

**Title: Diversity dynamics in Nymphalidae butterflies: Effect of phylogenetic uncertainty on diversification rate shift estimates**

**Authors: Carlos Peña<sup>1,\*</sup> and Marianne Espeland<sup>2,\*</sup>**

**Running title: Diversity dynamics in Nymphalidae butterflies**

*1 Laboratory of Genetics, Department of Biology, University of Turku, Turku, Finland. Email: [mycalesis@gmail.com](mailto:mycalesis@gmail.com)*

*2 Museum of Comparative Zoology and Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, USA. Email: [marianne.espeland@gmail.com](mailto:marianne.espeland@gmail.com), McGuire Center?*

**Corresponding author: [mycalesis@gmail.com](mailto:mycalesis@gmail.com)** Telephone: +358 417063065 Fax: +358 2 333 6680

**Comments:** 5814 words, 7 figures, 2 tables and 25 supplementary material files.

**Additional Supplementary materials:** <http://dx.doi.org/10.6084/m9.figshare.639208>

## **Abstract**

The speciose butterfly family Nymphalidae has been used to study evolutionary interactions between plants and insects. Theories of insect-hostplant dynamics predict accelerated diversification due to key innovations. We investigated whether phylogenetic uncertainty affects a commonly used method (MEDUSA, modelling evolutionary diversity using stepwise AIC) for estimating shifts in diversification rates in lineages, by extending the method across a random sample of trees from the posterior distribution of a Bayesian run. We found that phylogenetic uncertainty greatly affects diversification rate estimates. Different trees produced diversification rates ranging from high values to almost zero for the same clade, and both significant rate increase and decrease in some clades. Only four out of 18 significant shifts found on the maximum clade credibility tree were consistent across most of the sampled trees. Among these, we found accelerated diversification for Ithomiini butterflies. We used the binary speciation and extinction model (BiSSE) and found that a hostplant shift to Solanaceae is correlated with increased net diversification rates in Ithomiini, congruent with the diffuse cospeciation hypothesis. Our results show that taking phylogenetic uncertainty into account when estimating net diversification rate shifts is of great importance, and relying on the maximum clade credibility tree alone can potentially give erroneous results.

**Keywords:** diversification analysis, MEDUSA, BiSSE, speciation rate, insect-hostplant dynamics

## 29 Introduction

30 Hostplant shifts have been invoked to be responsible for a great part of the biodiversity of herbivorous  
31 insects (Mitter *et al.*, 1988). The study of the evolution of hostplant use has spawned several theories  
32 explaining the evolutionary interactions between plants and insects (reviewed by Nyman *et al.*, 2012).  
33 The “escape-and-radiate” hypothesis (Ehrlich & Raven, 1964), the “oscillation hypothesis” (Janz, 2011;  
34 Nylin *et al.*, 2013) or “diffuse cospeciation” (Nyman *et al.*, 2012).

35 The butterfly family Nymphalidae has been an important taxon for developing some of the mentioned  
36 hypotheses. Nymphalidae contains around 6000 species (Nieukerken *et al.*, 2011), and is the largest family  
37 within the true butterflies. The family most likely originated around 94 MYA in the mid Cretaceous.  
38 Diversification of the group began in the Late Cretaceous and most major radiations (current subfamilies)  
39 appeared shortly after the Cretaceous-Paleogene (K-Pg) boundary (Heikkilä *et al.*, 2012). Several studies  
40 have used time-calibrated phylogenies and diversification models to reconstruct the evolutionary history  
41 of the group to identify patterns of accelerated or decelerated diversification of some Nymphalidae clades  
42 (Elias *et al.*, 2009; Wahlberg *et al.*, 2009; Fordyce, 2010; Heikkilä *et al.*, 2012). For example, it has been  
43 suggested that climate change in the Oligocene and the subsequent diversification of grasses has led to  
44 diversification of the subfamily Satyrinae (Peña & Wahlberg, 2008) due to the abundance of grasses over  
45 extensive geographic areas (“resource abundance-dependent diversity dynamics” hypothesis). Fordyce  
46 (2010) found increased net diversification rates in some Nymphalidae lineages after a major hostplant  
47 shift, which appears to be in agreement with Ehrlich & Raven (1964) “escape-and-radiate” model of  
48 diversification.

49 Although it has been suggested that part of the great diversity of Nymphalidae butterflies is a result  
50 of hostplant-insect dynamics, it is necessary to use modern techniques to investigate whether the di-  
51 versification patterns of Nymphalidae are in agreement with the theoretical predictions. It is necessary  
52 to test whether the overall diversification pattern of Nymphalidae is congruent with events of sudden  
53 diversification bursts due to hostplant shift, climatic events or shifts to closely related hostplants (Nylin  
54 *et al.*, 2013; Ferrer-Paris *et al.*, 2013).

55 In this study, we used a time-calibrated genus-level phylogenetic hypothesis for Nymphalidae butterflies  
56 (taken from Wahlberg *et al.*, 2009) to investigate patterns of diversification. We applied MEDUSA  
57 (modelling evolutionary diversification using stepwise AIC, Alfaro *et al.*, 2009; Harmon *et al.*, 2011), a  
58 recently developed statistical method, to study the diversification pattern of Nymphalidae butterflies.  
59 MEDUSA fits likelihood models of diversification into a time-calibrated tree and tests whether allowing  
60 increases or decreases in speciation and extinction rates within the tree produces better fit of the models.  
61 MEDUSA is able to take into account unsampled extant species diversity during model fitting and  
62 it is normally applied to the maximum clade credibility phylogenetic tree. Particularly, we wanted

to study the effects of phylogenetic uncertainty and by using the extended MEDUSA method called MultiMEDUSA (version 093.4.3 Alfaro *et al.*, 2009). We also tested whether hostplant association dynamics can explain the diversification patterns of component Nymphalidae lineages by testing whether character states of hostplant use affected the diversification pattern of those lineages employing the method BiSSE as implemented in the R package `diversitree` (FitzJohn, 2012).

## Methods

### Data

For analyses, we used the phylogenetic trees from the study of Wahlberg *et al.* (2009) that were generated using DNA sequence data from 10 gene regions for 398 of the 540 valid genera in Nymphalidae. We employed Wahlberg *et al.* (2009) maximum clade credibility tree (MCC tree) (Fig. 1) as well as a random sample of 1000 trees from their BEAST run after burnin. Their original BEAST run was for 40 million generations. We used a burnin of 25 million generations and took a random sample of 1000 trees using Burntrees v.0.1.9 <http://www.abc.se/~nylander/> (supp. mat. 5) in order to correct for phylogenetic uncertainty when performing the diversification analyses.

Species richness data for Nymphalidae genera were compiled from several sources including the specialist-curated lists on <http://tolweb.org>, Lamas (2004) and curated lists of Global Butterfly Names project (<http://www.ucl.ac.uk/taxome/gbn/>). We assigned the species numbers of genera not included in the phylogeny to the closest related genus that was included in Wahlberg *et al.* (2009) study according to available phylogenetic studies (Penz, 1999; Freitas & Brown Jr., 2004; Peña *et al.*, 2006, 2010, 2011; Silva-Brandão *et al.*, 2008; Brower *et al.*, 2010; Kodandaramaiah *et al.*, 2010; De-Silva *et al.*, 2010; Matos-Maraví *et al.*, 2013; Ortiz-Acevedo & Willmott, 2013), taxonomical classification and morphological resemblance when no phylogenies were available (supp. mat. 3).

Hostplant data for Nymphalidae species were compiled from several sources including Ackery (1988), HOSTS database (<http://bit.ly/YI7nwW>), Dyer & Gentry (2002), Beccaloni *et al.* (2008), Janzen & Hallwachs (2009) and others (supp. mat. 15–16) for a total of 6586 hostplant records, including 428 Nymphalidae genera and 143 plant families and 1070 plant genera. It was not possible to find any hostplant data for 35 butterfly genera.

### Analyses of Diversification

We used the statistical software R version 3.0.1 (Core Team, 2013) in combination with the APE (Popescu *et al.*, 2012), GEIGER (Harmon *et al.*, 2008), MEDUSA (Alfaro *et al.*, 2009) and `diversitree` (FitzJohn, 2012) packages along with our own scripts to perform the analyses (included as supplementary materials

94 2, 6, 10). All analyses were run on the 1000 random trees from Wahlberg et al. (2009), on the maximum  
95 clade credibility tree derived from these and the MCC tree from Wahlberg et al. (2009).

## 96 **Detecting diversification shifts on phylogenetic trees**

97 Patterns of diversification in Nymphalidae were analyzed by using MEDUSA version 093.4.33 (Alfaro *et*  
98 *al.*, 2009) on the maximum clade credibility tree from Wahlberg et al. (2009). MEDUSA fits alternative  
99 birth-death likelihood models to a phylogenetic tree in order to estimate changes in net diversification  
100 rates along branches. MEDUSA estimates likelihood and AIC scores for the simplest birth-death model,  
101 with two parameters, the rates  $\mathbf{r}$ : net diversification = speciation ( $\mathbf{b}$ ) - extinction ( $\mathbf{d}$ ) and  $\varepsilon$ : relative  
102 extinction =  $\mathbf{d}/\mathbf{b}$ . The AIC scores of the two-parameter model are then compared with incrementally  
103 more complex models until the addition of parameters do not improve the AIC scores beyond a cutoff  
104 value. MEDUSA finds the likelihood of the models after taking into account branch lengths and number  
105 of species per lineage (Alfaro *et al.*, 2009). Most studies using MEDUSA only run the method on a single  
106 tree, usually the maximum clade credibility tree, which makes the assumption that this tree is correct. We  
107 wanted to study the effects of phylogenetic uncertainty on estimation of net diversification rate shifts and  
108 therefore selected 1000 random genus-level trees from the posterior distribution (MultiMEDUSA, supp.  
109 mat. 5). We calculated a new MCC tree from the selection of trees, run MEDUSA and MultiMEDUSA  
110 on the selection of 1000 trees and summarized the estimated changes in net diversification rates for nodes  
111 across all trees. Patterns of change in net diversification rates are significant if they are found at the  
112 same node in at least 90% of the trees. We also expected to find similar values of net diversification ( $\mathbf{r} =$   
113  $\mathbf{b} - \mathbf{d}$ ) and relative extinction rate ( $\varepsilon = \mathbf{d}/\mathbf{b}$ ) values across the 1000 trees for the nodes where changes in  
114 diversification tempo occurs. We used the AICc threshold of 7.8 units, as estimated by MEDUSA, as the  
115 limit for a significantly better fit to select among increasingly complex alternative models. We did not  
116 set a higher limit in the number of net diversification rate shifts that could be estimated by MEDUSA  
117 as we used the corrected AICc threshold recommended by the software.

## 118 **Estimation of trait-dependent speciation rates**

119 As the MEDUSA and MultiMEDUSA approaches estimated an increase in net diversification in the clade  
120 Ithomiini, we tested whether this pattern can be explained by increase in birth-rate due to hostplant  
121 use and performed analyses with the “binary state speciation and extinction” (BiSSE; Maddison *et al.*,  
122 2007) Bayesian approach as implemented in the R package *diversitree* (FitzJohn, 2012). MuSSE  
123 (FitzJohn, 2012) is designed to examine the joint effects of two or more traits on speciation. Because  
124 most of Nymphalidae butterflies are restricted to use one plant family as hostplant, the character states  
125 can be coded as presence/absence, for which the BiSSE analysis is better suited. BiSSE was designed to

test whether a binary character state has had any effect on increased net diversification rate for a clade (Maddison *et al.*, 2007). We used our compiled data of hostplant use to produce a binary dataset for the character “feeding on the plant family Solanaceae” with two states (absence = 0; presence = 1) (supp. mat. 17), which is the main hostplants of the diverse Ithomiini butterflies and closest relatives (Willmott & Freitas, 2006). Other hostplant shifts were not tested for effect on net diversification rates due to the low support for diversification shifts found in the MEDUSA analyses. We analyzed the data using BiSSE employing the Markov Chain Monte Carlo algorithm on the maximum clade credibility tree, taking into account missing taxa by using the parameter “sampling factor” (`sampling.f`) in `diversitree`. We also used constrained analyses forcing no effect of hostplant use on diversification and likelihood ratio tests to find out whether the hypothesis of effect on diversification has a significantly better likelihood than the null hypothesis (no effect). The analyses were run across a sample of 250 trees from the posterior distribution. The records of *Vanessa* and *Hypanartia* feeding on Solanaceae (Scott, 1986; Dyer & Gentry, 2002) might be incorrect as it is unlikely that these species can be feeding on this plant family. We tested whether coding these two genera as not feeding on Solanaceae affected our results.

## Results

### Detecting diversification shifts on the maximum clade credibility tree

The MEDUSA analysis on the MCC tree in combination with richness data estimated 18 changes in the tempo of diversification in Nymphalidae history (Fig. 1; Table 1). The corrected AICc acceptance threshold for adding subsequent piecewise birth-death processes to the overall model was set to 7.8 units, as prescribed by MEDUSA. In all MEDUSA analyses, the maximum number of inferred diversification splits in all trees was 26. The background net diversification rate for Nymphalidae was estimated as  $r = 0.092$  lineages per Million of years and the AICc score for the best fit model was 5090.5 (Table 1). MEDUSA also estimated that the basic constant birth–death model was not a better explanation for our data (AICc = 5449.3).

Some of the 18 changes in diversification correspond to rate increases in very species-rich genera: *Ypthima* ( $r = 0.311$ ), *Charaxes* ( $r = 0.291$ ), *Callicore* + *Diaethria* ( $r = 0.220$ ), *Pedaliodes* ( $r = 0.196$ ) and *Taenaris* ( $r = 0.234$ ). We found rate increases for other clades as well such as: Mycalesina ( $r = 0.191$ ), Oleriina + Ithomiina + Napeogenina + Dircennina + Godyridina ( $r = 0.187$ ), Satyrini ( $r = 0.116$ ), Phyciodina in part ( $r = 0.241$ ) and Satyrina ( $r = 0.221$ ), *Coenonympha* ( $r = 0.209$ ), *Caeruleptychia* + *Magneptychia* ( $r = 0.312$ ) and *Taenaris* ( $r = 0.312$ ). We also found decreases in net diversification rates for Limenitidinae + Heliconiinae ( $r = 0.0541$ ), part of Danaini ( $r = 0.0423$ ), Pseudergolinae ( $r = 0.024$ ) and Coenonymphina ( $r = 0.065$ ) (Table 1).

## 158 Phylogenetic uncertainty in the MultiMEDUSA approach

159 We tested MEDUSA to find out whether taking into account the phylogenetic signal from a random  
160 sample of 1000 trees from the posterior distribution can return similar estimates of diversification to the  
161 values obtained from the MCC tree. We run MultiMEDUSA on a random sample of 1000 trees (supp.  
162 mat. 5) from the posterior distribution and compared the results with a MEDUSA analysis on the MCC  
163 tree derived from this sample (supp. mat. 9).

164 We found that the analysis by MultiMEDUSA on the 1000 trees estimated lower median net diversifica-  
165 tion rates for the diversification shifts found by MEDUSA on the MCC tree from Wahlberg et al. (2009)  
166 (Table 2). Although the diversification pattern found by MEDUSA and MultiMEDUSA was the same,  
167 the latter consistently estimated lower rates. Thus, the shifts recovered with low net diversification rate  
168 on the MCC were recovered with negative net diversification rate by MultiMEDUSA. The background  
169 diversification and all shifts found by MEDUSA on the 1000 trees are provided as an R object in supp.  
170 mat. 7).

171 We also compared the results from MultiMEDUSA (derived from the sample of 1000 trees) with the  
172 splits found by MEDUSA on the MCC tree derived from this random sample. In the summary statistics,  
173 MultiMEDUSA reports the frequency of the diversification shifts found in the trees (parameter `sum.prop`).  
174 Thus, if a node is found in only half of the 1000 trees, but the phylogenetic signal was strong enough to  
175 be picked up by MEDUSA and a node shift was found most of the time, then the `sum.prop` should be  
176 close to 1. For example the node 625 consists of the clade **Charaxes** + **Polyura** and it was found in only  
177 256 trees, however MultiMEDUSA was consistently able to find a diversification shift for that node and  
178 the `sum.prop` value is 0.996.

179 We found different rates for diversification shifts even for the nodes that appear in most of the sample  
180 of 1000 trees (Table 2) as many of the nodes also found in the MCC tree had a probability of being  
181 recovered higher than 0.90.

182 There were four diversification shifts found with more than 90% of probability in the trees from the  
183 random sample (Table 2): (i) net diversification rate increase in the genus *Ypthima*; (ii) net diversification  
184 rate increase in the genus *Charaxes*; (iii) rate increase in Ithomiini subtribes Oleriina + Ithomiina +  
185 Napeogenina + Dircennina + Godryridina; and (iv) rate increase in *Callicore* + *Diaethria*.

186 MultiMEDUSA provided mean and standard deviation statistics for the diversification values found on  
187 the shifts on the 1000 trees (supp. mat. 13–14), and found that some of the changes in net diversification  
188 rate values had great variation across the posterior distribution of trees. A boxplot of the net diversifica-  
189 tion rate values estimated for the clades that appear in the MCC tree shows that some shifts are estimated  
190 as increased or slowed diversification pace depending on the tree used for analysis (Fig. 2). This varia-

tion is especially wide for the clade formed by the genera *Magneptychia* and *Caeruleptychia* because MEDUSA estimated diversification values from six times the background net diversification rate ( $r = 0.5234$ ) to almost zero ( $r = 1.22e-01$ ). The rates for *Taenaris* were between 0.14 and 0.44 (mean value 0.25). Similar degrees of variation were found in the nodes for *Ypthima*, *Charaxes* and *Coenonympha* (Fig. 2). The net diversification rates estimates for the clades (Oleriina + Ithomiina + Napeogenina + Dircennina + Godyridina), Limenitidinae + Heliconiinae and Pseudergolinae are relatively consistent across the 1000 trees (Fig. 2).

It is also evident that not all the diversification shifts estimated on the MCC tree are consistently recovered in most of the 1000 trees. Some of the splits in the MCC tree are recovered in very few trees, for example the split for the clade Satyrini (Euptychiina + Pronophilina + Satyrina + Maniolina) is recovered with a probability of 18% (Fig. 3).

## Estimation of trait-dependent speciation rates

The MEDUSA analyses taking into account phylogenetic uncertainty estimated a net diversification rate increase in part of the clade Ithomiini across more than 95% of the trees. Our BiSSE analysis found a positive effect of the character state “feeding on Solanaceae” on the net net diversification rate on part of Ithomiini (Oleriina + Ithomiina + Napeogenina + Dircennina + Godyridina) (Fig. 4). The Markov Chain Monte Carlo algorithm was run for 10000 generations discarding the first 7500 as burnin. The estimated mean net diversification rate for taxa that do not feed on Solanaceae was  $r = 0.11$  while the net diversification rate for the Solanaceae feeders was  $r = 0.16$  (see Fig. S1 for a boxplot of speciation and extinction values for the 95% credibility intervals). The same analysis considering *Vanessa* and *Hypanartia* as non-Solanaceae feeders due to dubious records produced the same pattern and net diversification rates (Fig. S2). Therefore, the rest of the analyses were performed assuming these two genera as Solanaceae feeders. We constrained the BiSSE likelihood model to force equal rates of speciation for both character states in order to test whether the model of different speciation rates is a significantly better explanation for the data. A likelihood ratio test found that the model for increased net diversification rate for nymphalids feeding on Solanaceae is a significantly better explanation than this character state having no effect on diversification ( $p < 0.001$ ) (Table 3; character states available in supp. mat. 17, code in supp. mat. 18, and mcmc run in supp. mat. 19). We combined the post-burnin mcmc generations from running BiSSE on 250 trees from the posterior distribution and found the same pattern as the BiSSE analysis on the maximum clade credibility tree (combined mcmc run in supp. mat. 20; profiles plot of speciation rates in Fig. S3; boxplot of 95% credibility intervals in Fig. S4). A BiSSE analysis to test whether the trait “feeding on Apocynaceae” had any effect on increased net diversification rates found similar speciation rates for lineages feeding on Apocynaceae and other

plants (Fig. S5). It has been shown that BiSSE performs poorly under certain conditions (Davis *et al.*, 2013). However, our data has adequate number of taxa under analysis (more than 300 tips), adequate speciation bias (between 1.5x and 2.0x), character state bias (around 8x) and extinction bias (around 4x) for the analysis of Solanaceae hostplants. Thus, BiSSE is expected to produce robust results (Davis *et al.*, 2013).

## Discussion

### Effects of phylogenetic uncertainty on the performance of MEDUSA

The MEDUSA method has been used to infer changes in net diversification rates along a phylogenetic tree. Since its publication (Alfaro *et al.*, 2009) the results of using MEDUSA on a single tree, the maximum clade credibility tree, have been used for generation of hypotheses and discussion (e.g. Litman *et al.*, 2011; Heikkilä *et al.*, 2012; Ryberg & Matheny, 2012). However, we found that MEDUSA estimated different diversification shifts and different rates of diversification for certain lineages when phylogenetic uncertainty was taken into account by using MEDUSA on a random sample of trees from the posterior distribution of a Bayesian run. We found that some diversification splits, estimated on the Nymphalidae maximum clade credibility tree, were found with a very low probability in a random sample of 1000 trees from the posterior distribution (Fig. 3, Table 2). We also found that, even though MEDUSA could estimate the same diversification splits on two or more trees, the estimated net diversification rates could vary widely (Fig. 2). For example, in our Nymphalidae trees, we found that the split for *Magneptychia* and *Caeruleptychia* had a variation from  $r = 0.5234$ , higher than the background net diversification rate, to almost zero. This means that observed patterns and conclusions can be completely contradictory depending on tree choice.

In this study, the effect of phylogenetic uncertainty on the inferred diversification splits by MEDUSA is amplified because some Nymphalidae taxa appear to be strongly affected by long-branch attraction artifacts (Peña *et al.*, 2011). Thus, the Bayesian runs are expected to recover alternative topologies on the posterior distribution of trees, resulting in low support and posterior probability values for the nodes. For example, posterior probability values for clades in Satyrini are very low (0.5 to 0.6; Wahlberg *et al.*, 2009). As a result, MEDUSA inferred a net diversification rate increase for part of Satyrini in the maximum clade credibility tree, but this was recovered only in 13% of probabilities in the MultiMEDUSA analysis on the random sample of trees.

If there is strong phylogenetic signal for increases or decreases in net diversification rates for a node, it is expected that these splits would be inferred by MEDUSA in most of the posterior distribution of trees. However, weak phylogenetic signal for some nodes can cause some clades to be absent in some trees



and MEDUSA will be unable to estimate any diversification shift (due to a non-existent node). This is the reason why MEDUSA estimated net diversification rate splits with more than 90% of probabilities in the sample of trees for only four splits: the genus *Charaxes*, the genus *Ypthima*, part of Ithomiini and the clade *Callicore* + *Diaethria* (Fig. 3), while estimating splits for other lineages with much lower probability.

The clade Ithomiini and the non-basal danaiids are well supported by high posterior probability values in Wahlberg et al. (2009). Therefore our MEDUSA analyses recovered an increase in net diversification rate with probability higher than 90% in the posterior distribution of trees (Fig. 3).

## Hostplant use and diversification in Nymphalidae

**Ithomiini** Keith Brown suggested that feeding on Solanaceae was an important event in the diversification of Ithomiini butterflies (Brown Jr., 1987). Ithomiini butterflies are exclusively Neotropical and most species feed on Solanaceae hostplants during larval stage (supp. mat. 17; Willmott & Freitas, 2006). Optimizations of the evolution of hostplant use on phylogenies evidence a probable shift from Apocynaceae to Solanaceae in the ancestor of the tribe (Brower *et al.*, 2006; Willmott & Freitas, 2006). Fordyce (2010) found that the Gamma statistics, a LTT plot of an Ithomiini phylogeny and the fit of the density-dependent model of diversification are consistent with a burst of diversification in Ithomiini following the shift from Apocynaceae to Solanaceae.

We investigated whether the strong signal for an increase in net diversification rate for Ithomiini (found by MEDUSA) can be explained due to the use of Solanaceae plants as hosts during larval stage. For this, we used a Bayesian approach (BiSSE; FitzJohn *et al.*, 2009) to test whether the trait “feeding on Solanaceae” had any effect on the diversification of the group.

Our BiSSE analysis, extended to take into account missing taxa and phylogenetic uncertainty, shows a significantly higher net diversification rate for Ithomiini taxa, which can be attributed to the trait “feeding on Solanaceae hostplants” (Fig. 4). This is in agreement with the findings of Fordyce (2010) using other statistical methods. Due to the fact that Ithomiini are virtually the only nymphalids using Solanaceae as hostplants, it is possible that the trait responsible for a higher diversification of Ithomiini might not be the hostplant character. As noted by Maddison et al. (2007), the responsible trait might be a codistributed character such as a trait related to the ability to digest secondary metabolites.

Solanaceae plants contain chemical compounds and it has been suggested that the high diversity of Ithomiini is consistent with the “escape-and-radiate scenario” due to a shift onto Solanaceae (Fordyce, 2010) and radiation scenarios among chemically different lineages of Solanaceae plants (Brown Jr., 1987; Willmott & Freitas, 2006). According to this theory, the shift from Apocynaceae to Solanaceae allowed

288 Ithomiini to invade newly available resources due to a possible key innovation that allowed them cope  
289 with secondary metabolites of the new hosts. Additional studies are needed to identify the actual enzymes  
290 that Ithomiini species might be using for detoxification of ingested food as they have been found in other  
291 butterfly groups (for example Pieridae larvae feeding on Brassicales hosts; Wheat *et al.*, 2007).

292 The increase in diversification inferred by MEDUSA occurred after the probable shift from Apocynaceae  
293 to Solanaceae, as the Solanaceae feeders in the subtribes Melinaeina and Mechanitina are not included  
294 in the diversification shift (shift number 5 in Fig. 1). The apparent conflicting results from MEDUSA  
295 and BiSSE can be explained by the low species-richness of the subtribes Melinaeina and Mechanitina  
296 compared to the other subtribes included in the shift (52 versus 272 species). It can be that MEDUSA  
297 is more conservative than BiSSE and is not including Melinaeina and Mechanitina in the shift due to  
298 low species numbers.

299 Although the Solanaceae genera used by the Ithomiini clades are well known (Willmott & Freitas, 2006),  
300 we do not have any understanding on the physiological routes involved in the detoxification of Solanaceae  
301 compounds by the several lineages of Ithomiini. We can speculate that older lineages exploiting a novel  
302 toxic resource (as the subtribes Melinaeina and Mechanitina; Willmott & Freitas, 2006; Wahlberg *et*  
303 *al.*, 2009) might not be too efficient in metabolizing plant toxins and that younger lineages are able to  
304 deal with toxins more efficiently, so that host switching events within Solanaceae are possible, which  
305 can lead to higher diversification. Studies in *Papilio* species have reported that detoxification enzymes  
306 can become more efficient in metabolizing toxins than ancestral configurations of the proteins, providing  
307 more opportunities for hostplant switch (Li *et al.*, 2003). This might be the reason why the basal  
308 Ithomiini subtribes Melinaeina and Mechanitina are so species-poor and restricted to few Solanaceae  
309 hosts (Willmott & Freitas, 2006), while recent subtribes are species-rich and have expanded their host  
310 range into several Solanaceae lineages (Willmott & Freitas, 2006). It might be that the switch to feeding  
311 in Solanaceae was an important event in the evolutionary history of Ithomiini, but the actual radiation  
312 occurred after critical physiological changes (a probable key innovation) allowed efficient detoxification  
313 of Solanaceae toxins.

314 The diffuse cospeciation hypothesis predicts almost identical ages of insects and their hostplants, while  
315 the “resource abundance-dependent diversity” and the “escape-and-radiate” hypotheses state that insects  
316 diversify after their hostplants (Ehrlich & Raven, 1964; Janz, 2011; Nyman *et al.*, 2012). Wheat *et al.*  
317 (2007) found strong evidence for a model of speciation congruent with Ehrlich and Raven’s hypothesis in  
318 Pieridae butterflies due to, in addition to the identification of a key innovation, a burst of diversification  
319 in glucosinolate-feeding taxa shortly afterwards (with a lag of ~10 MY). According to a recent dated  
320 phylogeny of the Angiosperms (Bell *et al.*, 2010), the family Solanaceae split from its sister group about  
321 59 (49-68) MYA and diversification started (crown group age) around 37 (29-47) MYA. Wahlberg *et al.*

(2009) give the corresponding ages for Ithomiini as 45 (39-53) and 37 (32-43) MYA, respectively. Thus, current evidence shows that Solanaceae and Ithomiini might have diversified around the same time, during the Late Eocene and Oligocene, and this is congruent with the diffuse cospeciation hypothesis.

The uplift of the Andes was a major tectonic event that underwent higher activity during the Oligocene (Somoza, 1998). This caused climatic changes in the region that affected the flora and fauna of the time, which coincides with the diversification of modern montane plant and animal taxa (Hoorn *et al.*, 2010) including Ithomiini butterflies and Solanaceae hostplants. Moreover, all Solanaceae clades currently present in New World originated in South America (Olmstead, 2013) as well as Ithomiini butterflies (Wahlberg *et al.*, 2009). This suggests a process of “diffuse cospeciation” of Ithomiini and hostplants.

**Danaini** Our MultiMEDUSA approach gave a significant slowdown in net diversification rate in the subtribe Danaina of the Danini. Both Danaina and the sister clade Euploeina feed mainly on Apocynaceae and thus a hostplant shift should not be responsible for the observed slowdown of diversification in the Danaina. As expected, our BiSSE analysis of Apocynaceae feeders shows that there is no effect of feeding on this plant family on the net diversification rates of Nymphalidae lineages. Many of the Danaina are large, strong fliers, highly migratory and involved in mimicry rings. Among them is for example the monarch (*Danaus plexippus*), probably the most well known of all migratory butterflies. The causes for a lower net diversification rate in the Danaina remains to be investigated, but their great dispersal power might be involved in preventing allopatric speciation. It has been found in highly vagile species in the nymphalid genus *Vanessa* that dispersal has homogenized populations due to gene flow, as old and vagile species seem to be genetically homogeneous while younger widespread species show higher genetic differentiation in their populations (Wahlberg & Rubinoff, 2011).

**Charaxes** The genus *Charaxes* contain 193 species distributed in the Old World with highest diversity in the Afrotropical region. These butterflies are also very strong fliers, but contrary to Danaina, which are specialized Apocynaceae feeders, *Charaxes* are known to feed on at least 28 plant families in 18 orders (Ackery, 1988) and some species appear to be polyphagous (Müller *et al.*, 2010). Aduse-Poku *et al.* (2009) showed that most of the diversification of the genus occurred from the late Oligocene to the Miocene when there were drastic global climatic fluctuations, indicating that the diversity mainly is driven by climate change. Müller *et al.* (2010) found that climatic changes during the Pliocene, Pleistocene as well as dispersal and vicariance might have been responsible for the high diversification of the genus.

**Satyrini** The diverse tribe Satyrini radiated simultaneously with the radiation of their main hostplant, grasses, during the climatic cooling in the Oligocene (Peña & Wahlberg, 2008). Thus, it is somewhat surprising that part of Satyrini (the subtribes Euptychiina, Satyrina and Pronophilina) were found to

354 have accelerated diversification in only 13% of probabilities in the trees from the posterior distribution.  
355 Although this can be attributed to low phylogenetic signal (posterior probability value = 0.6 for this  
356 clade in Wahlberg *et al.*, 2009), the clade Satyrini is very robust (posterior probability value = 1.0  
357 for this clade in Wahlberg *et al.*, 2009) and MEDUSA failed to identify any significant accelerated net  
358 diversification rate for Satyrini. It appears that the radiation of Satyrini as a whole was not remarkably  
359 fast and therefore not picked up by MEDUSA.

360 The origin of the tribe Satyrini is not completely clear (originated either in the Neotropical or Eastern  
361 Palaearctic, Oriental and/or Indo-Australian regions (Peña *et al.*, 2011) and their radiation involved  
362 colonizing almost all continents starting from their place of origin. For butterflies, a prerequisite for  
363 colonizing new areas is that suitable hostplants are already present. Therefore, it should be expected  
364 that the diversification of Satyrinae occurred in a stepwise manner, with pulses or bursts of diversification  
365 for certain lineages but unlikely for the tribe Satyrini as a whole.

366 We found that even though MEDUSA estimated several diversification shifts in the maximum clade  
367 credibility tree of Nymphalidae, only a few of these splits were found in more than 90% of the trees from  
368 the posterior distribution. In the literature, it is common practice that conclusions are based on the splits  
369 estimated on the maximum clade credibility tree. However, by using a MultiMEDUSA approach, we  
370 found that some of this splits might be greatly affected by phylogenetic uncertainty. Moreover, some of  
371 these splits can be recovered either as increases or decreases in net diversification rate depending on the  
372 tree from the posterior distribution that was used for analysis. This means that contradictory conclusions  
373 would be made if only the maximum clade credibility tree was used for analysis. MEDUSA appears to  
374 be sensitive to the number of nodes with high posterior probability and width of age confidence intervals.  
375 For our data, it would be necessary to obtain a posterior distribution of trees with no conflicting topology,  
376 and very similar estimated ages for nodes in order to consistently recover most of the diversification splits  
377 on the posterior distribution of trees that were inferred by MEDUSA on the MCC tree.

378 Our MultiMEDUSA approach to perform analyses on the posterior distribution of trees found strong  
379 support for an increase in net diversification rate in the tribe Ithomiini and the genus *Charaxes*, and for a  
380 decrease in net diversification rate in the subtribe Danaina. Due to phylogenetic uncertainty, we did not  
381 obtain strong support for other diversification splits in Nymphalidae. Our BiSSE analysis corroborated  
382 other studies in that the trait “feeding on Solanaceae”, or a codistributed character, was important in the  
383 diversification of Ithomiini butterflies. However, by applying MEDUSA we found that a critical character  
384 in the radiation of the group might have appeared after the shift from Apocynaceae to Solanaceae. We  
385 also found that the trait “feeding on Apocynaceae” is not responsible for the slowdown of diversification  
386 in Danaina. Ithomiini and Solanaceae diversified near simultaneously, which is in agreement with the  
387 diffuse co-speciation hypothesis (Janz, 2011; Nyman *et al.*, 2012).

## 388 Acknowledgments

389 We are thankful to Mark Cornwall for help with the script to extend MEDUSA to include phylogenetic  
390 uncertainty, Niklas Wahlberg for commenting on the manuscript and giving us the posterior distribution  
391 of trees, Luke Harmon for commenting on the manuscript and anonymous reviewers for their comments,  
392 which greatly improved the manuscript, Jessica Slove Davidson and Niklas Janz for access to their host-  
393 plant data. The study was supported by a Kone Foundation grant (awarded to Niklas Wahlberg), Finland  
394 (C. Peña) and the Research Council of Norway (grant no. 204308 to M. Espeland). We acknowledge  
395 CSC-IT Center for Science Ltd. (Finland) for the allocation of computational resources.

## 396 Figure legends

397 **Figure 1.** Results of the MEDUSA analysis run on the maximum clade credibility tree from Wahlberg  
398 et al. (2009). Rate shifts were estimated for the following nodes (besides the background rate): 2) Limeni-  
399 tidinae + Heliconiinae, 3) *Ypthima*, 4) *Charaxes*, 5) Ithomiini in part, 6) Satyrini, 7) Coenonymphina  
400 8) Phyciodina in part, 9) Danaini in part, 10) *Coenonympha*, 11) *Caeruleptychia* + *Magneptychia*,  
401 12) *Callicore* + *Diaethria*, 13) Satyrina, 14) Mycalesinae, 15) *Pedaliodes*, 16) *Dryas* + *Dryadula*, 17)  
402 *Taenaris*, 18) Pseudergolinae.

403 **Figure 2.** Boxplot of the range of diversification values for tips or clades estimated by MEDUSA on  
404 the 1000 random trees from the posterior distribution of the Nymphalidae phylogeny. The tips or clades  
405 shown are those present on the maximum clade credibility tree from Wahlberg et al. (2009).

406 **Figure 3.** Results of the MultiMEDUSA analysis on 1000 randomly sampled trees from the posterior  
407 distribution of the Nymphalidae phylogeny. Bars show the probability for nodes of being recovered as  
408 significant increases or decreases in net diversification rates by MultiMEDUSA.

409 **Figure 4.** BiSSE analysis of diversification of nymphalids due to feeding on Solanaceae hostplants.  
410 Speciation and net diversification rates are significantly higher in Solanaceae feeders ( $\lambda_1$ ,  $r_1$ ).

## 411 Tables

412 **Table 1.** Significant net diversification rate shifts found in the MEDUSA analysis of the Nymphalid  
413 maximum clade credibility tree. Split.Node = node number, Model = preferred diversification model by  
414 MEDUSA, r = net diversification rate, LnLik.part = log likelihood value.

415 **Table 2.** Differences in rates estimated by MEDUSA on the MCC tree from Wahlberg and the Multi-  
416 MEDUSA approach on 1000 random trees from the posterior distribution.

417 Table 3. Likelihood ratio test between the model of increased diversification of nymphalids feeding on  
418 Solanaceae against a model forcing equal speciation rates (no effect on diversification).

## 419 **Supp. mat.**

420 `supp_mat_01_genus.nex`: MCC Nymphalidae tree from Wahlberg et al. (2009).

421 `supp_mat_02_run_medusa_on_mct.R`: Run MEDUSA on MCC tree from Wahlberg et al. (2009).

422 `supp_mat_03_richness.nex`: Species richness for lineages in Nymphalidae.

423 `supp_mat_04_MEDUSA_results.txt`: Results of running MEDUSA on the MCC tree.

424 `supp_mat_05_1000_random_trees_no_outgroups.nex`: 1000 random trees from Wahlberg et al. (2009).

425 `supp_mat_06_run_multimedusa.R`: Run MultiMEDUSA on 1000 random trees from Wahlberg et al.  
426 (2009).

427 `supp_mat_07_multimedusa_on_1000_trees.txt`: Raw results from the MultiMEDUSA run on the ran-  
428 dom sample of trees from the posterior distribution.

429 `supp_mat_08_multimedusa_result01.pdf`: Summary of MultimedUSA analysis on the 1000 trees from  
430 the posterior distribution.

431 `supp_mat_09_mcc_tree_from_random_sample_of_1000.tree`: MCC tree from the 1000 random trees  
432 selected from the posterior distribution.

433 `supp_mat_10_run_medusa_on_mct_from_1000_trees.R`: Run MEDUSA on MCC tree from the 1000  
434 random trees selected from the posterior distribution.

435 `supp_mat_11_results_of_MEDUSA_run_on_mcc_from_1000_trees.txt`: Summary results from a  
436 MEDUSA run on the MCC tree from the 1000 random trees selected from the posterior distribution.

437 `supp_mat_12_MEDUSA_mcc_tree_from_random_sample_of_1000_trees.pdf`: Figure for MEDUSA run  
438 on MCC tree from random 1000 trees.

439 `supp_mat_13_results_of_Multi-MEDUSA_run_on_1000_trees.txt`: Summary results from a MultiME-  
440 DUSA run on the 1000 random trees selected from the posterior distribution.

441 `supp_mat_14_multimedusa_result_prob_nodes.pdf`: Probability of nodes with estimated rates from  
442 a MultiMEDUSA run on the 1000 random trees selected from the posterior distribution.

443 `supp_mat_15_host_plant_data.csv`: Hostplants of Nymphalidae butterflies recorded from the litera-  
444 ture.

445 `supp_mat_16_host_plant_data_references.csv`: References for hostplants data.

446 `supp_mat_17_hostplant_states.csv`: Data matrix with character states for hostplant use.

447 `supp_mat_18_bisse_source.R`: R code for running the BiSSE analysis.

448 `supp_mat_19_bisse_mcmc_run.csv`: Raw results for the BiSSE analysis.

449 `supp_mat_20_combined_bisse.csv`: Raw results for the combined BiSSE analysis.

450 **Figure S1.** Boxplot of speciation ( $\lambda$ ) and extinction ( $\mu$ ) values for the 95% credibility intervals  
 451 of values estimated by BiSSE analysis of diversification due to feeding on Solanaceae plants.

452 **Figure S2.** BiSSE analysis of diversification of nymphalids due to feeding on Solanaceae hostplants  
 453 assuming *Vanessa* and *Hypanartia* as non-Solanaceae feeders. The same pattern is recovered, speciation  
 454 and net diversification rates are significantly higher for Solanaceae feeders ( $\lambda_1$ ,  $r_1$ ).

455 **Figure S3.** Net diversification rates of nymphalids feeding on Solanaceae plants as estimated by com-  
 456 bining post-burnin runs of BiSSE on the 1000 trees from the posterior distribution.

457 **Figure S4.** Boxplot of speciation and extinction values for the 95% credibility intervals of values  
 458 estimated by BiSSE analysis of diversification due to feeding on Solanaceae plants on the combined  
 459 post-burnin runs on 1000 trees from the posterior distribution.

460 **Figure S5.** BiSSE analysis of diversification of nymphalids due to feeding on Apocynaceae hostplants.  
 461 Speciation and net diversification rates are similar.

## 462 References

- 463 Ackery, P.R. 1988. Hostplants and classification: a review of nymphalid butterflies. *Biol. J. Linn. Soc.*  
 464 **33**: 95–203.
- 465 Aduse-Poku, K., Vingerhoedt, E. & Wahlberg, N. 2009. Out-of-africa again: a phylogenetic hypothesis  
 466 of the genus *Charaxes* (Lepidoptera: Nymphalidae) based on 5 gene regions. *Mol. Phylogenet. Evol.* **53**:  
 467 463–478.
- 468 Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A. & Rabosky, D.L. *et al.* 2009. Nine  
 469 exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *P. Natl. Acad.*  
 470 *Sci. USA* **106**: 13410–13414.
- 471 Beccaloni, G., Vitoria, Á., Hall, S. & Robinson, G. 2008. *Catalogue of the hostplants of the neotropical*  
 472 *butterflies*. The Natural History Museum, London.
- 473 Bell, C.D., Soltis, D.E. & Soltis, P.S. 2010. The age and diversification of the angiosperms re-revisited.  
 474 *Am. J. Bot.* **97**: 1296–1303.

475 Brower, A.V.Z., Freitas, A.V.L., Lee, M.-M., Silva-Brandão, K.L., Whinnett, A. & Willmott, K.R.  
476 2006. Phylogenetic relationships among the ithomiini (lepidoptera: nymphalidae) inferred from one  
477 mitochondrial and two nuclear gene regions. *Syst. Entomol.* **31**: 288–301.

478 Brower, A.V.Z., Wahlberg, N., Ogawa, J.R., Boppré, M. & Vane-Wright, R.I. 2010. Phylogenetic rela-  
479 tionships among genera of danaine butterflies (lepidoptera: nymphalidae) as implied by morphology and  
480 DNA sequences. *Syst. Biodivers.* **8**: 75–89.

481 Brown Jr., K.S. 1987. Chemistry at the Solanaceae/Ithomiinae interface. *Ann. Mo. Bot. Gard.* **74**:  
482 359–397.

483 Core Team, R. 2013. *R: a language and environment for statistical computing*. R Foundation for  
484 Statistical Computing, Vienna, Austria.

485 Davis, M.P., Midford, P.E. & Maddison, W. 2013. Exploring power and parameter estimation of the  
486 BiSSE method for analyzing species diversification. *BMC Evol. Biol.* **13**: 38.

487 De-Silva, D.L., Day, J.J., Elias, M., Willmott, K.R., Whinnett, A. & Mallet, J. 2010. Molecular phyloge-  
488 netics of the neotropical butterfly subtribe oleriina (nymphalidae: danainae: ithomiini). *Mol. Phylogenet.*  
489 *Evol.* **55**: 1032–1041.

490 Dyer, L.A. & Gentry, G.L. 2002. Caterpillars and parasitoids of a tropical lowland wet forest.

491 Ehrlich, P.R. & Raven, P.H. 1964. Butterflies and plants: a study in coevolution. *Evolution* **18**: 586–608.

492 Elias, M., Joron, M., Willmott, K.R., Silva-Brandão, K.L., Kaiser, V. & Arias, C.F.*et al.* 2009. Out of  
493 the andes: patterns of diversification in clearwing butterflies. *Mol. Ecol.* **18**: 1716–1729.

494 Ferrer-Paris, J.R., Sánchez-Mercado, A., Vilorio, Á.L. & Donaldson, J. 2013. Congruence and diversity  
495 of butterfly-host plant associations at higher taxonomic levels. *PloS one* **8**: e63570. Public Library of  
496 Science.

497 FitzJohn, R.G. 2012. Diversitree: comparative phylogenetic analyses of diversification in r. *Method.*  
498 *Ecol. Evol.* **3**: 1084–1092.

499 FitzJohn, R.G., Maddison, W.P. & Otto, S.P. 2009. Estimating trait-dependent speciation and extinction  
500 rates from incompletely resolved phylogenies. *Syst. Biol.* **58**: 595–611.

501 Fordyce, J.A. 2010. Host shifts and evolutionary radiations of butterflies. *P. Roy. Soc. B* **277**: 3735–  
502 3743.

503 Freitas, A.V.L. & Brown Jr., K.S. 2004. Phylogeny of the nymphalidae (lepidoptera). *Syst. Biol.* **53**:  
504 363–383.

505 Harmon, L.J., Rabosky, D.L., FitzJohn, R.G. & Brown, J.W. 2011. *turboMEDUSA: MEDUSA: modeling*  
506 *evolutionary diversification using stepwise AIC*.



507 Harmon, L.J., Weir, J.T., Brock, C.D., Glor, R.E. & Challenger, W. 2008. GEIGER: investigating  
508 evolutionary radiations. *Bioinformatics* **24**: 129–131.

509 Heikkilä, M., Kaila, L., Mutanen, M., Peña, C. & Wahlberg, N. 2012. Cretaceous origin and repeated  
510 tertiary diversification of the redefined butterflies. *P. Roy. Soc. B* **279**: 1093–9.

511 Hoorn, C., Wesselingh, F.P., Steege, H. ter, Bermudez, M.A., Mora, A. & Sevink, J.*et al.* 2010. Amazonia  
512 through time: andean uplift, climate change, landscape evolution, and biodiversity. *Science* **330**: 927–  
513 931.

514 Janz, N. 2011. Ehrlich and raven revisited: mechanisms underlying codiversification of plants and  
515 enemies. *Annu. Rev. Ecol. Evol. S.* **42**: 71–89.

516 Janzen, D.H. & Hallwachs, W. 2009. Dynamic database for an inventory of the macrocaterpillar fauna,  
517 and its food plants and parasitoids, of area de conservacion guanacaste (aCG), northwestern costa rica.

518 Kodandaramaiah, U., Lees, D.C., Müller, C.J., Torres, E., Karanth, K.P. & Wahlberg, N. 2010. Phy-  
519 logenetics and biogeography of a spectacular old world radiation of butterflies: the subtribe mycalesina  
520 (lepidoptera: nymphalidae: satyrini). *BMC Evol. Biol.* **10**: 172.

521 Kodandaramaiah, U., Peña, C., Braby, M.F., Grund, R., Müller, C.J. & Nylin, S.*et al.* 2010. Phylo-  
522 genetics of coenonymphina (nymphalidae: satyrinae) and the problem of rooting rapid radiations. *Mol.*  
523 *Phylogenet. Evol.* **54**: 386–394.

524 Lamas, G. 2004. *Checklist: part 4A. hesperioidea-papilionoidea*. Gainesville.

525 Li, W., Schuler, M.A. & Berenbaum, M.R. 2003. Diversification of furanocoumarin-metabolizing cy-  
526 tochrome p450 monooxygenases in two papilionids: specificity and substrate encounter rate. *P. Natl.*  
527 *Acad. Sci. USA* **100**: 14593–14598.

528 Litman, J.R., Danforth, B.N., Eardley, C.D. & Praz, C.J. 2011. Why do leafcutter bees cut leaves? new  
529 insights into the early evolution of bees. *P. Roy. Soc. B* **278**: 3593–3600.

530 Maddison, W.P., Midford, P.E. & Otto, S.P. 2007. Estimating a binary character’s effect on speciation  
531 and extinction. *Syst. Biol.* **56**: 701–710.

532 Matos-Maraví, P.F., Peña, C., Willmott, K.R., Freitas, A.V. & Wahlberg, N. 2013. Systematics and  
533 evolutionary history of butterflies in the “Taygetis clade” (nymphalidae: satyrinae: euptychiina): towards  
534 a better understanding of neotropical biogeography. *Mol. Phylogenet. Evol.* **66**: 54–68.

535 Mitter, C., Farrell, B. & Wiegmann, B. 1988. The phylogenetic study of adaptive zones: has phytophagy  
536 promoted insect diversification? *Am. Nat.* **132**: 107–128.

537 Müller, C.J., Wahlberg, N. & Beheregaray, L.B. 2010. ‘After africa’: the evolutionary history and

systematics of the genus *Charaxes* *ochsenheimer* (Lepidoptera: Nymphalidae) in the Indo-Pacific region.  
*Biol. J. Linn. Soc.* **100**: 457–481.

Nieukerken, E.J. van, Kaila, L., Kitching, I.J., Kristensen, N.P., Lees, D.C. & Minet, J.*et al.* 2011. Order  
 Lepidoptera. In: *Animal biodiversity: an outline of higher-level classification and survey of taxonomic  
 richness*, pp. 212–221. Zootaxa.

Nylin, S., Slove, J. & Janz, N. 2013. Host plant utilization, host range oscillations and diversification in  
 nymphalid butterflies: a phylogenetic investigation. *Evolution* in press. Wiley Online Library.

Nyman, T., Linder, H.P., Peña, C., Malm, T. & Wahlberg, N. 2012. Climate-driven diversity dynamics  
 in plants and plant-feeding insects. *Ecol. Lett.* **15**: 889–898.

Olmstead, R.G. 2013. Phylogeny and biogeography in Solanaceae, Verbenaceae and Bignoniaceae: a  
 comparison of continental and intercontinental diversification patterns. *Bot. J. Linn. Soc.* **171**: 80–  
 102.

Ortiz-Acevedo, E. & Willmott, K.R. 2013. Molecular systematics of the butterfly tribe Preponini  
 (Nymphalidae: Charaxinae). *Syst. Entomol.* no–no.

Penz, C.M. 1999. Higher level phylogeny for the passion-vine butterflies (Nymphalidae, Heliconiinae)  
 based on early stage and adult morphology. *Zool. J. Linn. Soc.* **127**: 277–344.

Peña, C. & Wahlberg, N. 2008. Prehistorical climate change increased diversification of a group of  
 butterflies. *Biol. Lett.* **4**: 274–278.

Peña, C., Nylin, S. & Wahlberg, N. 2011. The radiation of Satyrini butterflies (Nymphalidae: Satyrinae):  
 a challenge for phylogenetic methods. *Zool. J. Linn. Soc.* **161**: 64–87.

Peña, C., Nylin, S., Freitas, A.V.L. & Wahlberg, N. 2010. Biogeographic history of the butterfly subtribe  
 Euptychiina (Lepidoptera, Nymphalidae, Satyrinae). *Zool. Scr.* **39**: 243–258.

Peña, C., Wahlberg, N., Weingartner, E., Kodandaramaiah, U., Nylin, S. & Freitas, A.V.L.*et al.* 2006.  
 Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data.  
*Mol. Phylogenet. Evol.* **40**: 29–49.

Popescu, A.-A., Huber, K.T. & Paradis, E. 2012. ape 3.0: new tools for distance-based phylogenetics  
 and evolutionary analysis in R. *Bioinformatics* **28**: 1536–1537. Oxford Univ Press.

Ryberg, M. & Matheny, P.B. 2012. Asynchronous origins of ectomycorrhizal clades of Agaricales. *P. Roy.  
 Soc. B* **279**: 2003–2011.

Scott, J.A. 1986. *The butterflies of North America*. Stanford University Press.

Silva-Brandão, K.L., Wahlberg, N., Francini, R.B., Azeredo-Espin, A.M.L., Brown Jr., K.S. & Paluch,

569 M.*et al.* 2008. Phylogenetic relationships of butterflies of the tribe acraeini (lepidoptera, nymphalidae,  
570 heliconiinae) and the evolution of host plant use. *Mol. Phylogenet. Evol.* **46**: 515–31.

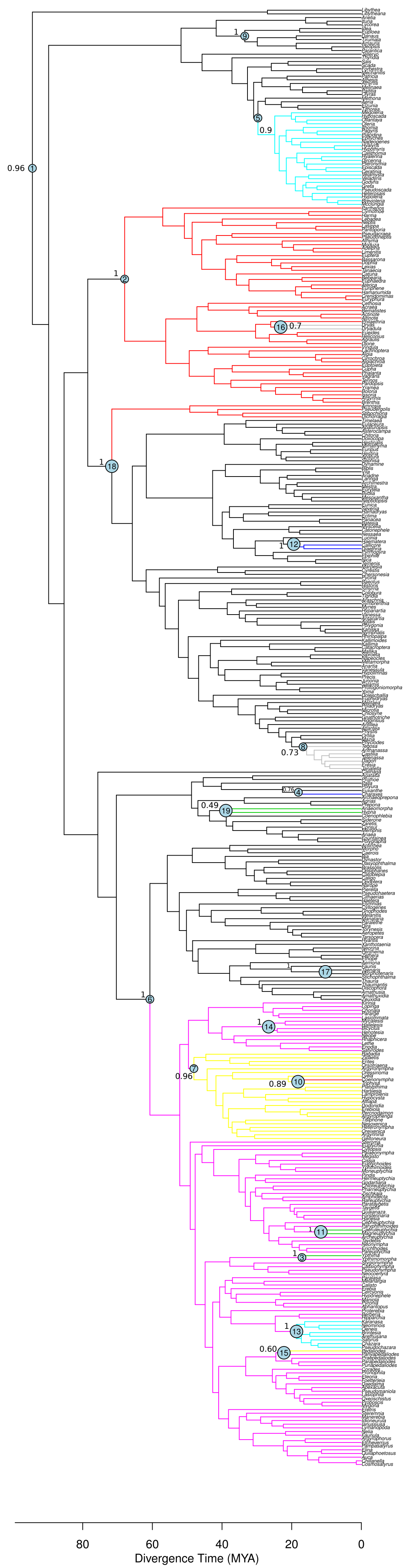
571 Somoza, R. 1998. Updated nazca (farallon) - south america relative motions during the last 40 my:  
572 implications for mountain building in the central andean region. *J. S. Am. Earth Sci.* **11**: 211–215.

573 Wahlberg, N. & Rubinoff, D. 2011. Vagility across vanessa (lepidoptera: nymphalidae): mobility in  
574 butterfly species does not inhibit the formation and persistence of isolated sister taxa. *Syst. Entomol.*  
575 **36**: 362–370.

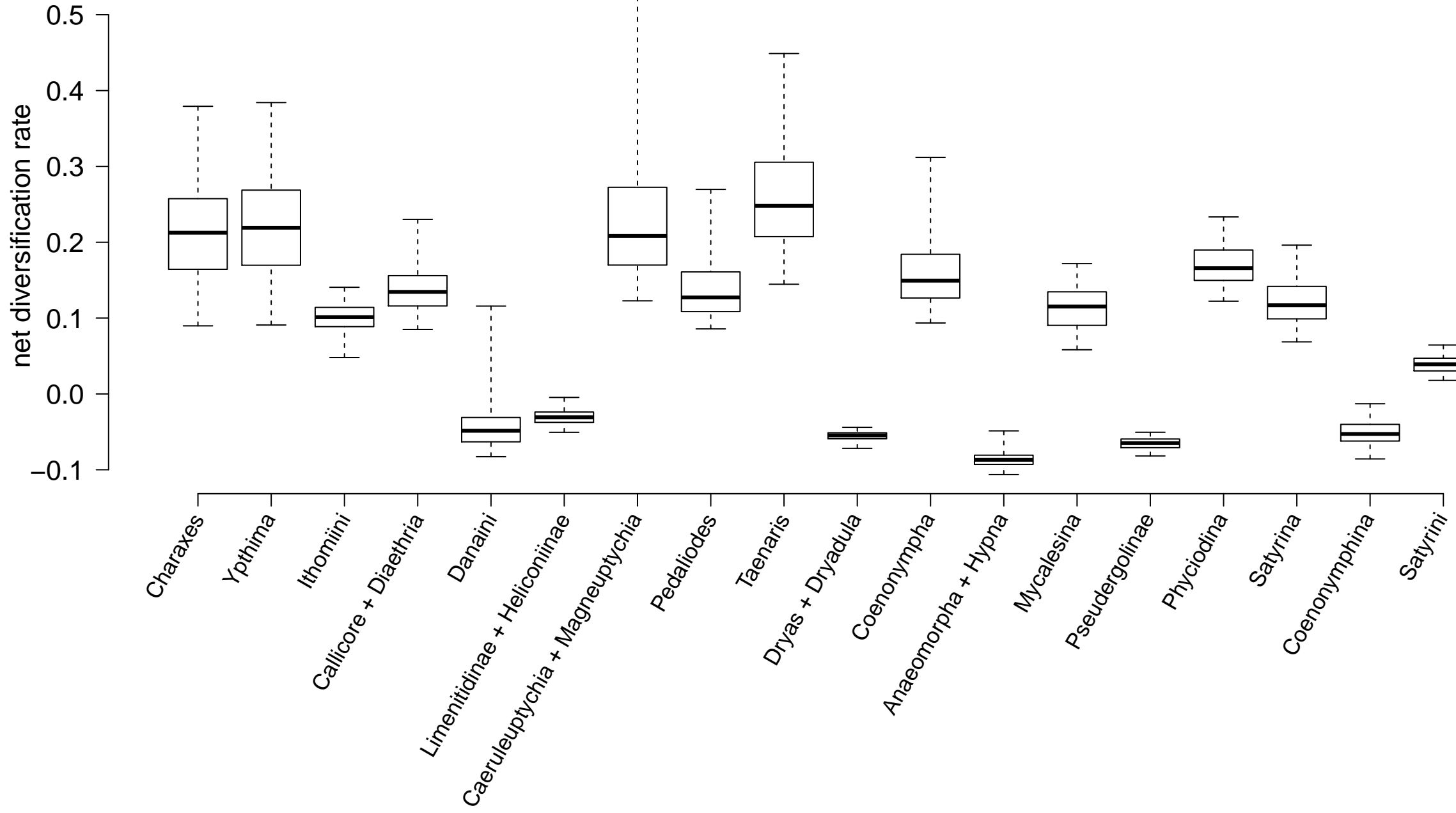
576 Wahlberg, N., Leneveu, J., Kodandaramaiah, U., Peña, C., Nylin, S. & Freitas, A.V.L.*et al.* 2009.  
577 Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. *P. Roy.*  
578 *Soc. B* **276**: 4295–4302.

579 Wheat, C.W., Vogel, H., Wittstock, U., Braby, M.B., Underwood, D. & Mitchell-Olds, T. 2007. The  
580 genetic basis of a plant-insect coevolutionary key innovation. *P. Natl. Acad. Sci. USA* **104**: 20427–  
581 20431.

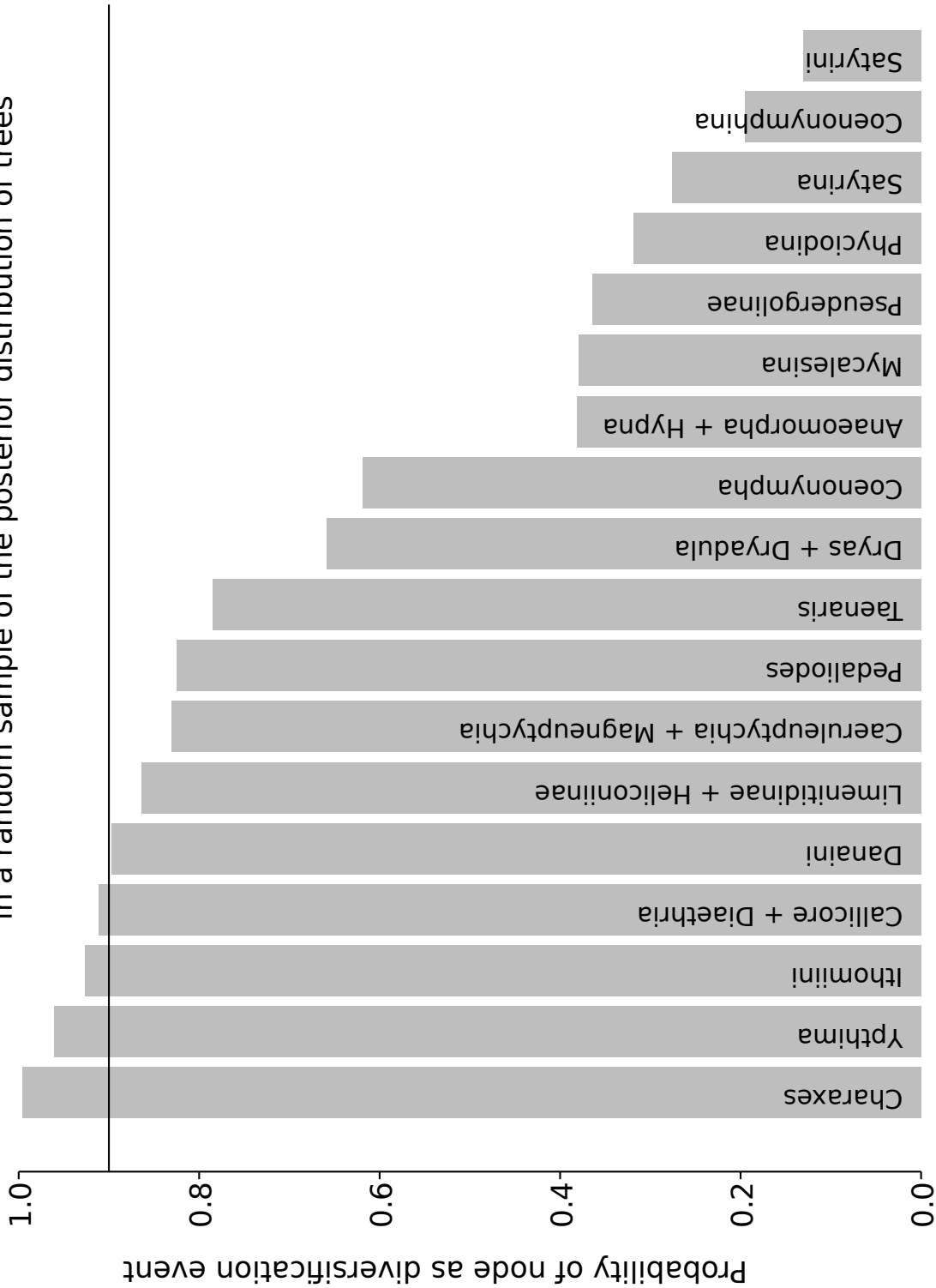
582 Willmott, K.R. & Freitas, A.V.L. 2006. Higher-level phylogeny of the ithomiinae (lepidoptera: nymphal-  
583 idae): classification, patterns of larval hostplant colonization and diversification. *Cladistics* **22**: 297–368.



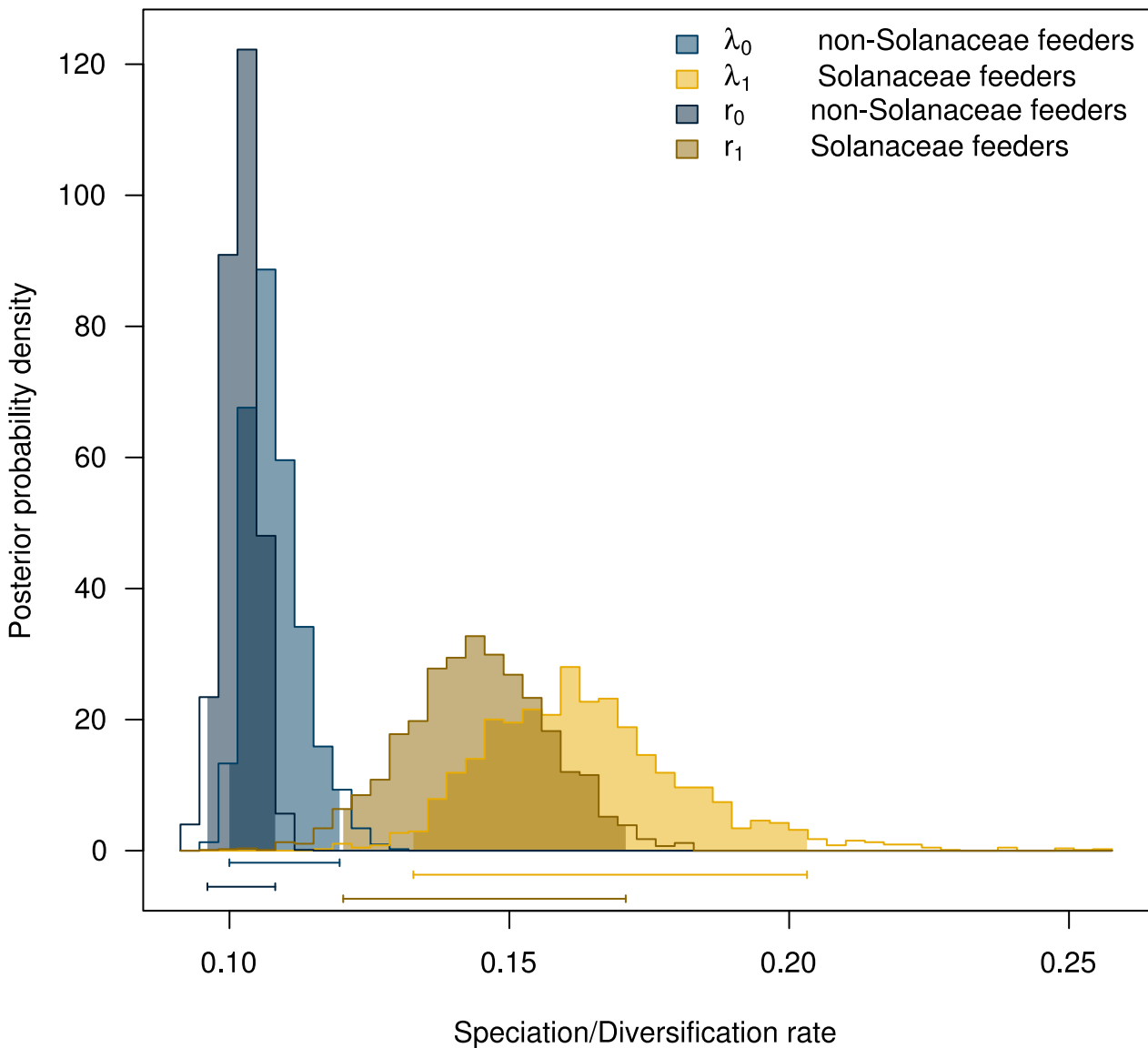
**Boxplot of diversification values for each node with significant bursts  
as found by MEDUSA, only clades/tips present in MCC tree**



Probability of finding nodes as diversification events by MultiMEDUSA in a random sample of the posterior distribution of trees



Nodes of MCC phylogenetic tree



# Sheet1

Table 1

Significant diversification rate shifts found in the Turbo-MEDUSA analysis of the Nymphalid maximum credibility tree. Split.node = node number, r = net diversification rate, LnLik.part = log likelihood value, AICc = corrected AIC score.

| Split.Node | Model    | r        | LnLik.part | AICc     | Taxa   |
|------------|----------|----------|------------|----------|--|
| 1          | 399 yule | 0.092459 | -1055.957  |          | Nymphalidae (root)   |
| 2          | 691 bd   | 0.054129 | -406.3703  |          | Limenitidinae + Heliconiinae                                     |
| 3          | 299 yule | 0.311199 | -6.3058    |          | <i>Ypthima</i>   |
| 4          | 224 yule | 0.290989 | -6.2601    |          | <i>Charaxes</i>  |
| 5          | 750 yule | 0.186913 | -147.4146  |          | Oleriina + Ithomiina +<br>Napeogenina + Dircennina +<br>Godyrina |
| 6          | 405 yule | 0.116252 | -555.0276  |          | Satyrini   |
| 7          | 495 yule | 0.064656 | -124.9143  |          | Coenonymphina  |
| 8          | 609 yule | 0.240562 | -35.0908   |          | Phyciodina in part   |
| 9          | 787 bd   | 0.042332 | -43.7819   |          | Danaini in part  |
| 10         | 231 yule | 0.209416 | -4.7955    |          | <i>Coenonympha</i>   |
| 11         | 478 yule | 0.311684 | -9.2218    |          | <i>Caeruleptychia</i> + <i>Magneptychia</i>                      |
| 12         | 659 yule | 0.219253 | -11.2686   |          | <i>Callicore</i> + <i>Diaethria</i>                              |
| 13         | 444 yule | 0.220615 | -43.7812   |          | Satyrina   |
| 14         | 524 yule | 0.190754 | -26.0651   |          | Mycalesina   |
| 15         | 355 yule | 0.234041 | -6.2013    |          | <i>Pedaliodes</i>  |
| 16         | 714 yule | 0        | 0          |          | <i>Dryas</i> + <i>Dryadula</i>                                   |
| 17         | 377 yule | 0.311671 | -4.1986    |          | <i>Taenaris</i>  |
| 18         | 688 yule | 0.024724 | -17.5256   |          | Pseudergolinae   |
| 19         | 583 yule | 0        | 0          | 5090.492 | <i>Anaeomorpha</i> + <i>Hypna</i>                                |



# Sheet1

Table 2

Differences in rates estimated by MEDUSA on the MCC tree from Wahlberg and the MultiMEDUS, approach on 1000 random trees from the posterior distribution.

| Split.Node | rate by MEDUSA | Median rate by MultiMEDUSA | probability of being recovered |
|------------|----------------|----------------------------|--------------------------------|
| 1          | 0.090          | not found                  | 0.000                          |
| 2          | 0.054          | -0.030                     | 0.864                          |
| 3          | 0.311          | 0.219                      | <b>0.961</b>                   |
| 4          | 0.290          | 0.212                      | <b>0.996</b>                   |
| 5          | 0.187          | 0.101                      | <b>0.927</b>                   |
| 6          | 0.116          | 0.039                      | 0.131                          |
| 7          | 0.064          | -0.052                     | 0.195                          |
| 8          | 0.240          | 0.166                      | 0.319                          |
| 9          | 0.042          | -0.049                     | 0.897                          |
| 10         | 0.209          | 0.149                      | 0.619                          |
| 11         | 0.311          | 0.208                      | 0.831                          |
| 12         | 0.219          | 0.135                      | <b>0.911</b>                   |
| 13         | 0.220          | 0.117                      | 0.276                          |
| 14         | 0.190          | 0.115                      | 0.379                          |
| 15         | 0.234          | 0.127                      | 0.825                          |
| 16         | 0.000          | -0.005                     | 0.659                          |
| 17         | 0.311          | 0.248                      | 0.785                          |
| 18         | 0.024          | -0.064                     | 0.024                          |
| 19         | 0.000          | -0.087                     | 0.381                          |

Table 3. Likelihood ratio test for increase of diversification rates for Solanaceae feeders versus no-effect model

|              | Df | lnLik   | AIC    | ChiSq | p       |
|--------------|----|---------|--------|-------|---------|
| full         | 6  | -1613.3 | 3238.5 |       |         |
| equal.lambda | 5  | -1619.4 | 3248.9 | 12.3  | 0.00045 |