Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 31 (2004) 1181–1196
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Tempo of speciation in a butterfly genus from the Southeast Asian tropics, inferred from mitochondrial and nuclear DNA sequence data

Hendrik-Jan Megens,[a,b,*] Coline H.M. van Moorsel,[b] William H. Piel,[b] Naomi E. Pierce,[c] and Rienk de Jong[a]

[a] *National Museum of Natural History, Leiden, The Netherlands*
[b] *Institute of Evolutionary and Ecological Sciences, Leiden University, Leiden, The Netherlands*
[c] *Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA*

Received 14 July 2003; revised 2 October 2003

Abstract

Molecular systematics is frequently beset with phylogenetic results that are not fully resolved. Researchers either state that the absence of resolution is due to character conflict, explosive speciation, or some combination of the two, but seldom do they carefully examine their data to distinguish between these causes. In this study, we exhaustively analyze a set of nuclear and mitochondrial nucleotide data for the Asian tropical butterfly genus *Arhopala* so as to highlight the causes of polytomies in the phylogenetic trees, and, as a result, to infer important biological events in the history of this genus. We began by using non-parametric statistical methods to determine whether the ambiguously resolved regions in these trees represent hard or soft polytomies. In addition we determined how this correlated to number of inferred changes on branches, using parametric maximum likelihood estimations. Based on congruent patterns in both mitochondrial and nuclear DNA sequences, we concluded that at two stages in the history of *Arhopala* there have been accelerated instances of speciation. One event, at the base of the phylogeny, generated many of the groups and subgroups currently recognized in this genus, while a later event generated another major clade consisting of both Oriental and Papuan species groups. Based on comparisons of closely related taxa, the ratio of instantaneous rate of evolution between mitochondrial and nuclear DNA evolution is established at approximately 3:1. The earliest radiation is dated between 7 and 11 Ma by a molecular clock analysis, setting the events generating much of the diversity of *Arhopala* at well before the Pleistocene. Periodical flooding of the Sunda plateau during interglacial periods was, therefore, not responsible for generating the major divisions in the genus *Arhopala*. Instead, we hypothesize that large-scale climatic changes taking place in the Miocene have induced the early acceleration in speciation.

Keywords: Rapid radiation; Hard polytomy; Molecular clock; Tropical diversity; *Arhopala*

1. Introduction

Phylogenetic analysis is increasingly used to address research questions in tropical ecology (Moritz et al., 2000; Wollenberg et al., 1996). With the growing wealth of DNA sequence data at hand, molecular phylogenies of extant taxa offer the opportunity to examine the tempo and mode of speciation (Pagel, 1998; Purvis et al., 1995). These analyses can enhance our understanding of the evolution in groups for which fossil data are lacking (Paradis, 1998), as is the case for most tropical rainforest taxa (Moritz et al., 2000). In addition, quantification of DNA substitutions under the assumptions of a molecular clock allows estimation of divergence dates (Rambaut and Bromham, 1998). Estimation of approximate branching times can help correlate novel traits or radiation events to geological events (Bromham et al., 1999; Shank et al., 1999; Xiang et al., 2000), and allow testing of ecological hypotheses (Pellmyr et al., 1998).

* Corresponding author. Present address: Cell Biology and Immunology, Department of Animal Sciences, Wageningen University, P.O. Box 338, 6700AH Wageningen, The Netherlands. Fax: +31-317-783955.
E-mail address: hendrik-jan.megens@wur.nl (H.-J. Megens).

1055-7903/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2003.10.002

Despite careful phylogenetic analysis using multiple gene regions and thousands of base pairs, the historical relationships of many groups of organisms remain unresolved (e.g., Arnaiz-Villena et al., 1999; Lessa and Cook, 1998; Waits et al., 1999). These enigmatic cases have led researchers to postulate rapid radiation events, where the lack of characters supporting an internal node is explained by insufficient time for substitutions to occur (Mardulyn and Whitfield, 1999; Waits et al., 1999). Polytomies of this type represent one or more internal branches at which no substitutions have accumulated. These branches therefore have near zero length, represent a factual quality of phylogenetic history, and are referred to as hard polytomies (Maddison, 1989). In contrast, soft polytomies occur when phylogenetic signal is obscured by homoplasy caused by multiple substitutions, insufficient taxon sampling, directional mutation, or other causes of character conflict.

In a number of recent studies (Halanych and Robinson, 1999; Jackman et al., 1999; Mardulyn and Whitfield, 1999; Richardson et al., 2001) rapid radiation was tested using molecular phylogenetic analyses. These investigations usually sought to use relative dating of branching events: (1) to show that internal nodes occurred in rapid succession and (2) to investigate the nature of the polytomies (hard or soft; Jackman et al., 1999; but also see Mardulyn and Whitfield, 1999; Waits et al., 1999). The two approaches can also be combined in order to demonstrate a correlation, as branch lengths and amount of phylogenetic information are influenced by the same process (e.g., rapid diversification: Halanych and Robinson, 1999).

1.1. The tropical butterfly genus Arhopala

The butterfly genus *Arhopala* (Lepidoptera: Lycaenidae) is among the most species rich butterfly genera in the Sunda region. With over 120 species present (Corbet and Pendlebury, 1992; Eliot, 1963, 1972; Seki et al., 1991), the genus *Arhopala* makes up almost 10% of all the butterfly species in this region. In total over 200 species have been described (Bridges, 1988), and many subspecies have been proposed to accommodate geographical variation (Evans, 1957). *Arhopala* is largely restricted to the Oriental and Australasian regions and the vast majority of species occur only in perhumid and seasonal tropical forests (Corbet, 1946). Therefore, this genus provides an excellent example of high entomological diversity in a tropical ecosystem.

Two explicit evolutionary explanations for why *Arhopala* contains so many species have been postulated. Corbet (1946) speculated that periodical flooding of the Malayan region during interglacials caused isolation of populations, and subsequent allopatric speciation. This protocol is generally thought to play an important role in tropical diversity generation (e.g., Bush, 1994). In contrast, Maschwitz et al. (1984) hypothesized that host plant co-speciation powered the diversification of at least one of the groups of species within the *Arhopala amphimuta* subgroup. Host plant co-evolution represents an important force in generating insect diversity (e.g., Farrell, 1998).

Both molecular and morphological studies show difficulties in resolving the basal phylogenetic relationships in the genus *Arhopala*. The earliest investigations on *Arhopala* already showed the difficulty of establishing taxonomic subdivisions in this large genus (e.g., Doherty, 1889 as cited in Evans, 1957). In the most recent comprehensive revision, Eliot (1963) set out to make 'natural' species groups. He constituted a large number of groups, but decided that it was impossible to find morphological characters upon which to infer relationships between them. In a recent molecular taxonomic study, Megens et al. (in press) presented a hypothesis of the higher phylogeny of *Arhopala*. The phylogenetic analysis was carried out using over 2 kb of mitochondrial (Cytochrome Oxidase 1 and 2) and nuclear (fragment of the *wingless* gene) protein coding genes. A number of systematic relationships were inferred with confidence, confirming previous hypotheses as well as establishing new ideas about *Arhopala* systematics. However, the phylogeny did not resolve at certain nodes, creating a large internal basal polytomy.

Several authors (e.g., Corbet, 1946; Eliot, 1963) have hinted that the problems of resolving the taxonomy of *Arhopala* are linked with its mode of evolution. Here we investigated two hypotheses. The first is that rapid radiation events are responsible for obscuring the genealogical relationships in *Arhopala*. We assessed unresolved nodes for being either hard or soft polytomies by estimating the amount of unambiguous phylogenetic signal and reconstructing the amount of substitutional changes (estimation of branch lengths). We also determined congruence between nuclear and mitochondrial partitions. For several species multiple specimens were included to investigate the amount of divergence between specimens of the same species and between closely related species, because instantaneous rates of evolution are best estimated between such recently diverged lineages (Moriyama and Powell, 1997). From these data we estimated the difference between the rate of substitution in mtDNA and nuclear *wingless* gene. In addition we dated internal nodes based on genetic divergences between species, thus testing the second hypthesis, that Pleistocene climate changes coincided with the diversification of *Arhopala*.

2. Material and methods

2.1. Taxon sampling and data alignment

We sampled 48 specimens, representing 27 species of *Arhopala*, one species of *Flos* (*F. anniella*) which was

regarded as belonging to the ingroup (Corbet, 1941; Megens et al., in press), and two outgroup species (*Semanga superba* and *Surendra vivarna*). This sample covered almost all of the larger groups within *Arhopala* as recognized by Eliot (1963) and thus should provide an even representation of diversity throughout the entire genus. We sequenced and aligned parts of the mitochondrial genes Cytochrome Oxidase 1, Cytochrome Oxidase 2, and $tRNA^{Leu}$. These sequences correspond with *Drosophila yakuba* (Clary and Wolstenholme, 1985) mitochondrial genome sites 1777 through 2170, 2229 through 2712, 2804 through 3265, and 3292 trough 3769. Primer pairs were, respectively, Ron-m/Nancy, Tonya/Hobbes, George/Phyllis, and Strom/Eva-m (Brower, 1994b; Simon et al., 1994). Primers indicated with '-m' were slightly modified by Megens (Ron-m: GGA GCT CCT GAC ATA GCA TTC CC; Eva-m: ATT ACT TGC TTT CAG TCA TCT). We also sequenced and aligned a 393 bp fragment of the nuclear gene *wingless* using primer pair LepWG1 and LepWG2 (Brower and DeSalle, 1998), corresponding to sites 1381–1744 in *Drosophila melanogaster wg* gene (Rijsewijk et al., 1987). All sequences were obtained by direct sequencing of PCR products.

2.2. Phylogenetic inference and character support

Phylogenetic inference was carried out with unweighted Maximum Parsimony (MP) analysis in PAUP* 4.04a (Swofford, 2000) using the heuristic search option (1000 replicates, steepest descent). Bootstrap analyses (Felsenstein, 1988) were performed using informative characters only, and 1000 replicates. Maximum Parsimony was preferred over other methods to analyze mitochondrial and nuclear sequences together because it can be viewed as a model free method, in addition allowing for explicit testing of incongruence of partitions that may represent different processes or histories (cf. Reed and Sperling, 1999). CO1, CO2, and *wingless* datasets were tested for incongruence using the incongruence length difference (ILD) test (Farris et al., 1995), implemented in PAUP* as the 'partition homogeneity test,' with 500 replicates to generate the null distribution.

Maximum Likelihood (ML) methods were used for the purpose of evaluating topologies found in the combined (mitochondrial DNA and *wingless*) MP analysis (cf. Yang, 1996), using PAUP4.04a. Mitochondrial and nuclear partitions were analysed separately using different parameter estimations, because an explicit model of evolution is assumed for each (Yang, 1996). The $-\ln$ likelihoods were summed since partitions can be assumed to be independent (Yang, 1996). We implemented a General Time Reversible model (GTR; substitution rates estimated), with base frequencies estimated empirically, using a Γ-shape (α estimated; 4 rate categories with 7 partitions independently evaluated, 7 rate categories when nuclear and mitochondrial partitions independently evaluated), and invariable sites (I estimated) (parameters estimated by PAUP*). For testing significance of differences in likelihood between topologies, the Kishino and Hasegawa (1989) test could not be used due to its parametric nature (M. Sanderson, pers. com.). Instead, significance values were determined with the nonparametric Templeton test (Templeton, 1983), using the differences in likelihood per site. The Templeton test was carried out using the Wilcoxon sign rank test as implemented in the statistical program SPSS 8.0. The single tree resulting from this procedure allowed comparing mitochondrial and nuclear partitions for branch length estimates without the topology of the tree being a variable.

The mitochondrial and nuclear partitions were separately analyzed using MrBayes 2.01 (Huelsenbeck and Ronquist, 2001), for the purpose of generating an additional confidence value and as an alternative evaluation of congruence in phylogenetic signal among the partitions. We implemented a GTR model, with rates site-specific and sites corresponding to codon positions, and 50,000 replicates performed after the burn in. In evaluating the congruence among partitions, we added posterior probabilities for nodes. A combined score of >150 indicates that in both analyses a particular node has a larger than 50% (and hence the largest) posterior probability in both partitions.

2.3. Estimation of branch lengths, molecular clock, and dating

A ML approach for estimating branch lengths is often preferred over parsimony (e.g., Halanych and Robinson, 1999; Monteiro and Pierce, 2001), because the implementation of realistic models of substitutional change outperform parsimony when multiple substitutions are being estimated (Yang and Nielsen, 2000). In addition, parsimony methods are found to perform relatively poorly under biased evolutionary scenarios, such as inequality of evolutionary rates among branches (Huelsenbeck, 1995), or under unequal base frequencies (Eyre-Walker, 1998). One of seven MP trees was chosen for further analysis, because it had the best likelihood score. The topology of this tree was used to estimate branch lengths, implementing a GTR model (substitution parameters estimated) with invariable sites and Γ-shape (7 rate categories; all parameters estimated).

Constraining a ML analysis to a molecular clock (Rambaut and Bromham, 1998) invokes computational limitations regarding estimation of parameters in the GTR-Γ-I model. Therefore, parameters established under corresponding non-clock ML estimations were implemented in clock-constrained evaluations. The same parameter estimations were used when subsequent clock evaluations required pruning of taxa in order to comply with a clock-like scenario. The significance in the differences between clock and non-clock scenarios for trees

were calculated with a likelihood ratio test ($2\Delta \ln L$ tested with χ^2 distance; Felsenstein, 1988).

Although the nuclear and mitochondrial sequences did not prove to represent different process partitions in a phylogenetic analysis, we did not a priori assume this would be true as well for a ML evaluation. In addition, given that there was a difference between partitions in compositional bias, the branch length estimations were performed for mitochondrial and nuclear sequences separately (cf. Yang, 1996).

The difference in instantaneous substitution rates between *wingless* and the mitochondrial DNA was estimated on different levels of divergence as recommended by Moriyama and Powell (1997). We used intraspecific comparisons of *Arhopala achelous*, *Arhopala epimuta*, and *Arhopala atosia*, between the species pair *A. atosia* and *A. epimuta*, and between closely related species of the *amphimuta* subgroup and *centaurus* group. The species pair *A. atosia* and *A. epimuta* was seen as a 'special case' of closely related species. Taxonomically they are virtually indistinguishable apart from a single morphological character that has certain phylogenetic lability (tails at hindwing vein 2; it is easilly lost, but not easilly regained; see Megens et al., in press). There is evidence that the tailless species is a recent offshoot of the tailed (Megens, 2002).

2.4. Testing for hard polytomies

Testing for hard polytomies (Maddison, 1989) as proposed by Jackman et al. (1999) is performed on datasets where saturated characters are removed. To estimate potential loss of phylogenetic signal due to saturation we employed two analyses. Uncorrected pairwise distances were plotted against the uncorrected pairwise distance using only transversions, under the assumption that transversion rate will go up particularly in compositionally biased sequences (e.g., Gomez-Zurita et al., 2000). As a second measure, we studied how Tamure-Nei (TN93) corrected distances compare with uncorrected distances (Jackman et al., 1999) for each codon position separately for both *wingless* and Cytochrome Oxidase (Reed and Sperling, 1999). For this purpose, each of the partitions (tRNA and the first, second and third codon positions of the mitochondrial and nuclear sequences) were treated separately. TN93 was preferred over ML correction because it is nonparametric and would therefore be less susceptible to unreliable estimates for partitions with relatively small numbers of variable sites (i.e., first and second codon positions) (cf. Yang, 1996).

We subsequently selected only those parts of the genes that were qualified as not significantly saturated. These partitions were analyzed for phylogenetic information with respect to particular internal nodes (Jackman et al., 1999). Two nodes were selected that were assumed to be difficult to resolve due to rapid speciation. A single species from each of the clades originating from these putatively unresolvable nodes was selected (Hillis and Huelsenbeck, 1992). Nodes and selected taxa for evaluation are shown in Fig. 1. Based on this sample of taxa we subsequently evaluated 10^6 random trees to estimate skewness parameter g1 (Hillis and Huelsenbeck, 1992), and a randomization test at 1000 replicates, PTP method (Archie, 1989; Faith and Cranston, 1991; Lyons-Weiler and Hoelzer, 1999) under the MP criterion as implemented in PAUP* 4.0.

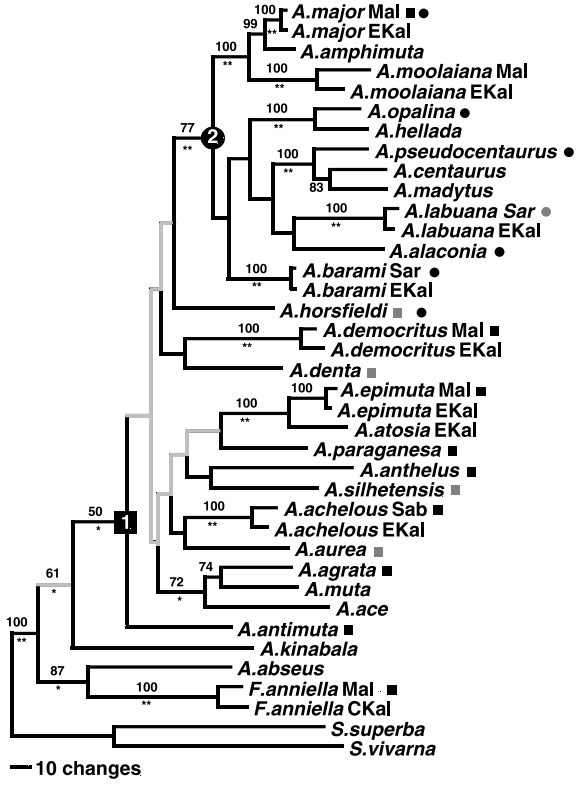

Fig. 1. One of seven shortest trees from the parsimony search using both mtDNA and *wingless* (treelength 2114 steps, CI = 0.42, RI = 0.50, RC = 0.21). This tree had the highest likelihood. Branches that collapsed in the strict consensus of the seven MP trees are in gray. MP bootstrap values are indicated above the branches; congruence among phylogenies derived from the Bayesian inference (combined posterior probabilities either by a *, indicating a > 150 combined score, or **, indicating a > 190 combined score) below the branches. We used this topology for further evaluations. The nodes designated with 1 and 2 were investigated for hard polytomies. The taxa labeled with a square or circle were used to investigate node 1 and 2, respectively, for hard polytomies, solid symbols used in the most restricted taxon evaluation, solid + gray in the more elaborate taxon sampling.

3. Results

3.1. Phylogenetic analysis

An aligned dataset was created consisting of 1778 bp of mtDNA and 393 bp of *wingless* from 27 species of

Arhopala (eight of which were represented by two specimens) and 1 species of *Flos*, and with *Su. vivarna* and *Se. superba* as outgroups. This resulted in a complete dataset of 38 specimens for the phylogenetic analysis, and sequences were complete for all these specimens. We further obtained sequence data for 10 specimens of *A. achelous*, *A. atosia*, and *A. epimuta*. However, since the CO1 fragment 1777 through 2170 failed to amplify for these specimens, they were excluded for the phylogenetic analysis, and included only in estimating instantaneous rates of evolution (see below). In addition, the *wingless* sequence for one *A. achelous* and mitochondrial fragment 2804 through 3265 for one *A. achelous* and one *A. epimuta* failed to amplify. Sequences are deposited in GenBank, accession numbers are shown in Appendix A. The aligned dataset is submitted to Treebase (www.treebase.org; accession number pending).

Seven bases were excluded for *Se. superba* at the junction of CO1 and $tRNA^{Leu}$ because these represent a unique insertion. A small portion of CO2 (corresponding to sites 3440 through 3470 of the *Drosophila yakuba* complete mtDNA sequence) was also discarded because it contained a highly variable indel-rich region. Despite the fact that the indels corresponded with complete codons, the alignment for this region could not be performed unambiguously. The ILD test results indicated that mtDNA and *wingless* ($p = 0.20$), and CO1, CO2, and *wingless* separately ($p = 0.18$) did not represent incongruent partitions. Subsequent unweighted parsimony analysis of the combined data set yielded 7 trees of equal length (2114 steps). Evaluation of these trees under a ML model with separate evaluations for mtDNA and *wingless* showed that one topology had the lowest $-\ln L$ score (12844.96); the tree is shown in Fig. 1. A Templeton (1983) test performed on likelihood per site showed no significant difference between this tree and the other 6 trees. The strict consensus of the 7 MP trees can be inferred from Fig. 1 by collapsing the hatched branches.

There was no major incongruence between the two Bayesian analyses, meaning that no group with higher than 50% posterior probability was present which was in conflict with the results from the other partition. Only one minor conflict occurred with regard to the species of the *agelastus* group: *A. alaconia* clusters with *A. labuana* (88% posterior probability) in the analysis based on the mtDNA, while in the analysis based on *wingless* it clusters with *A. barami* (98%). The presence of only this single minor incongruence indicates that there is no major conflict in phylogenetic information between the mitochondrial and nuclear genes.

3.2. Branch length estimates, molecular clock, and dating

The Maximum Likelihood estimation of branch lengths based on the overall most likely tree for the *wingless* and mtDNA data sets separately is shown in Fig. 2. There is a substantial separation between outgroup and ingroup species, more so in the *wingless* data than in the mtDNA data. In the ingroup both the *A. abseus*/*F. anniella* clade and the *A. kinabala* clade are well differentiated for both nuclear and mtDNA partitions. The branches subdividing other *Arhopala* groups generally have extremely short reconstructed branch lengths, effectively collapsing many nodes into a large polytomy (number 1 in Fig. 1). The congruence between nuclear and mtDNA partitions is striking, and the polytomy that would be created is congruent with the polytomy from the consensus of the 7 MP trees (Fig. 1, gray branches).

From this polytomy originates a well-established clade (number 2 in Fig. 1) which contains several larger groups that were recognized by Eliot (1963; *amphimuta* subgroup, the *agelastus*, *camdeo*, *wildei*, and *centaurus* groups). At the base of this clade branch lengths again are short, in particular with regard to the *wingless* data, indicating that this node may also effectively represent a polytomy. Congruence between mtDNA and *wingless* regarding the inferred branch lengths in the two polytomies is visualized by Fig. 3. All 12 branches (indicated by the open circles) which are effectively collapsed in the two putative polytomies consistently show near zero inferred branch lengths for both partitions.

ML evaluations of the tree for CO1, CO2 (including tRNA), and *wingless* separately under a clock model, resulted in significant differences between clock and non-clock scenarios (all $p \ll 0.01$). Base composition heterogeneity can seriously impair the assessment of molecular phylogenetic processes (Lyons-Weiler and Hoelzer, 1999). Therefore, we investigated the presence of compositional heterogeneity. The third codon position, which in the Cytochrome Oxidase is the most heavily biased position, shows significant compositional heterogeneity among ingroup taxa ($\chi^2 = 135.4$; df $= 105$; $p < 0.025$). CO1 is the most compositionally biased partition, and shows a much higher degree of compositional heterogeneity compared to CO2. To investigate the effect of variance in base composition on ML estimates, we plotted GC content against the observed (uncorrected) and ML-corrected distances from the root (represented by *Su. vivarna*) for CO1. We compared this with *wingless*, a data partition which shows very little compositional bias and no significant compositional heterogeneity. For the CO1 sequences there was a clear correlation especially when ML corrected distances were used (Fig. 4). By pruning six taxa with a deviant GC content for the CO1 sequences from the tree (taxa are indicated in Fig. 4A), compositional heterogeneity was no longer significant in the CO third codon positions combined ($\chi^2 = 78.0$; df $= 87$; $p = 0.74$).

After elimination of the effect of variance in GC content on the ML estimates by pruning taxa from the

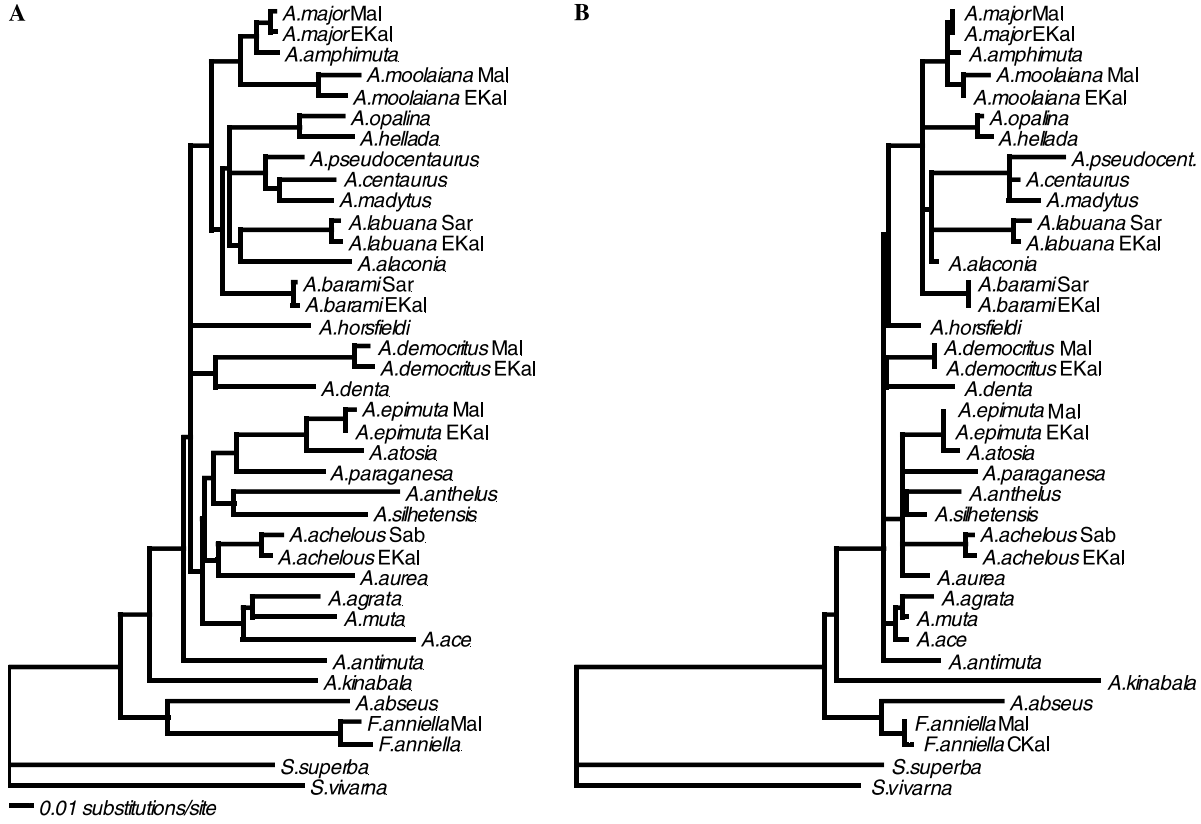

Fig. 2. Branch length reconstruction of the tree topology of Fig. 1. Cytochrome Oxidase (A) and *wingless* (B) were evaluated separately. The branch lengths of both reconstructions are drawn to the same scale.

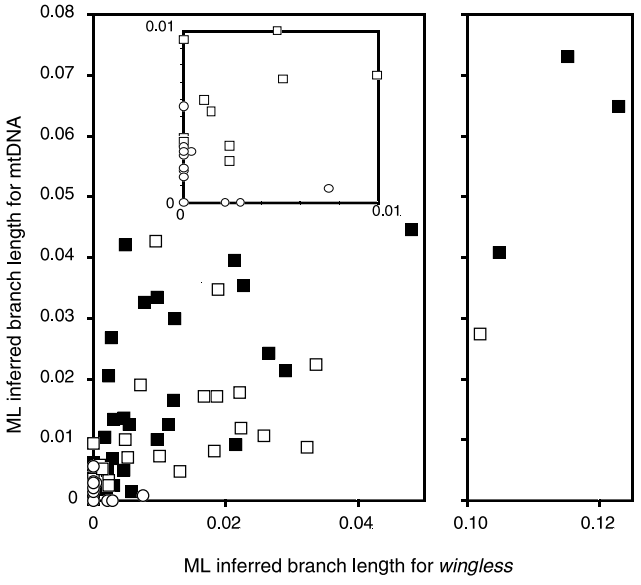

Fig. 3. The inferred proportional (branch length divided by summed branch lengths over the whole tree) branch lengths for *wingless* plotted versus mtDNA, based on the tree topology of Fig. 1. Open symbols (○, □) show internal branches, solid squares (■) the external branches. Open circles (○) indicate the 12 internal branches that are effectively collapsed in polytomy 1 and 2 (see Fig. 1). The inset shows an enlargement of the graph for smaller than 1% proportional branch length. Fig. 2 shows the corresponding phylograms.

tree, the clock hypothesis could not be rejected as significantly worse than the non-clock hypothesis at the 0.01 level ($2\Delta \ln L = 44.91$, df $= 28$, $p = 0.023$ for CO1; $2\Delta \ln L = 47.29$, df $= 28$, $p = 0.013$ for CO2). For the *wingless* sequences such a correlation could not be found (Fig. 4B, and pruning taxa with a deviating GC content did not result in a better fit to a clock model. We subsequently estimated branch lengths for the CO1 and CO2 data based on the pruned topology shown in Fig. 1, under the same ML model (parameters estimated), but now assuming a clock. The resulting estimates are shown in Fig. 5.

3.3. Testing for hard polytomies

The third codon position of the CO1 and CO2 genes shows a substantial increase in proportion of transversions at higher uncorrected pairwise distances (Fig. 6), which indicates that these positions are saturated (Arnaiz-Villena et al., 1999). The proportion of transversions in CO increases with increasing pairwise distance, and at the higher pairwise distances greatly exceeds that of the *wingless* gene.

Assessment of the corrected versus uncorrected genetic distances between pairs of species suggests no

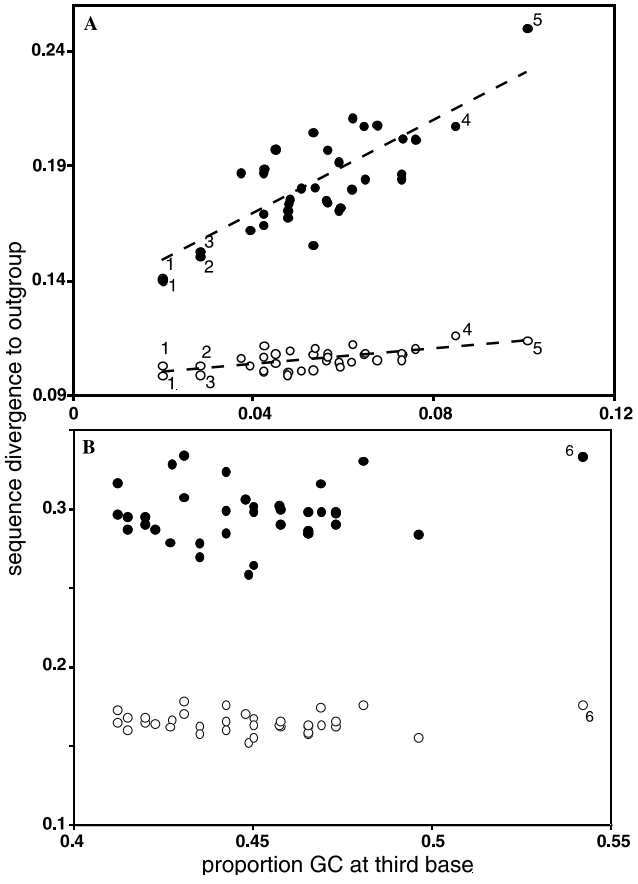

Fig. 4. Relationship between distance to root and proportion of GC at third codon position for CO1 (A) and *wingless* (B). The distance to the root is calculated as the pairwise ML corrected (●) and uncorrected (○) to *S. vivarna* for all ingroup taxa. For CO1, the outliers were: (1) *A. major* (two specimens); (2) *A. amphimuta*; (3) *A. madytus*; (4) *A. silhetensis*; and (5) *A. ace*. For *wingless* the most conspicuous outlier was *A. kinabala* (no. 6).

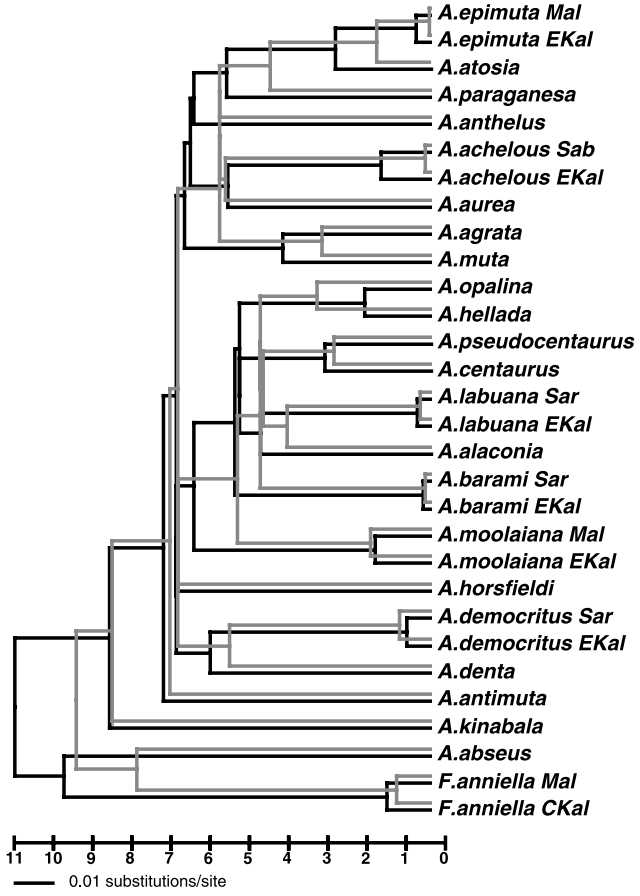

Fig. 5. A molecular clock tree for Cytochrome Oxidase 1 and 2 based on the topology of Fig. 1, for the selection of taxa for which a clock hypothesis could not be discarded. The black tree is from the CO1 estimations, the gray tree from CO2.

substantial saturation for first and second positions of *wingless*, for the first and second positions of the CO1 and CO2 genes, and for the tRNA (Fig. 7; tRNA not shown). The third codon position of *wingless* and the CO1 and CO2 genes shows higher levels of uncorrected pairwise divergence and a higher proportion of inferred additional mutations. The corrected divergence of the third codon positions of *wingless*, however, is never more than about 50% higher than the uncorrected. In addition, there was no increase in transversions over transitions, indicative of the absence of saturation. The CO third positions can have a corrected distance that is more than twice, and even going up to three times, the uncorrected distance. In addition, it is displaying a higher variance, suggesting that the correction is unreliable. The CO third codon positions were therefore not included in the hard polytomy test.

The g1 test results (Table 1) show insignificant signal for polytomy 1. Based on g1 results for polytomy 2, some phylogenetic information is suggested to be present in the *wingless* partition. However, the

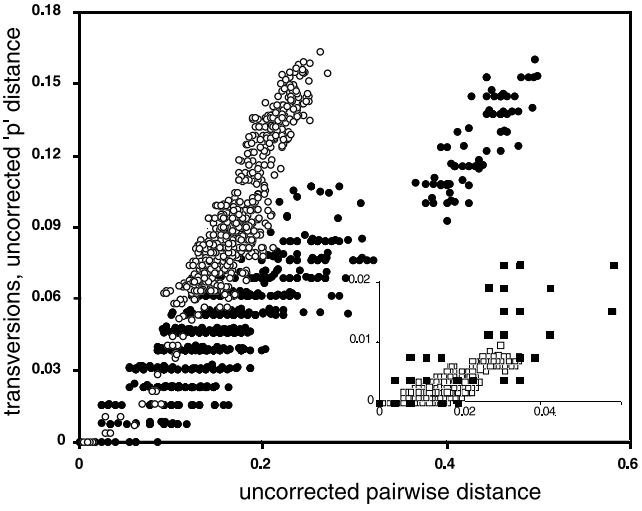

Fig. 6. A plot of the uncorrected pairwise distance versus the pairwise distances when only transversions were taken. Large graph depicts the third codon positions of *wingless* (●) and Cytochrome Oxidase 1 and 2 (○). The small inset graph shows the first + second positions plot for *wingless* (■) and CO1 and 2 (□).

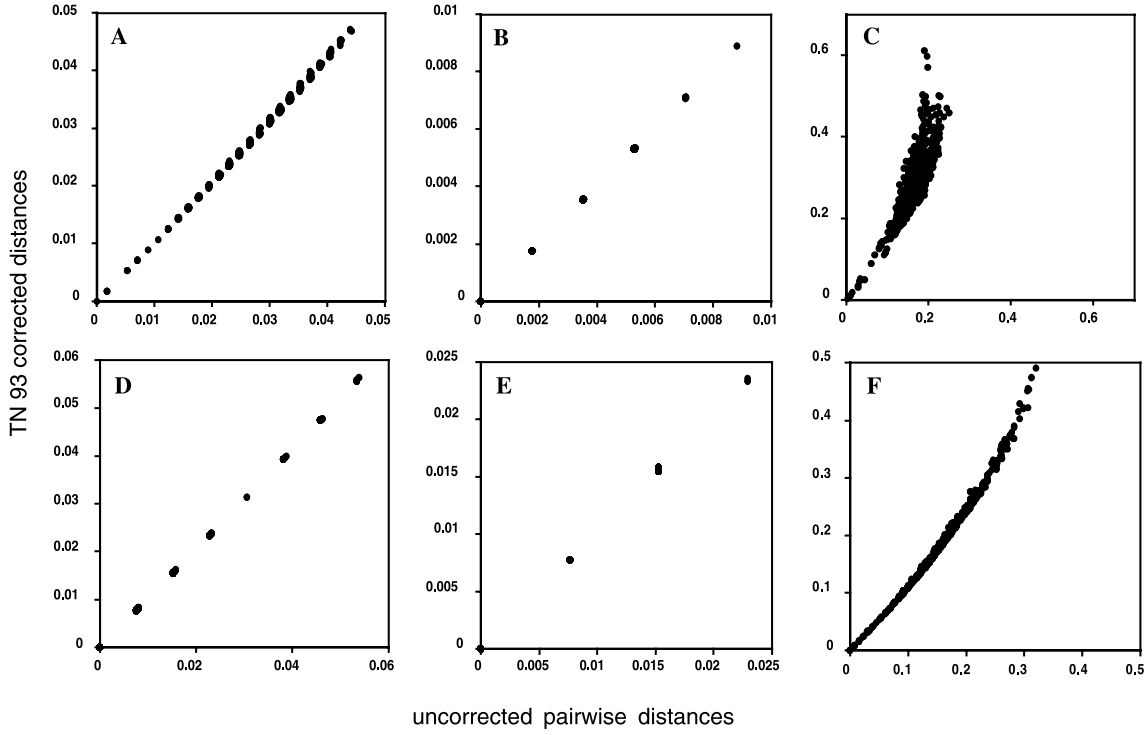

Fig. 7. Plots showing uncorrected pairwise distances in relation to pairwise TN93 reconstructed distances between ingroup taxa, for Cytochrome Oxidase 1 and 2 first codon positions (A), second codon positions (B) and third codon positions (C), and *wingless* first codon position (D) second codon positions (E) and third codon position (F).

Table 1
Results of the 'hard polytomy test'

	Polytomy 1						Polytomy 2					
	8 taxa ■			12 taxa ■ ■			5 taxa ●			6 taxa ● ●		
	PTP	g1	Var. sites	PTP	g1	Var. sites	PTP	g1	Var. sites	PTP	g1	Var. sites
CO, first and second	0.93	−0.12	53(15)	0.86	−0.1	70(26)	0.17	−0.48	32(7)	0.36	−0.22	37(10)
wg	0.47	−0.16	52(20)	0.11	−0.26	62(25)	0.47	−1.18*	28(3)	0.07	−0.74*	40(7)
wg + CO, first and second	0.16	−0.25	105(35)	0.50	−0.24	132(51)	0.66	−0.07	60(10)	0.05	−0.79**	77(17)

The gray and black squares and circles refer to the taxon sampling depicted in Fig. 1. A * indicates significance values <0.05, ** values <0.01. CO, cytochrome oxidase; first and second refers to codon positions of CO; wg, *wingless*. Numbers between brackets indicate number of parsimony informative sites.

corresponding low number of parsimony informative sites in comparison to the actual number of variable sites (3 and 28, respectively) may lead to inaccurate interpretation of the table in Hillis and Huelsenbeck (1992). A PTP test on the same two nodes reveals that the actual length consistently falls within a randomly distributed dataset (results also in Table 1), indicating that at both polytomies there is no significant phylogenetic signal. Inclusion of *A. labuana* at polytomy 2 seems to enhance phylogenetic signal in the *wingless* partition as is suggested by the PTP test. Based on CO1 and CO2 first and second positions (including the small portion of tRNA) and on the *wingless* gene, we conclude that polytomy 1 is 'hard' and thus unresolvable using the data tested. Polytomy 2 also seems to represent unresolvable relationships, but the test results are not fully in agreement here.

3.4. Estimating instantaneous rates of evolution

The estimated intraspecific sequence divergence between mitochondrial haplotypes sampled here ranged from as little as 0.2% (between *A. barami* from Sarawak and East Kalimantan) to about 3% (between *A. moolaiana* from Kalimantan and the Malaysian peninsula; see Fig. 5). *A. atosia* and *A. epimuta* were treated

Table 2
Average uncorrected pairwise sequence divergence of mtDNA and *wingless*, including SE, for increasing levels of taxonomic diversity

	N	div mtDNA	stdev	div wg	stdev	Ratio
Same locality	19	0.0060	(0.0103)	0.0031	(0.0037)	1.94
Same species, different locality	15	0.0069	(0.0071)	0.0028	(0.0033)	2.45
Very close species: *epimuta–atosia*	18	0.0238	(0.0077)	0.0077	(0.0018)	3.10
Other putative close species	9	0.0324	(0.0122)	0.0170	(0.0079)	1.90
All ingroup	630	0.0618	(0.0135)	0.0533	(0.0242)	1.16
Between out- and ingroup	51	0.0937	(0.0051)	0.1590	(0.0091)	0.59
Mean of first, second, and third categories	52	0.0124	(0.0119)	0.0046	(0.0037)	2.71

The last row shows an average of among populations, within species, and between *A. atosia* and *A. epimuta* estimations (meaning the first three rows). The ratio mtDNA/*wingless* is given in the last column. *N*, number of comparisons; div, average pairwise divergence; stdev, standard deviation.

seperately from other 'closely related species' of the *centaurus* group and *amphimuta* subgroup (see Section 4). The average inferred instantaneous substitution rate for mtDNA versus *wingless* between very closely related specimens is estimated at approximately 2.7 (Table 2). There is an apparent difference between specimens of the same locality, geographically remote specimens, and the *A. atosia/epimuta* species pair, which in this order show an ascending inferred ratio between mtDNA and *wingless*, peaking at 3.10 for the *A. atosia/epimuta* species pair. A decline in the ratios is observed when comparing among less closely related species, among ingroup species, and between in- and outgroup species.

4. Discussion

There has been much discussion about 'the best evolutionary rate' for a gene to be suitable for phylogeny reconstruction (e.g., Yang, 1998). However, usefulness of particular DNA sequences also depends on tempo and mode of genealogical differentiation of the organisms carrying them. This in fact has led phylogeneticists to use the reverse argument, and to discern particular events in the history (other than the branching order) from the pattern of diversification (e.g., Pagel, 1998). The genes we have used for this study have been used before in other studies on butterfly phylogenetics. The use of mitochondrial genes in particular has been problematic in several cases since it is almost universally AT biased in insects (Wirth et al., 1999). Despite this problem, the Cytochrome Oxidases have proven to be valuable for inferring butterfly phylogenies (Brower, 1994b; Brower and Egan, 1997; Caterino and Sperling, 1999; Monteiro and Pierce, 2001; Rand et al., 2000). The same is true for the sequenced portion of the *wingless* gene, even though it is relatively short. It has proven to be useful at several levels of phylogenetic depth, both at the genus and tribus level (Brower and DeSalle, 1998) and at the subfamily and family level (Campbell et al., 2000). There seems to be no a priori reason to assume that this gene evolves with insufficient speed to resolve the phylogeny of *Arhopala*. The fact that very species rich butterfly groups can yield well resolved, stable, and congruent phylogenies using these genes indicates the need for a thorough investigation into the nature of the polytomies that are observed in this study.

4.1. Congruent results in branch length estimation and polytomy tests

The 'hard polytomy' test incorporates those parts of the dataset that are not much affected by homoplasious events. The non-parametric nature of generating a null-distribution and subsequent MP evaluation prohibits the use of substantially saturated partitions (Jackman et al., 1999). Therefore, we had to remove the third codon positions of Cytochrome Oxidase for this part of our analysis. There is some debate about the validity of parsimony analysis and non-parametric tests in biased sequence data (e.g., Eyre-Walker, 1998). However, Lyons-Weiler and Hoelzer (1999) showed that generally the amount of phylogenetic signal is overestimated in such biased sequences. As our results from the PTP test show an absence of phylogenetic signal, it is unlikely that biases cause problems in this study.

Parametric methods, such as Maximum Likelihood incorporating realistic models of substitutions, are found to perform best regarding reconstructing multiple changes (Yang, 1996; Yang and Nielsen, 2000). In this study we invoke a GTR-Γ-I model, which is currently among the most complex models implemented in phylogenetic analysis. Optimizing corresponding parameters of the model is advantageous since it takes care of the difficult task of weighting evolutionary pathways and correcting for multiple hits without the need for ad hoc approximations (Yang and Nielsen, 2000). Parameter estimation, however, is crucial for inferring correct probabilities of base change, and ML analysis may show sensitivity to heavily biased sequences (Lyons-Weiler and Hoelzer, 1999; but see Yang and Nielsen, 2000).

Our analyses indicate that absence of phylogenetic signal coincides with limited inferred mutational

changes on internal branches. These findings are consistent with earlier studies claiming rapid radiation events (Halanych and Robinson, 1999). Moreover, these patterns are strikingly congruent between *wingless* and Cytochrome Oxidase (Fig. 3). These sequences, originating from two separately evolving genomes, not only generate congruent phylogenetic signal (as indicated by the Bayesian analysis, Fig. 1), they also share the absence of clear phylogenetic signal for particular internal nodes. Well-supported nodes occur both below and above the polytomies. The most probable explanation is that the molecular evolution is reflected in the (species) evolution of this genus. On at least two occasions speciation proceeded so fast that given the data presented here there can be no accurate assessment of phylogenetic relationships, neither for the *wingless* gene, nor for the Cytochrome Oxidase.

4.2. Interpretation of the molecular clock

Evaluation of the sequence data indicates that the evolution of both mitochondrial and nuclear DNA has not proceeded in a clocklike manner in all taxa. In the mtDNA there is a clear relationship between GC content (at third bases), and inferred length to the root. As a consequence, pruning the most extreme data results in a better fit to a molecular clock scenario. The overall estimations of divergence are slightly higher for CO1 than for CO2 in the ML tree, but there is complete congruence for the two putative radiation events. This slightly higher divergence for CO1 contrasts with *Bicyclus*, where CO2 consistently gave higher estimates of divergence (Monteiro and Pierce, 2001).

It is accepted practice to estimate divergence events based on a fossil-calibrated molecular clock (Rambaut and Bromham, 1998). Unfortunately, for insects such calibrations are often not feasible, and hence we need to resort to comparative studies. This comparative dating is a common practice for groups for which fossil data is absent (e.g., Arnaiz-Villena et al., 1999; Halanych and Robinson, 1999; Shank et al., 1999). A widely accepted rate for mitochondrial cytochrome oxidase genes in insects, based on comparisons between very closely related species, is 2% per million years pairwise divergence (Brower, 1994a; DeSalle et al., 1987). This figure has been used in numerous other studies of insect molecular divergence (e.g., Juan et al., 1995, 1996; Monteiro and Pierce, 2001; Sandoval et al., 1998). Based on this calibration, the first major division within *Arhopala s.l.* (including *Flos*; Corbet, 1941) may have occurred as early as 11 million years ago. An estimate for the first putative radiation event taken place in *Arhopala* would be 7 million years ago, dating it prior to the Pliocene.

Substitution rates of mitochondrial genes are considered to be 10 times higher than nuclear genes in mammals (Moriyama and Powell, 1997). In insects, however, the ratio between mitochondrial and nuclear genes is presumed to be much smaller due to differences in metabolic rate. Furthermore, severe AT-bias makes it difficult to estimate differences in rates of synonymous substitution. Even when comparing fairly closely related species, AT bias may cause saturation in the mtDNA (Moriyama and Powell, 1997), leading to underestimation of the true level of divergence. Yet, by comparing populations, there is the possibility of detecting allelic divergence much older than the populations due to ancestral polymorphisms (Moriyama and Powell, 1997). Such polymorphisms are more likely to remain in nuclear genes compared to the mitochondrial haplotype since the latter effective population size is four times smaller than the former (in case of uniparental inheritance; Avise, 1994).

Ideally, one should compare mitochondrial and nuclear genes for different degrees of divergence (Moriyama and Powell, 1997). This approach is taken here. The results of Table 2 are in striking agreement with the predicted pattern; from very closely related specimens to distantly related species, the ratio goes up first, peaking at 3.10 for the extremely closely related species pair *A. atosia* an *A. epimuta*. This species pair is taken as a special case of 'closely related species,' since upon more detailed analysis the variation of *A. epimuta wingless* and mtDNA sequences seems to fall within that of *A. atosia*, suggesting that the former is a recent, yet morphologically well defined, offshoot of the latter (analysis not shown). From the *A. atosia/epimuta* species pair the ratio goes down again. This suggests that even within groups or subgroups a certain level of multiple substitutions for the mtDNA has to be assumed. This also confirms that *A. atosia* and *A. epimuta* are more closely related than the species of the *centaurus* and *amphimuta* (sub)groups. According to Moriyama and Powell (1997), the highest ratio in such a sequence of comparisons from closely to less closely related is nearest to the correct one. Therefore, the ratio of 3.10 seems the most accurate estimation in our analysis.

The ratio becomes even smaller than 1 when comparing distantly related taxa, estimating a larger uncorrected sequence divergence for *wingless* compared to CO (Table 2). Also, for the ML corrected sequence divergence, the differences in estimation of divergence between the mitochondrial and nuclear DNA become larger as more distantly related species are compared, as can be seen in Fig. 4. At the earliest radiation (polytomy 1) there is an inferred sequence divergence of about 14% for CO (Fig. 5), corresponding to an age of about 7 Ma. However, using the same procedure for

wingless, the ML corrected sequence difference at this particular node is about 7%. This corresponds to approximately 11 Ma, assuming that the instantaneous rate of evolution is three times faster for CO. Going even further back in time, the origin of the genus would be estimated at more than double the sequence divergence for *wingless* (approximately 15%) and only about 50% more for CO.

The extreme compositional bias is likely to be of importance in this regard, as the number of substitutions as inferred by ML analysis seems to be correlated to GC content at third base (see Fig. 4A. As the underestimation seems to be more substantial for mitochondrial DNA deeper into the tree, this indicates that ML correction may not be able to infer the actual number of changes that occurred. The level of divergence as shown in Fig. 5 based on the mitochondrial genes therefore represent an underestimate, and thus the assessment of 7 Ma for the initial radiation in *Arhopala* is likely to be an underestimate as well.

4.3. Patterns of diversification in the phylogeny of *Arhopala*

Our analyses support the hypothesis that one or several rapid radiations occurred during the evolution of the tropical genus *Arhopala*. This observation is at odds with the hypothesis that diversity in the tropics results from low extinction rates in combination with gradual partitioning of species over long periods of time in a stable environment. Over the last years several studies have indicated that rapid radiations do occur in tropical rainforest habitats (Blattner et al., 2001; Garcia-Moreno et al., 1999; Richardson et al., 2001). Periodical flooding of the Sunda plateau during Pleistocene inter-glaciations was not responsible for massive species formation early in the history of the genus *Arhopala*. This theory, postulated by Corbet (1946), does not correlate well with the inferred date of at least 7 Ma for the early radiation. Most of the species groups appear to have been present already at the onset of the Pleistocene.

Rapid diversification early in *Arhopala* evolution appears so suddenly that it must be explained by one or several dramatic changes. Such a change could be the acquisition of a novel life history trait, followed by adaptive radiation. The second radiation (node 2 in Fig. 1) coincides well with a change in life history traits. The overall trend in this clade is towards greater myrmecophily and away from feeding chiefly on Fagaceae (Megens, 2002). Such a correlation is, however, not present at the earlier radiation. Feeding on oaks is the dominant preference in most clades apart from those present in node 2, but seems to have originated prior to the radiation in node 1.

Alternatively, climatic or environmental changes can induce massive speciation (Janis et al., 2000; Morley, 2000; Wilf et al., 2001). The molecular clock analysis suggests that the initial radiations in *Arhopala* may have occurred at least 7 Ma, which would put this event somewhere near the end of the Miocene. Yet, it is likely that the CO based molecular clock underestimates the true amount of divergence, particularly when inferring early events (see above in the discussion). Comparisons with the *wingless* sequence suggest this underestimation could be larger than 50%, which would push the initial radiation further back still. From the middle Miocene onward, the SE Asian forests were in a substantial state of flux due to dramatic climatic changes in the form of global cooling by 15 Ma (Zachos et al., 2001) and resulting in intensified monsoons around 8 Ma (Zachos et al., 2001). During this interval, the dipterocarp lowland rainforests, until recently covering the whole of the Sunda Plateau, came into existence (Morley, 2000). Such events can trigger rapid and large scale evolution in animals depending on host plants, as has been demonstrated by changes in diversity of herbivorous animals during other Tertiary climatic changes (Janis et al., 2000; Wilf et al., 2001). Given the coincidence of a basal rapid radiation in *Arhopala* generating almost all the major groups in the genus, so close in time with major climate and floristic changes, we must consider this event as a possible cause for the inferred radiation. Recently Slik (2001) and Blattner et al., 2001) suggested that the speciose euphorb genus *Macaranga* originated in the middle Miocene. This genus, which is now closely associated with the lowland dipterocarp forests of SE Asia, seems to be derived from savanna species of the genus *Mallotus* that were probably more abundant in dryer Early Miocene times. A comparative, phylogenetic analysis on other species rich groups of animals and plants occurring in the region may reveal that the diversification of *Arhopala* has been part of a massive environmental change acting upon all forms of life in this region.

Acknowledgments

We thank Johan van Nes for technical assistance, Man-Wah Tan collected many specimens and assisted in identifications. Danny Cleary contributed specimens and together with Ferry Slik assisted in the field work. Barbara Gravendeel provided useful comments to the manuscript. The research was funded through NWO Priority Program Biodiversity in Disturbed Ecosystems (Grant 895-10-002P) and further supported through grants from NWO (Grant SIR14-2218), Jan Joost ter Pelkwijk Fonds en Uyttenboogaart-Eliasen Stichting (all to the first author).

Appendix A

Species and species group	Collector and sample code	Collection date and locality	Accession numbers for mtDNA (1777–2170, 2229–2712, 2804–3265, and 3292–3769) and *wingless*
Arhopala			
abseus group			
A. abseus	MWT 93-E051	VII-1993 Gn. Serapi, Malaysia	AY235896, AY235992, AY235848, AY235940, AY236044
agelastus group			
A. labuana	NP 95-Z298	IX-1995 Sarawak, Malaysia	AY235872, AY235966, AY235823, AY235914, AY236018
	HJM 97240522	V-1997 Berau, Kaltim, Indonesia	AY235873, AY235967, AY235824, AY235915, AY236019
A. alaconia	HJM 97280512	V-1997 Berau, Kaltim, Indonesia	AY235876, AY235970, AY235827, AY235918, AY236022
A. barami	NP 95-Z318	IX-1995 Sarawak, Malaysia	AY235874, AY235968, AY235825, AY235916, AY236020
	HJM 9725054	V-1997 Berau, Kaltim, Indonesia	AY235875, AY235969, AY235826, AY235917, AY236021
agesias group			
A. kinabala	NP 95-Y263	IX-1995 Pahang, Malaysia	AY235895, AY235991, AY235847, AY235939, AY236043
alitaeus group			
A. denta	HJM 9728051	V-1997 Berau, Kaltim, Indonesia	AY235885, AY235979, AY235836, AY235927, AY236031
amphimuta group, *amphimuta* subgroup			
A. major	MWT 93-B055	VII-1993 FRIM, Malaysia	AY235861, AY235955, AY235812, AY235904, AY236007
	HJM 9729057	V-1997 Berau, Kaltim, Indonesia	AY235862, AY235956, AY235813, AY235905, AY236008
A. amphimuta	MWT 93-C034	VII-1993 FRIM, Malaysia	AY235863, AY235957, AY235814, AY235906, AY236009
A. moolaiana	NP 95-Y246	IX-1995 Pahang, Malaysia	AY235864, AY235958, AY235815, AY235907, AY236010
	HJM 97280518	V-1997 Berau, Kaltim, Indonesia	AY235865, AY235959, AY235816, AY235908, AY236011
amphimuta group, *muta* subgroup			
A. muta	NP 95-Z297	IX-1995 Sarawak, Malaysia	AY235889, AY235984, AY235841, AY235932, AY236036
amphimuta group, *perimuta* subgroup			
A. antimuta	HJM 9729053	V-1997 Berau, Kaltim, Indonesia	AY235894, AY235989, AY235845, AY235937, AY236041
anthelus group			
A. achelous	MWT 93-D017	VII-1993 Kokol, Sabah, Malaysia	AY235880, AY235974, AY235831, AY235922, AY236026
	HJM 97270513	V-1997 Berau, Kaltim, Indonesia	—, AY235998, AY235854, AY235946, AY236050
	HJM 97270514	V-1997 Berau, Kaltim, Indonesia	AY235881, AY235975, AY235832, AY235923, AY236027
	HJM 97270517	V-1997 Berau, Kaltim, Indonesia	—, AY235999, AY235855, AY235947,—
	HJM 97240521	V-1997 Berau, Kaltim, Indonesia	—, AY235997, AY235853, AY235945, AY236049

Appendix A (*continued*)

Species and species group	Collector and sample code	Collection date and locality	Accession numbers for mtDNA (1777–2170, 2229–2712, 2804–3265, and 3292–3769) and *wingless*
	HJM 97240528	V-1997 Berau, Kaltim, Indonesia	—, AY236000, —, AY235948, AY236051
A. anthelus	NP 95-Z316	IX-1995 Sarawak, Malaysia	AY235890, AY235985, AY235842, AY235933, AY236037
aurea group			
A. aurea	DC 97733	1997 Sangai, Kalten, Indonesia	AY235892, AY235987, AY235844, AY235935, AY236039
camdeo group			
A. opalina	MWT 93-C071	VII-1993 Genting, Malaysia	AY235867, AY235961, AY235818, AY235909, AY236013
A. hellada	MWT 93-E050	VII-1993 Gn. Serapi, Malaysia	AY235868, AY235962, AY235819, AY235910, AY236014
centaurus group			
A. centaurus	KD 94-T070	VII-1994 Queensland, Australia	AY235871, AY235965, AY235822, AY235913, AY236017
A. madytus	KD 95-Z559	XI-1995 Queensland, Australia	AY235869, AY235963, AY235820, AY235911, AY236015
A. pseudocentaurus	NP 95-Z291	IX-1995 Sarawak, Malaysia	AY235870, AY235964, AY235821, AY235912, AY236016
cleander group			
A. ace	NP 95-Y261	IX-1995 Pahang, Malaysia	AY235888, AY235983, AY235840, AY235931, AY236035
A. agrata	MWT 93-E068	VII-1993 Gn. Serapi, Malaysia	AY235887, AY235982, AY235839, AY235930, AY236034
A. silhetensis	HJM 97240523	V-1997 Berau, Kaltim, Indonesia	AY235891, AY235986, AY235843, AY235934, AY236038
democritus group			
A. democritus	MWT 93-E075	VII-1993 Gn. Serapi, Malaysia	AY235883, AY235977, AY235834, AY235925, AY236029
	HJM 97240526	V-1997 Berau, Kaltim, Indonesia	AY235884, AY235978, AY235835, AY235926, AY236030
epimuta group			
A. epimuta	MWT 93-C014	VII-1993 Papah, Malaysia	AY235877, AY235971, AY235828, AY235919, AY236023
	HJM 9727055	V-1997 Berau, Kaltim, Indonesia	—, AY236004, AY235859, AY235952, AY236055
	HJM 9727058	V-1997 Berau, Kaltim, Indonesia	AY235878, AY235972, AY235829, AY235920, AY236024
	HJM 97270510	V-1997 Berau, Kaltim, Indonesia	—, AY236003, AY235858, AY235951, AY236054
	HJM 97270521	V-1997 Berau, Kaltim, Indonesia	—, AY236005, —, AY235953, AY236056
	HJM 97270524	V-1997 Berau, Kaltim, Indonesia	—, AY236006, AY235860, AY235954, AY236057
A. atosia	HJM 97270515	V-1997 Berau, Kaltim, Indonesia	—, AY236001, AY235856, AY235949, AY236052
	HJM 97250517	V-1997 Berau, Kaltim, Indonesia	—, AY236002, AY235857, AY235950, AY236053
	HJM 97270523	V-1997 Berau, Kaltim, Indonesia	AY235879, AY235973, AY235830, AY235921, AY236025

Appendix A (*continued*)

Species and species group	Collector and sample code	Collection date and locality	Accession numbers for mtDNA (1777–2170, 2229–2712, 2804–3265, and 3292–3769) and *wingless*
eumolphus group			
A. horsfieldi	NP 95-Y249	IX-1995 Pahang, Malaysia	AY235886, AY235981, AY235838, AY235929, AY236033
ganesa group			
A. paraganesa	MWT 93-D026	VII-1993 Kokol, Sabah, Malaysia	AY235882, AY235976, AY235833, AY235924, AY236028
Flos anniella	DC 977714	1997 Sangai, Kalten, Indonesia	AY235898, AY235994, AY235850, AY235942, AY236046
	MWT 93-C048	VII-1993 Genting Ridge, Malaysia	AY235897, AY235993, AY235849, AY235941, AY236045
Semanga superba	MWT 93-C070	VII-1993 Genting, Malaysia	AY235899, AY235995, AY235851, AY235943, AY236047
Surendra vivarna	MWT 93-B049	VII-1993 Tapah, Malaysia	AY235900, AY235996, AY235852, AY235944, AY236048

References

Archie, J.W., 1989. A randomization test for phylogenetic information in systematic data. Syst. Zool. 38, 239–252.

Arnaiz-Villena, A., Alvarez-Tejado, M., Ruiz-del-Valle, V., Garcia-de-la-Torre, C., Varela, P., Recio, M.J., Ferre, S., Martinez-Laso, J., 1999. Rapid radiation of canaries (genus *Serinus*). Mol. Biol. Evol. 16, 2–11.

Avise, J.C., 1994. Molecular Markers, Natural History, and Evolution. Chapman and Hall, London. 512 pp.

Blattner, F.R., Weising, K., Bänfer, G., Maschwitz, U., Fiala, B., 2001. Molecular analysis of phylogenetic relationships among myrmecophytic *Macaranga* species (Euphorbiaceae). Mol. Phylogenet. Evol. 19, 331–344.

Bridges, C.A., 1988. Catalogue of Lycaenidae and Riodinidae (Lepidoptera: Rhopalocera). Published by C.A. Bridges.

Bromham, L., Phillips, M.J., Penny, D., 1999. Growing up with dinosaurs: molecular dates and the mammalian radiation. Trends Ecol. Evol. 14, 113–117.

Brower, A.V.Z., 1994a. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. USA 91, 6491–6495.

Brower, A.V.Z., 1994b. Phylogeny of *Heliconius* butterflies inferred from Mitochondrial DNA sequences (Lepidoptera: Nymphalidae). Mol. Phylogenet. Evol. 3, 159–174.

Brower, A.V.Z., Egan, M.G., 1997. Cladistic analysis of Heliconius butterflies and relatives (Nymphalidae: Heliconiiti): a revised phylogenetic position for *Eueides* based on sequences from mtDNA and a nuclear gene. Proc. Roy. Soc. Lond. B 264, 969–977.

Brower, A.V.Z., DeSalle, R., 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of *wingless* as a source of characters for phylogenetic inference. Insect Mol. Biol. 7, 73–82.

Bush, M., 1994. Amazonian speciation: a necessarily complex model. J. Biogeogr. 21, 5–17.

Campbell, D.L., Brower, A.V.Z., Pierce, N.E., 2000. Molecular evolution of the Wingless gene and its implications for the phylogenetic placement of the butterfly family Riodinidae (Lepidoptera: Papilionoidae). Mol. Biol. Evol. 17, 684–696.

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.

Clary, D.O., Wolstenholme, D.R., 1985. The mitochondrial DNA molecular of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22, 252–271.

Corbet, S.A., 1941. A key to the Indo-Malayan species of *Arhopala* Boisduval (Lepidoptera: Lycaenidae). Proc. R. Entomol. Soc. B 10, 149–170.

Corbet, S.A., 1946. Observations on the Indo-Malayan species of the genus *Arhopala* Boisduval (Lepidoptera: Lycaenidae). Trans. R. Entomol. Soc. Lond. 96 (part 5), 73–88.

Corbet, S.A., Pendlebury, H.M., 1992. The Butterflies of the Malay Peninsula, fourth revised ed. Malayan Nature Society, Kuala Lumpur.

DeSalle, R., Freedman, T., Prager, E.M., Wilson, A.C., 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol. 26, 157–164.

Eliot, J.N., 1963. A key to the Malayan species of *Arhopala*. Malayan Nat. J. 17, 188–217.

Eliot, J.N., 1972. Some *Arhopala* from Borneo, with a revision of the *Arhopala cleander* group. (Lepidoptera: Lycaenidae). J. Nat. Hist. 6, 1–15.

Evans, W.H., 1957. A revision of the *Arhopala* group of Oriental Lycaenidae (Lepidoptera: Rhopalocera). Bull. Br. Mus. Nat. Hist. (Ent.) 5, 85–141.

Eyre-Walker, A., 1998. Problems with parsimony in sequences of biased base composition. J. Mol. Evol. 47, 686–690.

Faith, D.P., Cranston, P.S., 1991. Could a cladogram this short have arisen by chance alone?: on permutation tests for cladistic structure. Cladistics 7, 1–28.

Farrell, B.D., 1998. Inordinate fondness explained: Why are there so many beetles? Science 281, 555–559.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1995. Testing significance of incongruence. Cladistics 10, 315–319.

Felsenstein, J., 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22, 521–565.

Garcia-Moreno, J., Arctander, P., Fjeldså, J., 1999. A case of rapid Diversification in the Neotropics: phylogenetic relationships

among *Cranioleuca* Spinetails (Aves, Furnariidae). Mol. Phylogenet. Evol. 12, 273–281.

Gomez-Zurita, J., Juan, C., Petitpierre, E., 2000. The evolutionary history of the genus *Timarcha* (Coleoptera, Chrysomelidae) inferred from mitochondrial COII gene and partial 16S rDNA sequences. Mol. Phylogenet. Evol. 14, 304–317.

Halanych, K.M., Robinson, T.J., 1999. Multiple substitutions affect the phylogenetic utility of Cytochrome *b* and 12S rDNA data: examining a rapid radiation in leporid (Lagomorpha) evolution. J. Mol. Evol. 48, 369–379.

Hillis, D.M., Huelsenbeck, J.P., 1992. Signal, noise and reliability in molecular phylogenetic analysis. J. Hered. 83, 189–195.

Huelsenbeck, J.P., 1995. Performance of phylogenetic methods in simulations. Syst. Biol. 44, 17–48.

Huelsenbeck, J.P. and F. Ronquist 2001. MrBayes: A Program for the Bayesian Inference of Phylogeny. Department of Biology, University of Rochester.

Jackman, T.R., Larson, A., de Queiroz, K., Losos, J.B., 1999. Phylogenetic relationships and tempo of early diversification in *Anolis* lizards. Syst. Biol. 48, 254–285.

Janis, C.M., Damuth, J., Theodor, J.M., 2000. Miocene ungulates and terrestrial primary productivity: where have all the browsers gone? Proc. Natl. Acad. Sci. USA 97, 7899–7904.

Juan, C., Oromi, P., Hewitt, G.M., 1995. Mitochondrial DNA phylogeny and sequential colonization of Canary Islands by darkling beetles of the genus *Pimelia* (Tenebrionidae). Proc. R. Soc. Lond. B 261, 173–180.

Juan, C., Oromi, P., Hewitt, G.M., 1996. Phylogeny of the genus *Hegeter* (Tenebrionidae, Coleoptera) and its colonization of the Canary Islands deduced from Cytochrome Oxidase I mitochondrial sequence. Heredity 76, 392–403.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidae. J. Mol. Evol. 29, 170–179.

Lessa, E.P., Cook, J.A., 1998. The molecular phylogenetics of Tucotucos (genus *Ctenomys*, Rodentia: Octodontidae) suggests an early burst of speciation. Mol. Phylogenet. Evol. 9, 88–99.

Lyons-Weiler, J., Hoelzer, G.A., 1999. Null model selection, compostional bias, character state bias, and the limits of phylogenetic information. Mol. Biol. Evol. 16, 1400–1405.

Maddison, W.P., 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5, 365–377.

Mardulyn, P., Whitfield, J.B., 1999. Phylogenetic signal in the CO1, 16S, and 28S genes for inferring relationships among genera of Microgastrinae (Hymenoptera; Braconidae): evidence of a high diversification rate in this group of parasitoids. Mol. Phylogenet. Evol. 12, 282–294.

Maschwitz, U., Schroth, M., Hänel, H., Pong, T.Y., 1984. Lycaenids parasitizing symbiotic plant–ant partnerships. Oecologia 64, 78–80.

Megens, H.-J., van Nes, W.J., van Moorsel, C.H.M., Pierce, N.E., de Jong, R., in press. Molecular phylogeny of the Oriental butterfly genus *Arhopala* (Lycaenidae, Theclinae) inferred from mitochondrial and nuclear genes. Syst. Entomol.

Megens, H.-J., 2002. Understanding the diversity of the tropical butterfly genus *Arhopala*: a molecular phylogenetic approach. PhD thesis Leiden University, Leiden, The Netherlands.

Monteiro, A., Pierce, N.E., 2001. Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII and EF-1α gene sequences. Mol. Phylogenet. Evol. 18, 264–281.

Moritz, C., Patton, J.L., Schneider, C.J., Smith, T.B., 2000. Diversification of Rainforest Faunas: an integrated molecular approach. Annu. Rev. Ecol. Syst. 31, 533–563.

Moriyama, E.N., Powell, J.R., 1997. Synonymous substitution rates in *Drosophila*: mitochondrial versus nuclear genes. J. Mol. Evol. 45, 378–391.

Morley, R.J., 2000. Origin and Evolution of Tropical Rain Forests. John Wiley and Sons.

Pagel, M., 1998. Inferring evolutionary processes from phylogenies. Zool. Scr. 26, 331–348.

Paradis, E., 1998. Detecting shifts in diversification rates without fossils. Am. Nat. 152 (2), 176–187.

Pellmyr, O., Leebens-Mack, J., Thompson, J.N., 1998. Herbivores and molecular clocks as tools in plant biogeography. Biol. J. Linn. Soc. 63, 367–378.

Purvis, A., Nee, S., Harvey, P.H., 1995. Macroevolutionary inferences from primate phylogeny. Proc. R. Soc. Lond. B 260, 329–333.

Rambaut, A., Bromham, L., 1998. Estimating divergence dates from molecular sequences. Mol. Biol. Evol. 15, 442–448.

Rand, D.B., Heath, A., Suderman, T., Pierce, N.E., 2000. Phylogeny and life history evolution of the genus *Chrysoritis* within the Aphnaeini (Lepidoptera: Lycaenidae), inferred from mitochondrial Cytochrome Oxidase I sequences. Mol. Phylogenet. Evol. 17, 85–96.

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.

Richardson, J.E., Pennington, R.T., Pennington, T.D., Hollingsworth, P.M., 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. Science 293, 2242–2245.

Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D., Nusse, R., 1987. The *Drosophila* homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. Cell 50, 649–657.

Sandoval, C., Carmean, D.A., Crespi, B.J., 1998. Molecular Phylogenetics of sexual and parthenogenetic *Temema* walking sticks. Proc. R. Soc. Lond. B 265, 589–595.

Seki, Y., Takanami, Y., Otsuka, K., 1991. Butterflies of Borneo, vol. 2: Lycaenidae. Tobishima corp. 112 pp.

Shank, T.M., Black, M.B., Halanych, K.M., Lutz, R.A., Vrijenhoek, R.C., 1999. Miocene radiation of deep-sea hydrothermal vent shrimp (Caridae: Bresiliidae): evidence from mitochondrial Cytochrome Oxidase subunit 1. Mol. Phylogenet. Evol. 13, 244–254.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–701.

Slik, J.W.F., 2001. *Macaranga* and *Mallotus* (Euphorbiaceae) as indicators for disturbance in the lowland dipterocarp forests of East Kalimantan, Indonesia. Tropenbos—Kalimantan Series 4, Tropenbos International, Wageningen, The Netherlands, pp. 224.

Swofford, D.L., 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.04a. Sinauer Associates, Sunderland, Massachusetts.

Templeton, A.R., 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and apes. Evolution 37, 221–244.

Waits, L.P., Sullivan, J., O'Brian, S.J., Ward, R.H., 1999. Rapid radiation events in the family Ursidae indicated by likelihood phylogenetic estimation from multiple fragments of mtDNA. Mol. Phylogenet. Evol. 13, 82–92.

Wilf, P., Labandeira, C.C., Johnson, K.R., Coley, P.D., Cutter, A.D., 2001. Insect herbivory, plant defense, and early Cenozoic climate change. Proc. Natl. Acad. Sci. USA 98, 6221–6226.

Wirth, T., Le Guellec, R., Veuille, M., 1999. Directional substitution and evolution of nucleotide content in the Cytochrome Oxidase II gene in Earwigs (Dermapteran Insects). Mol. Biol. Evol. 16, 1645–1653.

Wollenberg, K., Arnold, J., Avise, J.C., 1996. Recognizing the forest for the trees: testing temporal patterns of cladogenesis using a

null model of stochastic diversification. Mol. Biol. Evol. 13, 833–849.

Xiang, Q.-Y., Soltis, D.E., Soltis, P.S., Manchester, S.R., Crawford, D.J., 2000. Timing the Eastern Asian–Eastern North American Floristic disjunction: molecular clock corroborates paleontological estimates. Mol. Phylogenet. Evol. 15, 462–472.

Yang, Z.H., 1996. Maximum Likelihood Models for combined analyses of multiple sequence data. J. Mol. Evol. 42, 587–596.

Yang, Z.H., 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. 47, 125–133.

Yang, Z.H., Nielsen, R., 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. Mol. Biol. Evol. 17, 32–43.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693.