutaceae: Aurantioideae) in Australasia and New Caledonia. Am. J. Bot. 95, 1621–1631.
Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Pybus, O.G., Harvey, P.H., 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. Proc. Biol. Sci. 267, 2267–2272.
de Queiroz, A., 2005. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20, 68–73.
Rabosky, D.L., 2006a. Likelihood methods for detecting temporal shifts in diversification rates. Evolution 60, 1152–1164.
Rabosky, D.L., 2006b. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. Evol. Bioinform. Online 2, 257–260.
Rabosky, D.L., Lovette, I.J., 2008. Density-dependent diversification in North American wood warblers. Proc. R. Soc. B, 275, 2363–2371.
Ree, R.H., Sanmartín, I., 2009. Prospects and challenges for parametric models in historical biogeographical inference. J. Biogeogr. 36, 1211–1220.
Ree, R.H., Smith, S.A., 2008. Maximum-likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.
Ree, R.H., Moore, B.R., Webb, C.O., Donoghue, M.J., 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59, 2299–2311.

Reed, R.D., Sperling, F.A.H., 1999. Interaction of process partitions in phylogenetic analysis: an example from the swallowtail butterfly genus *Papilio*. Mol. Biol. Evol. 16, 286–297.
Rindal, E., Brower, A.V.Z., 2011. Do model-based phylogenetic analyses outperform parsimony? A test with empirical data. Cladistics 27, 331–334.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Salvo, G., Ho, S.Y.W., Rosenbaum, G., Ree, R., Conti, E., 2010. Tracing the temporal and spatial origins of island endemics in the Mediterranean region: a case study from the citrus family (*Ruta* L., Rutaceae). Syst. Biol. 59, 705–722.
Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.
Sanmartín, I., Ronquist, F., 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. 53, 216–243.
Sanmartin, I., van der Mark, P., Ronquist, F., 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. J. Biogeogr. 35, 428–449.
Scotese, C.R., 2001. Atlas of Earth History, Vol. 1. PALEOMAP Project, Arlington, TX.
Sechrest, W., Brooks, T.M., da Fonseca, G.A.B., Konstant, W.R., Mittermeier, R.A., Purvis, A., Rylands, A.B., Gittleman, J.L., 2002. Hotspots and the conservation of evolutionary history. Proc. Natl. Acad. Sci. USA 99, 2067–2071.
Shimogori, Y., 1997. *Achillides* Butterflies (Papilionidae), Vol. 6. Endless Science, Tokyo.
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, 113–137.
Sperling, F.A.H., 2003. Butterfly molecular systematics: from species definitions to higher-level phylogenies. In: Boggs, C.L., Watt, W.B., Ehrlich, P.R. (Eds), Butterflies: Ecology and Evolution Taking Flight. Chicago University Press, Chicago, IL, pp. 431–458.
Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657.
Vane-Wright, R.I., de Jong, R., 2003. The butterflies of Sulawesi: annotated checklist for a critical island fauna. Zool. Verh. Leiden 343, 3–267.
Voris, H.K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J. Biogeogr. 27, 1153–1167.
Wahlberg, N., Leneveu, J., Kodandaramaiah, U., Peña, C., Nylin, S., Freitas, A.V.L., Brower, A.V.Z., 2009. Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. Proc. Biol. Sci. 276, 4295–4302.
Wallace, A.R., 1860. On the zoological geography of the Malay Archipelago. J. Proc. Linn. Soc. Zool. 4, 172–184.
Wallace, A.R., 1863. On the physical geography of the Malay Archipelago. J. Roy. Geographic. Soc. London 33, 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.
Wiens, J.J., Donoghue, M.J., 2004. Historical biogeography, ecology and species richness. Trends Ecol. Evol. 19, 639–644.
Wilson, M.E., Moss, S.J., 1999. Cenozoic palaeogeographic evolution of Sulawesi and Borneo. Palaeogeogr. Palaeoclimatol. Palaeoecol. 145, 303–337.
Winkler, I.S., Mitter, C., Scheffer, S.J., 2009. Repeated climate-linked host shifts have promoted diversification in a temperate clade of leaf-mining flies. Proc. Natl. Acad. Sci. USA 106, 18103–18108.
Woodruff, D.S., Turner, L.M., 2009. The Indochinese–Sundaic zoogeographic transition: a description and analysis of terrestrial mammal species distributions. J. Biogeogr. 36, 803–821.
Yagi, T., Sasaki, G., Takebe, H., 1999. Phylogeny of Japanese papilionid butterflies inferred from nucleotide sequences of the mitochondrial ND5 gene. J. Mol. Evol. 48, 42–48.
Yagi, T., Sasaki, G., Omoto, K., 2006. Phylogeny, biogeography and wing pattern convergence of the subgenus *Achillides* (Lepidoptera, Papilionidae, *Papilio*) revealed by a mitochondrial DNA sequence analysis. Trans. Lepido. Soc. Japan 57, 137–147.
Yang, Z., Rannala, B., 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23, 212–226.
Yoshimito, H., 1998. *Papilio bianor* and *Papilio dehaanii*, two distinct species. Butterflies 20, 45–49.
Zachos, J.C., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to Present. Science 292, 686–693.
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.
Zeuner, F.E., 1943. Studies in the systematics of *Troides* Hübner (Lepidoptera Papilionidae) and its allies: distribution and phylogeny in relation to the geological history of the Australasian Archipelago. Trans. Zool. Soc. London 25, 107–184.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Map showing the region of study, distribution of the *Papilio* subgenus *Achillides* (black line) and delimited areas (red lines) used for biogeographical analyses. The region comprises 11 component areas: (A) East Palearctic (comprising central and northern China); (B) Japan archipelago; (C) India (encompassing Pakistan and Himalaya); (D) Indo-China (including non-peninsula Thailand, Laos, Cambodia, Vietnam, southern China and Myanmar); (E) Taiwan; (F) Sundaland (peninsular Thailand and Malaysia, Sumatra, Borneo and Java); (G) Philippines and Palawan; (H) Wallacea (including Lesser Sunda with all islands between Lombok and Tanimbar, including Timor and Wetar and the Moluccas with the islands of Ambon, Bacan, Buru, Halmahera, Morotai, Obi and Seram); (I) Irian Jaya, Papua New Guinea and Queensland (including Aru island); (J) Solomon Islands; and (K) New Caledonia and the Loyalty islands. Species richness is indicated by red dots as the number of species/subspecies that occur in each region. In the bottom-left corner, a map shows the delimitation of Wallacea, including various divisions of biotic affinities within the Indo-Australian Archipelago.

Appendix S2. Host plant preferences of the *Papilio* subgenus *Achillides*. These data are taken from Igarashi (1984) and Igarashi and Fukuda (2000). Pictures below show caterpillars of *P. dehaanii*, *P. paris* and *P. ulysses* on their host plant.

Appendix S3. Taxon sampling of *Papilio* subgenus *Achillides* used for this study. For each specimen, we specified the identification code (ID), the genus, species, author and date of description, subspecies if known, author and date of description, country, locality of capture and GenBank accession number for each molecular marker we used.

Appendix S4. DNA-based species delimitation with an extended-barcode marker (COI gene only).

Appendix S5. Temporal diversification of the *Papilio* subgenus *Achillides*. A lineages-through-time plot shows the number of lineages versus time (Ma). Superimposed palaeotemperature estimates (red line) are approximated by benthic $\delta^{18}O$ (Zachos et al., 2001). At the top, horizontal bars indicate whether the climate was warm (red) or cool (blue). Main climate changes for the last 20 Myr are indicated. Vertical teal bars indicate the significant effect of past climate on the diversification of *Achillides*. Light blue vertical bars, around 10.3, 7.4 and 3.9 Ma, identify three major shifts in diversification. At the bottom, the chronogram of *Achillides* is placed with a geological time scale from the Miocene to Holocene (H).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.