Bayesian Poisson tree processes and multispecies coalescent models shed new light on the diversification of Nawab butterflies in the Solomon Islands (Nymphalidae, Charaxinae, Polyura)

EMMANUEL F. A. TOUSSAINT1,2*, JÉRÔME MORINIÈRE2, ATHENA LAM2, BERNARD TURLIN3 and MICHAEL BALKE FLS2,4

1Department of Ecology & Evolutionary Biology & Division of Entomology, Biodiversity Institute, University of Kansas, Lawrence, KS, 66045, USA
2SNSB-Bavarian State Collection of Zoology, Münchhausenstraße 21, 81247 Munich, Germany
314 Résidence du Nouveau Parc, 78570 Andrésy, France
4GeoBioCenter, Ludwig-Maximilians University, Munich, Germany

Received 22 July 2015; revised 17 December 2015; accepted for publication 7 February 2016

Butterflies of the genus Polyura form a widespread tropical group distributed from Pakistan to Fiji. The rare endemic Polyura epigenes Godman & Salvin, 1888 from the Solomon Islands archipelago represents a case of marked island polymorphism. We sequenced museum specimens of this species across its geographic range to study the phylogeography and genetic differentiation of populations in the archipelago. We used the Bayesian Poisson tree processes and multispecies coalescent models, to study species boundaries. We also estimated divergence times to investigate the biogeographic history of populations. Our molecular species delimitation and nuclear DNA network analyses unambiguously indicate that Malaita populations form an independent metapopulation lineage, as defined in the generalized lineage concept. This lineage, previously ranked as a subspecies, is raised to species rank under the name Polyura bicolor Turlin & Sato, 1995 stat. nov. Divergence time estimates suggest that this lineage split from its sister taxon in the late Pleistocene. At this time, the bathymetric isolation of Malaita from the rest of the archipelago probably prevented gene flow during periods of lower sea level, thereby fostering allopatric speciation. The combination of molecular species delimitation methods, morphological comparisons, and divergence time estimation is useful to study lineage diversification across intricate geographic regions.

doi: 10.1111/zoj.12413

comparative way; however, these recent developments are only starting to be reflected in the literature, and there is still a need to assess the performance and limits of such methods in an empirical framework (e.g. Toussaint et al., 2015).

The systematics and taxonomy of butterflies in the genus Polyura Billberg, 1820, have been revised recently based on a comprehensive molecular phylogeny combined with molecular species-delimitation analyses (Toussaint et al., 2015). However, a few difficult cases were not solved because of a lack of phylogenetic resolution and/or geographic sampling. The genus comprises 33 species of brush-footed butterflies in the Oriental region and Indo-Australian archipelago (Smiles, 1982; Toussaint et al., 2015, in press). Multiple island endemics occur from the Sunda Islands east to the Pacific Islands. In the past century, multiple taxa have been described in various taxonomic ranks, resulting in a thorny taxonomy (Smiles, 1982). The rare, endemic Polyura epigenes Godman & Salvin, 1888 from the Solomon Islands archipelago is one of the puzzling cases within Polyura. The Solomon Islands archipelago is a complex, geological assemblage of recent volcanic terranes and more ancient blocks resulting from Cenozoic plate subduction (Hall, 2002). During the Plio-Pleistocene, sea-level fluctuations allowed the connection of some islands via land bridges. It is therefore an appealing setting to test biogeographic and phylogeographic hypotheses (Mayr & Diamond, 2001). Polyura epigenes is found across most of the Solomon Islands archipelago, where three morphologically distinct populations were described as subspecies (Fig. 1). Among these Polyura epigenes bicolor Turlin & Sato, 1995 from Malaita presents the most derived morphology (Fig. 1). The female has two morphs: orange and white. The orange morph resembles females of Polyura gamma (Lathy, 1898). It is markedly different from the usual white female of P. epigenes. Males of P. epigenes bicolor also have a narrower shape of the wings compared with other subspecies, and have more spotting on the upperside of the hindwing. Some differences in the male genitalia of P. epigenes bicolor and the rest of the populations have also been suggested (Müller & Tennent, 1998). There has been some debate concerning the status of this rather unique taxon (Müller & Tennent, 1998; Turlin, 2001). Toussaint et al. (2015) obtained contrasting phylogenetic hypotheses depending on the optimality criterion used (Bayesian inference versus maximum likelihood), therefore preventing a well-supported hypothesis of boundaries in this species complex.

Here, we sequenced previously unavailable specimens of P. epigenes from Malaita and Bougainville islands (Fig. 1). We aim to reconstruct relationships among populations of this taxon with the same optimality criteria used in Toussaint et al. (2015), and use molecular species delimitation methods to investigate genetic differentiation of Malaita populations with respect to other island populations. To do so, we rely on a recently introduced Bayesian implementation of the Poisson tree processes model (PTP; Zhang et al., 2013) and unguided species delimitation, as implemented in BAYESIAN PHYLOGENETICS AND PHYLOGEOGRAPHY (BPP; Yang, 2015). Finally, we aim to estimate the divergence time between populations dwelling in the Solomon Islands in order to understand the impact of geologic and climatic disruptions on the genetic structure observed in the archipelago today.

MATERIAL AND METHODS

TAXON SAMPLING AND MOLECULAR BIOLOGY

We retrieved sequence data from Toussaint et al. (2015), and sequenced two markers from additional specimens of Polyura epigenes monocroma Niepelt, 1914 and three additional specimens of P. epigenes bicolor to improve our taxonomic sampling. One specimen of P. epigenes bicolor (ET0219) from Toussaint et al. (2015) was removed from further analyses because it was only sequenced for a small portion of the cytochrome c oxidase subunit I (COI) gene fragment. As a result, this specimen would branch in a variety of awkward positions, and was considered useless to test the monophyly of P. epigenes bicolor and for species delimitation purposes. Total genomic DNA was extracted from leg tissues of dried collection specimens using the DNeasy kit (Qiagen, Hilden, Germany). Using polymerase chain reaction (PCR) protocols from Müller, Wahlberg & Beheregaray (2010), we amplified and then sequenced the following gene fragments: COI (471 bp) and NADH dehydrogenase subunit 5 (ND5; 417 bp). Because the newly extracted specimens are old collection vouchers (on average >15 years old), we failed to amplify the two nuclear markers ribosomal protein S5 (RPS5) and wingless (WGL). DNA sequences were edited in GENEIOUS R8 (Biomatters, http://www.geneious.com), aligned using MUSCLE (Edgar, 2004), and with the reading frames inferred with MESQUITE 3.02 (http://mesquiteproject.org). The different data sets used to infer phylogenetic relationships were generated under MESQUITE. All new sequences were deposited in GenBank (accession nos KU980225–KU980234).

MOLECULAR PHYLOGENETICS

We used Bayesian inference (BI) and maximum likelihood (ML) to reconstruct the phylogenetic
relationships. Partitions and corresponding optimal models of substitution were searched under PartitionFinder 1.1.1 (Lanfear et al., 2012) using the GREEDY algorithm, either with the MrBayes or the RAxML set of models, because MrBayes 3.2.5 (Ronquist et al., 2012) and RAxML GUI 0.93 (Stamatakis, 2006; Silvestro & Michalak, 2012) implement different sets of substitution models. The corrected Akaike information criterion (AICc) was used to compare the fit of the different models. The BI analyses were performed using MrBayes 3.2.5 (Ronquist et al., 2012). Two simultaneous and independent runs, consisting of eight Metropolis-coupled Markov chain Monte Carlo simulations (MCMCs, one cold and seven incrementally heated), running 20 million generations, were used. Trees were sampled every 1000 generations to calculate posterior probabilities (PP). In order to investigate the convergence of the runs we investigated the split frequencies and effective sample size (ESS) of all the parameters, and plotted the log-likelihood of the samples against the number of generations in TRACER 1.6 (http://BEAST.bio.ed.ac.uk/Tracer). An ESS exceeding 200 was acknowledged as a good indicator of convergence. All trees sampled prior to reaching the log-likelihood plateau were discarded as burn-in, and the remaining samples were used to generate a 50% majority rule consensus tree. The ML analyses were conducted with the best partitioning scheme selected in PartitionFinder 1.1.1 (Lanfear et al., 2012), using RAxML GUI 0.93 (Stamatakis, 2006; Silvestro & Michalak, 2012). We performed 1000 thorough bootstrap replicates (BS) to investigate the level of support at each node.

**Molecular species delimitations**

Following Toussaint et al. (2015), we used different methods to investigate the genetic differentiation of...
populations of *P. epigenes* in the Solomon Islands archipelago.

First, we used the Bayesian implementation of the bPTP (Zhang *et al*., 2013) to infer molecular phylogenies. The PTP model allows us to distinguish speciation and coalescent processes along the branches of a topology. It requires a phylogenetic tree as an input, with branch lengths representing the number of mutations. The PTP model estimates the mean expected number of substitutions per site between two branching events using the branch length information of the input phylogeny. It then implements two independent classes of Poisson processes (intra- and interspecific branching events), and clusters the phylogenetic tree accordingly. The Bayesian implementation of the PTP model allows for the computation of posterior probabilities as: (the number of occurrences of all of the descendants under a given node)/(the number of samples from MCMC sampling). These represent the posterior probability of taxa to form one species under the PTP model with flat prior. Resulting support values are strongly correlated with the accuracy of the delimitation. The analyses were conducted on the web server for bPTP (http://species.h-its.org/ptp/) using the MrBayes and RAxML topologies. Each analysis consisted of 500 000 generations, with a thinning every 100 generations and a burn-in of 25%. We also performed separate analyses including or excluding *Charaxes fournierae* Le Moul, 1930 and *Polyura posidonius* (Leech, 1891) in order to test the impact of out-group specification on species delimitation consistency.

Second, we used Bayesian species delimitation as implemented in BPP 3.1 (Yang, 2015). This method uses the multispecies coalescent model to compare different models of species delimitation and species phylogeny in a Bayesian framework. BPP can account for incomplete lineage sorting resulting from ancestral polymorphism and gene tree–species tree conflicts (Yang & Rannala, 2010, 2014; Rannala & Yang, 2013; Yang, 2015). The model includes two types of parameters: the species divergence times (*τ* *s*) and the population size parameters for both modern and ancestral species (*θ* *s*). The gamma prior G*θ* *s*(*α*,*β*), with mean *α*/*β*, is used on the population size parameters (*θ* *s*). The parameters *α* and *β* represent the shape and the rate of the gamma prior, respectively. The age of the root in the species tree (*τ* *s*) is assigned the gamma prior G*τ* *s*(*α*,*β*), whereas the other divergence time parameters are assigned the Dirichlet prior (Yang & Rannala, 2010: equation 2). We performed two different types of analyses to test species boundaries using *P. epigenes bicolor* as a putative species, and therefore with all remaining *P. epigenes* specimens as a different species referred to as *P. epigenes*.

We conducted a first analysis where a reversible-jump MCMC (rjMCMC) algorithm is used to move between different species-delimitation models that are compatible with a fixed guide tree (Yang & Rannala, 2010; Rannala & Yang, 2013). This analysis named A10 in BPP was set up by specifying speciesdelimitation = 1 and speciestree = 0 in the BPP control file. We used *BEAST* 1.8.2 (Heled & Drummond, 2010) to estimate the species tree based on the four alignments, and assigned each specimen to its corresponding putative species. The optimal model of substitution for each gene fragment was inferred using PartitionFinder 1.1.1 (Lanfear *et al*., 2012) using the GREEDY algorithm and the BEAST set of models. We specified an uncorrelated lognormal prior for the clock of each gene fragment, a Yule process model as species tree prior, and a piecewise constant population size model. The analysis consisted of 50 million generations, with a sampling interval of 5000 and a conservative burn-in of 25%.

We also performed another analysis where the algorithm explores different species delimitation models using the rjMCMC algorithm and different species phylogenies using the nearest-neighbour interchange (NNI) or subtree pruning and regrafting (SPR) algorithms (Yang & Rannala, 2014). Under this analysis (A11 in BPP), different putative species can be merged into one but can never be split into additional putative species. We set up the A11 analysis by specifying speciesdelimitation = 1 and speciestree = 1 in the BPP control file.

As advocated by Leaché & Fujita (2010), we conducted different sets of analyses with different values of *α* and *β*, allowing *θ* *s* and *τ* *s* to account for: (i) large ancestral population sizes and deep divergence between species, using *Gθ* *(1,10)* and *Gτ* *(1,10)*; and (ii) small ancestral population sizes and shallow divergence between species, using *Gθ* *(2,2000)* and *Gτ* *(2,2000)*. All A10 and A11 analyses were performed with the following settings: algorithm = 1, *α* = 1, and *m* = 2; speciesmodelprior = 1; usedata = 1; and cleandata = 0. The rjMCMC analyses consisted of 100 000 generations (sampling interval of 10) with 25 000 samples being discarded as burn-in. Each analysis was run three times using different starting points to confirm consistency between runs.

In order to make decisions on the status of a particular taxon, we adopted the generalized lineage concept, as defined by De Queiroz (2005, 2007). Under this concept, criteria usually assumed as a single indicator of species boundaries (morphology, monophyly, or reproductive isolation) are treated as attributes that accumulate throughout the process of
lineage diversification (De Queiroz, 2005; Carstens et al., 2013). In our study, we use the consistency of molecular species delimitations and nuclear gene differences as proxies for lineage divergence.

**Nuclear DNA analysis**

We inferred a haplotype network based on a concatenated matrix comprising both nuclear genes for all specimens. The out-group Charaxes fournierae was removed from this analysis to enhance the resolution. The network was reconstructed using SplitsTree 4.13.1 (Huson & Bryant, 2006) with calculated uncorrected p-distances and the NeighborNet algorithm. We ran 1000 bootstrap replicates to test the robustness of the inferred relationships.

We also reconstructed a nuclear DNA phylogenetic hypothesis using RAxML GUI 0.93 (Stamatakis, 2006; Silvestro & Michalak, 2012). The concatenated nuclear data set was left unpartitioned and the best-fitting model of substitution was set to a GTR+Γ+I model, based on the result of PartitionFinder 1.1.1 (Lanfear et al., 2012).

**Divergence time estimation**

In order to obtain an estimation of divergence times between *P. epigenes bicolor* and other populations of *P. epigenes*, we used *BEAST* 1.8.2 (Heled & Drummond, 2010). We tested the molecular clock hypothesis using PAUP* (Swofford, 2003), and as it was significantly rejected (*P* < 0.001), we used an uncorrelated lognormal Bayesian relaxed clock for each gene fragment, allowing rate variation among lineages. We used two substitution rates calculated for the COI and ND5 genes in Andújar, Serrano & Gómez-Zurita (2012). In this paper, the authors estimated substitution rates in the genus Carabus (Coleoptera, Carabidae) using multiple geological and fossil calibrations. The *ucld.mean* parameter of the COI and ND5 genes in Andújar et al. (2012). In this paper, the authors estimated substitution rates in the genus Carabus (Coleoptera, Carabidae) using multiple geological and fossil calibrations. The *ucld.mean* parameter of the COI and ND5 genes in Andújar et al. (2012). Specifically, we enforced the 95% credibility interval of 0.008–0.0146 for the COI gene fragment *ucld.mean* using a soft-normal prior (*mean* = 0.0113, SD = 0.0016835). We also enforced the 95% credibility interval of 0.012–0.0198 for the ND5 fragment *ucld.mean* using a soft-normal prior (*mean* = 0.0159, SD = 0.00199). The other genes were each assigned an uninformative interval of 0.001–0.05, with a uniform distribution. All parameters and post-analytic procedures were identical to those performed in the *BEAST* analysis of the BPP species delimitation section.

**RESULTS**

**Phylogenetic relationships**

Our BI and ML analyses yielded identical phylogenetic trees (Fig. 2). The in-group relationships among species are well supported. For the outgroups, the sister relationship of *Polyura cognatus* Vollenhoven, 1861 and *Polyura caphontis* (Hewitson, 1863) + *Polyura sacco* Smart, 1977 was poorly supported in both analyses. *Polyura dehannii* (Westwood, 1850) is found as sister to *P. epigenes* in both analyses, with strong (BI) and moderate (ML) support. *Polyura epigenes monochromata* is found nested within *P. epigenes monochromata*. This cluster is sister to *P. epigenes bicolor*. The genetic differentiation of *P. epigenes bicolor* from the two other subspecies is supported in both BI and ML analyses.

The *BEAST* analyses resulted in a highly supported species tree (Fig. 2), with the unique difference compared with the BI and ML analyses being the placement of *P. cognata* as sister to *P. dehannii + P. epigenes*, with moderate support (PP = 0.91).

**Molecular species delimitations**

The results of the different species delimitation analyses conducted are summarized in Tables 2 and 3, and Figures 2 and 3. Both methods (bPTP and BPP) yielded consistent results with different levels of support, depending on the parameters and priors used. Both methods indicate that *P. epigenes bicolor* is an independent metapopulation lineage, as defined in the generalized lineage concept (De Queiroz, 2007), except for one analysis of bPTP based on the RAxML topology including all out-groups (model M1). In this analysis, all subspecies of *P. epigenes* are inferred to be unique species with low support (PP = 0.71). All other analyses performed in bPTP support *P. epigenes bicolor* as a valid species. The bPTP analyses based on the MrBayes topology resulted in better support compared with analyses based on the RAxML topology. Likewise, the level of support for *P. epigenes bicolor* as a valid species increased gradually as out-groups were removed (Table 2, Fig. 3). All BPP analyses suggested that *P. epigenes bicolor* is a distinct species with maximal support, both with and without a guide tree. The use of very loose $Gθa$ and $Gτ0$ priors did not influence the outcome of the BPP analyses.

**Nuclear DNA analysis**

The haplotype network reconstructed based on nuclear DNA sequence data is presented in Figure 4, along with the RAxML phylogeny based on the same data set. Both analyses recover a clear split between
Figure 2. Molecular phylogeny and molecular species delimitation results. Bayesian phylogeny as recovered from MrBayes analyses. The nodal supports of both Bayesian-inference (BI) and maximum-likelihood (ML) analyses are shown, with colour coding as indicated in the figure. The cloudogram of the *BEAST analysis showing all posterior species trees is presented at the bottom left. Denser regions indicate a robust support for the inferred relationship.
the unique Malaita specimen and the three other *P. epigenes*
specimens. A careful examination of the individual gene fragment
alignments indicates multiple substitutions between these two lineages,
as highlighted in Figure 4.

**DIVERGENCE TIME ESTIMATION**

The *BEAST* analyses converged, with all parameters having ESS values over 500. The resulting chronogram is presented in Figure 5, with the median ages and 95% height posterior distributions (95% HPDs) computed from the post burn-in posterior topologies. We estimate a split between *P. posidonius* and the group of species from the *Polyura pyrrhus* group at c. 3.1 Mya (95% HPD: 1.79–5.04 Mya). The split between *P. dehanii* and the lineages from the Solomon Islands is estimated at c. 1.3 Mya (95% HPD: 0.43–2.27 Mya). Finally, the split between *P. epigenes bicolor* and other populations of *P. epigenes* is estimated at c. 250 Kya (95% HPD: 0.07–0.52 Kya).

**TAXONOMY**

Based on these analyses, we raise the subspecies *P. epigenes bicolor* to species rank with the name:

*Polyura* Billberg, 1820

*Polyura bicolor* (Turlin & Sato, 1995) **stat. nov.**


The male holotype (‘Malaita Islands, Solomons archipelago, VIII-1990’) of this species is housed in the Museum of Nature and Human Activities (MNHA), Hyogo, Japan. Twenty-three paratypes were designated (one with the label ‘Allotype’), and are located in the MNHA, as well as in the collections of Bernard Turlin and Hidetsugu Sato, the original authors of this species (Turlin & Sato, 1995).

**Diagnosis**

*Polyura bicolor* males are more brownish than the darker *P. epigenes* (Fig. 6). The apex of the forewing is more pointed in *P. bicolor* and the outer margin is more deeply concave than in *P. epigenes*. The five discal spots and the eight submarginal dots of the upperside of the forewing are larger and white in *P. bicolor* compared with *P. epigenes*, where they are smaller and ochreous in ssp. *epigenes* and obsolete or even absent in ssp. *monochroma*. The hindwing submarginal spots are blue in *P. epigenes* but are light violet in *P. bicolor*. Male genitalia differ by the sociuncus deeply indented ventrally in *P. epigenes* (almost not in *P. bicolor*), and by the sharp projections of the valva directed anteriorly in *P. epigenes* (ventrally in *P. bicolor*). Females of the common form of *P. bicolor* and of *P. epigenes* are pretty similar. The wide median bar is cream in *P. epigenes* but is white in the common form of *P. bicolor*. The orange form

---

**Figure 3.** Comparison of species delimitation results. Graph showing the result of each analysis of species delimitation. The posterior probability of either *Polyura bicolor* or *Polyura epigenes* (including *Polyura bicolor*) as a valid species are shown.
cinereus is very distinct (Fig. 5), with all light areas ochreous, and the external suffusion to the median bar of the hindwing is golden brown.

With the limited taxon sampling of *P. epigenes epigenes*, we choose not to synonymize *P. epigenes monochroma* with *P. epigenes epigenes*. Based on the bathymetry of the archipelago, it is likely that these two populations are not genetically distinct and do constitute a unique lineage. Additional taxon sampling will be needed to assess the potential genetic differentiation of these two populations.

**DISCUSSION**

**SPECIES DELIMITATION METHOD: PERFORMANCE AND POSSIBLE BIASES**

We find that bPTP and BPP deliver consistent results, although there are differences in the way that priors affect the delimitation between the two methods. In bPTP, the input topology and, in particular, the pruning of out-groups has some impact on the resulting delimitation (Fig. 3; Table 1). When using RAxML or MrBayes topologies, the results were generally consistent; however, when using a reduced set of out-groups the level of support for the different species differs. It seems that bPTP gives more similar results compared with BPP when distant out-groups are removed. Additional empirical studies using topologies with more or fewer out-groups are needed to assess the real impact of these factors on the statistical support of delimited putative species. In BPP, the use of very different $G_h$ and $G_s$ priors did not change the results of species delimitations; however, Toussaint *et al.* (2015) found that these priors could have an impact in the pyrrhus complex, a clade comprising several closely related species but presenting reliable diagnostic morphological features. As recommended by Leaché & Fujita (2010), a comparative strategy should always be adopted when using BPP, especially in the case of cryptic complexes. In contrast with Toussaint *et al.* (2015), we did not run Bayesian Generalized Mixed Yule Coalescent (bGMYC) analyses for two reasons. First, the latter study showed that this method is clearly suboptimal to delineate species, compared with the two other methods used in this.
paper (BPP and PTP). Second, bGMYC uses a collection of posterior gene trees from an independent analysis as an input. In Toussaint et al. (2015), the failure of bGMYC to delimit Malaita populations was likely to be accounted for by the lack of COI sequence data for the dubious specimen ET0219, and by the presence of a unique specimen in the nuclear gene trees.

Here, we extended our taxon sampling to enhance the phylogenetic resolution and therefore the accuracy of molecular species delimitations; however, because *P. epigenes* is a rare species, and the Solomon Islands archipelago is a remote location, we were not able to sample populations from all islands where this species occurs (i.e. Choiseul, New Georgia islands, San Isabel, Shortland Islands, Vella Lavella). Theoretically, this lack of geographic sampling could generate an artefactual genetic gap in the data set; however, we argue that this is unlikely, as we sampled all three morphological subspecies of *P. epigenes*. The island populations that we did not sample belong to *P. epigenes monochroma*, which is

---

**Figure 5.** Divergence time estimates derived from the *BEAST* analysis. Chronogram derived from the posterior trees of dating analysis conducted in BEAST. The 95% credibility intervals are shown at each node of the phylogeny. A map of the Solomon Islands with bathymetry is shown at the top of the figure. The island of Malaita is highlighted in violet. A picture of a female *Polyura epigenes bicolor* f. *cinereus* (orange morph) is presented. Picture taken by Bernard Turlin. T in time bar stands for Tarantian.
widespread in the Solomon Islands archipelago (Fig. 1). These populations are morphologically identical to the subspecies that we sampled, and therefore are unlikely to represent independent metapopulation lineages, especially considering that the morphologically deviant *P. epigenes epigenes* is found nested within *P. epigenes monochroma*. It is difficult to completely rule out this hypothesis, however, and therefore it would be crucial to integrate these unsampled populations in the future to revisit the status of populations in the Solomon Islands archipelago. Although we fully agree with recent recommendations to conduct molecular species delimitation (Carstens *et al.*, 2013), we argue that it might be unrealistic in most cases to fulfill the strict criteria deemed compulsory in most cases (see e.g. Lim, Balke & Meier, 2012). In particular, assembling a large geographic sample and generating large molecular matrices (with more than ten gene fragments) is economically not realistic at present, and/or is unachievable when working with museum collections, for instance. This should not prevent taxonomic acts, however, as long as molecular species delimitations are performed in a comparative manner with the use of different models and priors to assess the consistency across methods (Table 2).

**ON SUBSPECIES AND SPECIES IN A GENERALIZED LINEAGE CONCEPT**

Populations from Malaita were originally described as a subspecies by Turlin & Sato (1995). In their paper, the authors noted that the new subspecies *P. epigenes bicolor* ‘differs strikingly’ from the two other subspecies *P. epigenes epigenes* and *P. epigenes monochroma*. The two authors made reference to external...
morphological characters such as coloration and the size of both females and males. In an independent paper, Müller & Tennent (1998) overlooked the original description of this taxon and described it as a valid species under the name *P. thayn*. The authors confirmed the clear differentiation of Malaita populations from the rest of the archipelago when looking at external morphological characters; however, they also dissected the male genitalia of both lineages, and found several notable differences, upon which they argued

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Subspecies</th>
<th>Country</th>
<th>Locality</th>
<th>Code</th>
<th>COI</th>
<th>ND5</th>
<th>RPS</th>
<th>wingless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charaxes</td>
<td>fournierae</td>
<td>jolybouyeri</td>
<td>Guinea</td>
<td>-</td>
<td>EV-0006</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Polyura</td>
<td>caphontis</td>
<td>caphontis</td>
<td>Fiji</td>
<td>Biasevu, Viti, Levu Island</td>
<td>ET165</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>cognata</td>
<td>bellona</td>
<td>Indonesia</td>
<td>Bantimurung, Maros, Southern Sulawesi</td>
<td>ET43</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>cognata</td>
<td>bellona</td>
<td>Indonesia</td>
<td>Bantimurung, Maros, Southern Sulawesi</td>
<td>ET66</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>cognata</td>
<td>cognata</td>
<td>Indonesia</td>
<td>Camba, Southern Sulawesi</td>
<td>ET44</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>cognata</td>
<td>cognata</td>
<td>Indonesia</td>
<td>Papayato, Northern Sulawesi</td>
<td>ET77</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>cognata</td>
<td>yunikoe</td>
<td>Indonesia</td>
<td>Peleng Island, Sulawesi</td>
<td>ET131</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>dehanii</td>
<td>dehanii</td>
<td>Indonesia</td>
<td>Gunung Halimun, West Java</td>
<td>ET132</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>dehanii</td>
<td>dehanii</td>
<td>Indonesia</td>
<td>Gunung Halimun, West Java</td>
<td>ET51</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>dehanii</td>
<td>sulthan</td>
<td>Indonesia</td>
<td>Gunung Sanggul, West Sumatra</td>
<td>ET86</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>dehanii</td>
<td>sulthan</td>
<td>Indonesia</td>
<td>Susuk, North Sumatra</td>
<td>ET183</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>bicolor</td>
<td>Solomon Islands</td>
<td>Maluu, North Malaita Island</td>
<td>ET41</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>bicolor</td>
<td>Solomon Islands</td>
<td>New Mera Village, Malaita</td>
<td>ET224</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>bicolor</td>
<td>Solomon Islands</td>
<td>Maluu, North Malaita Island</td>
<td>ET247</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>bicolor</td>
<td>Solomon Islands</td>
<td>Maluu, North Malaita Island</td>
<td>ET248</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>bicolor</td>
<td>Solomon Islands</td>
<td>Maluu, North Malaita Island</td>
<td>ET249</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>epigenes</td>
<td>Solomon Islands</td>
<td>Guadalcanal Island</td>
<td>ET215</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>monochroma</td>
<td>Papua</td>
<td>Bougainville Island</td>
<td>ET168</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>monochroma</td>
<td>Papua</td>
<td>Buka Island</td>
<td>ET211</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>monochroma</td>
<td>Papua</td>
<td>Buka Island</td>
<td>ET214</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>monochroma</td>
<td>Papua</td>
<td>Bougainville Island</td>
<td>ET241</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>epigenes</td>
<td>monochroma</td>
<td>Papua</td>
<td>Bougainville Island</td>
<td>ET244</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>posidonius</td>
<td>–</td>
<td>China</td>
<td>Chengdu, 2200 m a.s.l., Tibet</td>
<td>ET207</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Polyura</td>
<td>sacco</td>
<td>sacco</td>
<td>Vanuatu</td>
<td>–</td>
<td>ET213</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
that Malaita populations were a valid species. In order to clarify the taxonomic situation of Malaita populations, Turlin synonymized *P. thayn* with *P. epigenes bicolor* (Turlin, 2001). Using a comprehensive molecular data set, Toussaint et al. (2015) noted that Malaita populations were recognized as a valid species by some species delimitation methods, but not by others, arguing that the taxon sampling was likely to be the cause of these inconsistencies. In the present study, a broader taxon sampling helps to shed light on the clear genetic demarcation of Malaita populations from the rest of *P. epigenes* populations across the Solomon Islands. If the divergence between these two lineages seems unquestionable, the taxonomic ranking of Malaita populations as the valid species *P. bicolor*, rather than as a subspecies of *P. epigenes*, requires additional justification. Here, we base our decision on several criteria mostly derived from the generalized lineage concept, as defined by De Queiroz (2007), and a recent review of subspecies definition (Braby, Eastwood & Murray, 2012).

De Queiroz (2007) advocated the combined use of multiple lines of evidence to refine species boundaries, as different lines of evidence taken individually is likely to result in different species concepts. In their review, Braby et al. (2012) extended this principle to subspecies, and recommended that lineages for which the taxonomic rank is uncertain should be assessed with a series of criteria that would grant a species-level ranking. Examples of such criteria include, but are not restricted to, differences in external morphology and genitalia, mitochondrial lineage sorting (monophyly of lineages based on mitochondrial DNA only), or different caterpillar host plants. If the candidate species does not meet any of these criteria, it should be conservatively treated as a subspecies. In practice, once more comprehensive studies have been conducted on wider spread species, such cases might be common, in particular in the case of remote localities combined with a lack of data, a situation that gives rise to taxonomic dilemmas (Skale et al., 2012). In the case of *P. bicolor*, however, the species ranking is substantiated by a number of features compared with *P. epigenes*: different male genitalia, different external morphologies, mitochondrial lineage sorting, and nuclear differentiation (Table 3).

### Biogeography of Polyura butterflies in the Solomon Islands

The factors responsible for the isolation of *P. bicolor* from other *P. epigenes* populations in the Solomon Islands archipelago can be inferred by screening the geological and climatic history of the region. In this study, we recover a divergence time estimate for the split between *P. bicolor* and *P. epigenes* of c. 400 Kya (Fig. 5). To date there is no published, dated phylogeny of *Polyura*, but provisional dating of the full genus (E.F.A. Toussaint, unpubl. data), based on secondary calibrations derived from Wahlberg et al. (2009), suggests a similar age, with a divergence between *P. bicolor* and *P. epigenes* in the late Pleistocene. Speciation events dated from this period are not common in butterflies, although some cases of allopatric speciation have been suggested. This is the case, for example, of the Mediterranean species *Erebia triaria* (de Prunner, 1798) (*Nymphalidae*, *Satyrinae*), in which a genetically very distinct lineage was hypothesized to be the result of isolation in a refuge during a Pleistocene glacial maximum (Vila, Vidal-Romani & Björklund, 2005). In oceanic settings, few examples of such Quaternary allopatric speciation have been documented. *Delias* butterflies (*Pieridae*, *Pierinae*) comprise some pairs of sister species, the divergence of which was estimated from the late...
Pleistocene (Müller, Matos-Maravi & Beheregaray, 2013). Likewise, the divergence between the sister species Charaxes antonius Semper, 1878 (Nymphalidae, Charaxinae) and Charaxes sangana Schroder & Treadaway, 1988, both endemic to the Philippines, was dated from the mid-Pleistocene (Müller et al., 2010). In the tribe Troidini (Papilionidae, Papilioninae), the genera Ornithoptera, Trogonoptera, and Troides both comprise possible cases of allopatric speciation in the Indo-Australian archipelago (Condamine et al., 2015); however, these events of speciation are slightly older, with the splits between sister allopatric lineages being dated from the early to mid-Pleistocene. In Papilio (Heraclides) butterflies (Papilionidae, Papilioninae), the split between the Cuban endemic Papilio caiguanabus (Poey, 1851) and its sister species Papilio aristor Godart, 1819, endemic to Haiti and the Dominican Republic, has been dated to the early Pleistocene (Lewis et al., 2015). Other examples in oceanic butterfly clades represent comparatively older cases of such allopatric vicariance (e.g. Müller & Beheregaray, 2010; Condamine et al., 2013; Matos-Maravi et al., 2014). Hence, the divergence of *P. bicolor* and *P. epigenes* is an uncommon example of recent oceanic allopatric speciation in butterflies. Except from the biogeographic trigger of lineage divergence, the possible mechanisms responsible for such a fast speciation remain elusive. Adaptive radiation with respect to host plant preferences seems unlikely because of the time needed for insect–plant coevolution to occur, especially in archipelagic settings (Adler & Dudley, 1994). Other adaptive scenarios in relation to habitat and ecology are possible, although at this stage our data set is not suitable for in-depth testing of such hypotheses. A possible cause for such recent speciation is a small initial population size of the Malaita lineage, which would foster rapid genetic change as a result of enhanced genetic drift and selective pressures (e.g. Jordan & Snell, 2008); however, testing this hypothesis would also require additional data.

The geological arcs found in the Solomon Islands archipelago are the result of Cenozoic plate tectonics (Hall, 2002). All islands currently subaerial in the archipelago have never been connected to any continental land. The islands where *P. epigenes* occurs presently are the result of an intense volcanism that took place from the Pliocene to present, whereas Malaita, where *P. bicolor* dwells, is of more ancient volcanic origin between the Eocene and the early Miocene (Mayr & Diamond, 2001; Hall, 2002). Hence, it seems unlikely that the genetic differentiation of these two species was shaped by geological factors, as most of the archipelagic assemblage largely predates this speciation event; however, the paleobathymetry of the archipelago is particularly relevant in the evolution of these two lineages. Indeed, most islands of the archipelago are separated by shallow water corridors (< 200 m, most of these < 100 m), except for Malaita and San Cristobal, which are surrounded by deeper waters (Fig. 5). From the Pliocene to present, glaciation cycles have triggered global sea-level fluctuations (Miller et al., 2005). During glaciation periods, the sea level dropped as glaciers formed and polar ice caps expanded, connecting islands separated by such shallow waters. As a result, most islands surrounded by light blue in Figure 5 would have been periodically connected throughout the past 5 Myr, whereas Malaita and San Cristobal would have been isolated. The large string of islands regrouping Buka, Bougainville, Choiseul, Santa Isabel, and Ngella islands (the chain north of Guadalcanal), connected by land bridges, was coined ‘Greater Bukida’ by Mayr & Diamond (2001). Guadalcanal, although being very close, might have remained isolated unless tectonic activity allowed the formation of a bridge with the Ngella Islands that were connected to Greater Bukida dur-

<table>
<thead>
<tr>
<th></th>
<th>A10 M1</th>
<th>A10 M2</th>
<th>A11 M1</th>
<th>A11 M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyura posidonius</td>
<td>1.00</td>
<td>1.00</td>
<td>0.90</td>
<td>0.98–1.00</td>
</tr>
<tr>
<td>Polyura sacco</td>
<td>0.38–0.39</td>
<td>0.85–0.86</td>
<td>0.44–0.45</td>
<td>0.98–1.00</td>
</tr>
<tr>
<td>Polyura caphontis</td>
<td>0.38–0.39</td>
<td>0.85–0.86</td>
<td>0.44–0.45</td>
<td>0.98–1.00</td>
</tr>
<tr>
<td>Polyura cognata</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Polyura dehanii</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99–1.00</td>
</tr>
<tr>
<td>Polyura epigenes</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Polyura epigenes*</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>ssp. bicolor</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

A10, BPP model with fixed guide tree; A11, BPP model with estimated guide tree; M1, BPP model with $G_{0}^{(1,10)}$ and $G_{10}^{(1,10)}$; M2, BPP model with $G_{0}^{(2,2000)}$ and $G_{2000}^{(2,2000)}$.  

### Table 3. Results of the BAYESIAN PHYLOGENETICS AND PHYLOGEOGRAPHY (BPP) analyses, with posterior probability intervals for each lineage
ing these periods. As a result, individuals in populations from the different islands making up Greater Bukida were probably able to interbreed and maintain gene flow, thereby preventing genetic segregation. On the other side of the archipelago, populations restricted to Malaita were likely to be isolated from other populations during most of the late Pleistocene, allowing their genetic and morphological differentiation. Several studies of the avifauna of the Solomon Islands have revealed that the spatial structure of genetic diversity in multiple species of birds is consistent with these periods of low eustasy (Filardi & Smith, 2005; Smith & Filardi, 2007; Uy, Moyle & Filardi, 2009). Likewise, Hagen, Donnellan & Bull (2012), studying the prehensile-tailed skink Corucia zebrata Gray, 1856, and Austin et al. (2010), investigating crocodile skinks (Squamata, Scincidae), found highly structured phylogeographic patterns with strong demarcation between groups of islands that were probably disconnected during periods of low sea level. Melonycteris fruit bats also display such a pattern, with populations on islands belonging to Greater Bukida being closely related and with other populations being clearly delineated genetically. In these molecular studies, the populations of Malaita and San Cristobal are always found as sister to populations on the remainder of the islands, illustrating the Plio-Pleistocene isolation fostered by sea-level rise and fall. The fact that P. epigenes epigenes is recovered nested within populations from Bougainville and Buka suggests that Guadalcanal might have been occasionally connected to the extreme southern part of Great Bukida, where the Ngella Islands lie today. Incomplete lineage sorting and hybridization are also impossible to rule out on the sole basis of this data set, and additional taxon sampling from Choiseul, New Georgia, and San Isabel will be necessary to tackle this question. The reason why Malaita and other island populations were unable to interbreed remains open to discussion, however. Here, we suggest that the behavior of Pacific Polyura butterflies might play a role. Males usually fly all day in a restricted area and chase other males, whereas females inhabit the forest canopy and only descend to feed or reproduce. Therefore, these butterflies do not appear to be good candidates for strong dispersal over water, and the widespread distribution of the genus in the Oriental region and Indo-Australian archipelago might be the result of a complex combination of occasional island hopping and rare long-distance dispersal.

CONCLUSION

The combination of a larger taxon sample enhanced the resolution of previously inferred phylogenetic relationships among populations of P. epigenes across the Solomon Islands archipelago. Complimentary analyses of molecular species delimitation typically agree that populations of P. epigenes from Malaita are an independent metapopulation lineage. These results are corroborated by analyses of the nuclear data only, which also reveal clear genetic differentiation. Our divergence time estimates suggest that populations of P. epigenes were likely to be isolated on this island during the Pleistocene high sea level, thereby promoting allopatric speciation. Based on these results and on the morphological divergence of populations from Malaita with respect to the two other subspecies of P. epigenes, we raised the former taxon to species status: P. bicolor Turlin & Sato, 1995 stat. nov.

ACKNOWLEDGEMENTS

We are indebted to Chris Müller for sending specimens and commenting on the manuscript. We also wish to thank John Tennent for helpful discussion. Three anonymous reviewers are acknowledged for comments that greatly helped to improve a former version of this paper. This work was supported by German Science Foundation (DFG) grants BA2152/6–1, 7–1, 11–1, and 20–1.

REFERENCES


Reid NM, Carstens BC. 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evolutionary Biology 12: 196.


Skale A, Tanzler R, Hendrich L, Balke M. 2012. High mitochondrial DNA sequence divergence in New Guinea Carabodytes stream beetles and the taxonomist’s dilemma when other evidence is kind of subtle…and collecting localities are far away away). ZooKeys 247: 31–43.


