1 **Running head**

2 THE AGE OF BUTTERFLIES REVISITED (AND TESTED)

3 **Title**

4 The Trials and Tribulations of Priors and Posteriors in Bayesian Timing of

5 Divergence Analyses: the Age of Butterflies Revisited.

6

7 **Authors**

8 NICOLAS CHAZOT^{1*}, NIKLAS WAHLBERG¹, ANDRÉ VICTOR LUCCI FREITAS²,

9 CHARLES MITTER³, CONRAD LABANDEIRA^{3,4}, JAE-CHEON SOHN⁵, RANJIT KUMAR

10 SAHOO⁶, NOEMY SERAPHIM⁷, RIENK DE JONG⁸, MARIA HEIKKILÄ⁹

11 **Affiliations**

12 ¹*Department of Biology, Lunds Universitet, Sölvegatan 37, 223 62, Lund, Sweden.*

13 ²*Departamento de Biologia Animal, Instituto de Biologia, Universidade Estadual de*
14 *Campinas (UNICAMP), Cidade Universitária Zeferino Vaz, Caixa postal 6109,*
15 *Barão Geraldo 13083-970, Campinas, SP, Brazil.*

16 ³*Department of Entomology, University of Maryland, College Park, MD 20742, U.S.A.*

17 ⁴*Department of Paleobiology, National Museum of Natural History, Smithsonian*

18 *Institution, Washington, DC 20013, USA; Department of Entomology and BEES*

19 *Program, University of Maryland, College Park, MD 20741; and Key Lab of Insect*

20 *Evolution and Environmental Change, School of Life Sciences, Capital Normal*

21 *University, Beijing 100048,*

22 ⁵*Institute of Littoral Environment, Mokpo National University, Muan, Jeonnam 58554,*
 23 *Republic of Korea*

24 ⁶*IISER-TVM Centre for Research and Education in Ecology and Evolution (ICREEE),*
 25 *School of Biology, Indian Institute of Science Education and Research,*
 26 *Thiruvananthapuram, Kerala 695 551, India.*

27 ⁷*Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Campus*
 28 *Campinas, CTI Renato Archer - Av. Comendador Aladino Selmi, s/n - Amarais,*
 29 *Campinas - SP, 13069-901, Brazil.*

30 ⁸*Naturalis Biodiversity Center, Department of Entomology, PO Box 9517, 2300 RA*
 31 *Leiden, The Netherlands.*

32 ⁹*Finnish Museum of Natural History LUOMUS, Zoology Unit, P.O. Box 17, FI-*
 33 *00014 University of Helsinki, Finland.*

34

35 **Corresponding author (*):**

36 Nicolas Chazot

37 *Department of Biology, Lunds Universitet, Sölvegatan 37, 223 62, Lund, Sweden.*

38 email: chazotn@gmail.com

39

40 **Abstract**

41 The need for robust estimates of times of divergence is essential for downstream
42 analyses, yet assessing this robustness is still rare. We generated a time-calibrated
43 genus-level phylogeny of butterflies (Papilionoidea), including 994 taxa, up to 10
44 gene fragments and an unprecedented set of 12 fossils and 10 host-plant node
45 calibration points. We compared marginal priors and posterior distributions to assess
46 the relative importance of the former on the latter. This approach revealed a strong
47 influence of the set of priors on the root age but for most calibrated nodes posterior
48 distributions shifted from the marginal prior, indicating significant information in the
49 molecular dataset. We also tested the effects of changing assumptions for fossil
50 calibration priors and the tree prior. Using a very conservative approach we estimated
51 an origin of butterflies at 107.6 Ma, approximately equivalent to the Early
52 Cretaceous–Late Cretaceous boundary, with a credibility interval ranging from 89.5
53 Ma (mid Late Cretaceous) to 129.5 Ma (mid Early Cretaceous). This estimate was
54 robust to alternative analyses changing core assumptions. With 994 genera, this tree
55 provides a comprehensive source of secondary calibrations for studies on butterflies.

56 **Keywords**

57 Papilionoidea, butterflies, time-calibration, fossils, host plants, marginal prior

58 INTRODUCTION

59 An increasing amount of molecular information is allowing the inference of broad and
60 densely sampled phylogenetic hypotheses for species-rich groups. This effort,
61 combined with the emergence of a great number of methods investigating trait
62 evolution, historical biogeography, and the dynamics of diversification have increased
63 the need for time-calibrated trees. Estimating divergence times in molecular
64 phylogenetic work depends primarily on fossils to constrain models of heterogeneous
65 rates of substitutions. Consequently, the robustness of such estimates relies on the
66 quality of fossil information, involving age and taxonomic assignment (Parham et al
67 2012), the priors assigned to nodes that are calibrated in a Bayesian analysis
68 (Warnock et al 2012, Brown & Smith 2017), and the amount of information inherent
69 in the molecular dataset (Yang & Rannala 2006, Rannala & Yang 2007, dos Reis &
70 Yang 2013).

71 Fossils inform us of the minimum age of a divergence, imposing a temporal constraint
72 that is widely accepted. However, the constraint of a simple hard minimum age is
73 insufficient information for a proper analysis of times of divergence, particularly as
74 there is an absence of information about maximum ages for divergences, including the
75 root node. Often fossil information is modeled as a probability distribution, such as a
76 lognormal or exponential distribution, indicating our beliefs regarding how
77 informative a fossil is about the age of a divergence (Drummond et al 2006, Warnock
78 et al. 2015). The distributional shapes of these priors are often established without
79 justification (Warnock et al. 2012). Ideally, in node-based dating, fossil information is
80 used only as a minimum age constraint for a given divergence in the form of a
81 uniform prior with a minimum age equaling the fossil age and a maximum age

82 extending beyond the age of the clade in question. In such cases at least one
 83 maximum constraint is needed, often also based on fossil information. Another
 84 approach is use of extraneous additional information, such as using ages of host plant
 85 families as maximum constraints for highly specialized phytophagous insect clades
 86 (Wahlberg et al. 2009). In such cases, a uniform prior also can be used, with the
 87 maximum set to the age of the divergence of the host plant family from its sister
 88 group and the minimum set to the present time.

89 Brown & Smith (2017) recently have pointed out the importance of assessing the
 90 relative influence of priors over the actual amount of information contained in the
 91 molecular dataset. As noted above, users specify fossil calibrations using prior
 92 distributions by modeling the prior expectation about the age of the node constrained.
 93 However, the broader set of fossil constraints can interact with each other and with
 94 the tree prior, leading to marginal prior distributions at nodes that usually differ from
 95 the user's first intention (Warnock et al 2012). If relevant information were contained
 96 within the molecular dataset, one would expect the posterior distribution to shift from
 97 the marginal prior distribution. In the case of angiospermous plants, Brown & Smith
 98 (2017) showed that the marginal prior resulting from the interaction of all priors
 99 (fossils and the tree) excluded an Early Cretaceous origin, in effect giving such an
 100 origin zero probability. In addition, many calibrated internal nodes showed nearly
 101 complete overlap of marginal prior and posterior distributions, suggesting little
 102 information in the molecular dataset but a potentially strong influence of the set of
 103 priors.

104 With more than 18,000 species described and extraordinary efforts made to infer
 105 phylogenetic hypotheses based on molecular data, butterflies (Lepidoptera:

106 Papilionoidea) have become a model system for insect diversification studies.

107 Nevertheless, the paucity of information available to infer times of divergence in

108 butterflies questions the reliability of the various estimates (e.g. Garzón-Orduña et al.

109 2015). Heikkilä et al. (2012) for example, used only three fossils to calibrate a higher-

110 level phylogeny of the superfamily Papilionoidea. The shortage of fossil information

111 for calibrating large-scale phylogenies also means that, most of the time, species-level

112 phylogenies at a smaller scale rely on secondary calibration points extracted from the

113 higher-level time-trees (e.g. Peña et al. 2011, Matos-Maravi et al. 2013, Kozak et al.

114 2015, Chazot et al. 2016, Toussaint & Balke 2016).

115 In a recent paper, de Jong (2017) revisited the butterfly fossil record, providing a

116 discussion about the quality of the different fossil specimens as well as their

117 taxonomic placement. Using this information, we established an unprecedented set of

118 12 fossil calibration points across all butterflies, which we use in this study to revisit

119 the timescale of butterfly evolution in a comprehensive phylogenetic framework, and

120 investigate the robustness of this new estimate. We complement the minimum age

121 constraints of clades based on fossils with maximum age constraints based on the ages

122 of host plant families. Some clades of butterflies have specialized on specific groups

123 of angiosperm hosts for larval development, such that one may assume that

124 diversification of the associated butterfly clade only occurred after the appearance of

125 the host plant clade. We use this assumption as additional information to calibrate the

126 molecular clock by setting the age of specific clades of butterflies to be younger than

127 the estimated age of their host plant lineage. We restrained these calibrations to

128 higher-level host plant clades.

129 The most recent estimates of divergence times using representatives of all butterfly
130 families inferred a crown clade age of butterflies of 110 Ma (Heikkilä et al. 2012) and
131 104 Ma (Wahlberg et al. 2013). These two dates yield to a large discrepancy when
132 taking the fossil record into account, as the oldest known fossil butterfly is estimated
133 to be 55.6 Ma and can be confidently assigned to the extant family HesperIIDae (de
134 Jong 2016, 2017). Such discrepancy has been extensively debated for a similar case,
135 the origin of angiosperms, often estimated to have originated during the Triassic (252-
136 –201 Ma ago), while the oldest undisputed fossil is pollen dated at 136 Ma. Despite a
137 much more fragmentary fossil record for butterflies, the same questions remain. First,
138 are the previous estimates robust to a more comprehensive assemblage of fossils and
139 taxon sampling? Second, is the 50 million-year discrepancy between molecular clock
140 estimates and the fossil record accurate or the result of a lack of information
141 contained in the molecular dataset? In other words, how much does the set of priors
142 influence the results?

143 Here, we generated a genus-level phylogeny of Papilionoidea, including 994 taxa, in
144 order to maximize the number and position of fossil calibration points and increase
145 the potential amount of molecular information. By establishing the set of 12 fossils
146 and 10 host-plant calibration points, we time-calibrated the tree in order to provide a
147 revised estimate of the timing in diversification of butterflies. We then assessed the
148 robustness of these results to the assumptions made throughout the analysis, including
149 (i) different subsets of fossil constraints, (ii) the prior distributions of fossil constraints,
150 (iii) a different estimate for host plant ages, (iv) a Yule tree prior, (v) a reduced taxon
151 sampling and (vi) the addition of a mitochondrial gene fragment to the nine nuclear
152 gene regions.

153 Finally, we compared the user specified priors, marginal prior and posterior
154 distributions of different analyses, to assess the influence of our set of constraints on
155 the estimated timing of divergence.

156

157 MATERIALS AND METHODS

158 *Molecular Dataset*

159 When designing our dataset, we aimed at building a genus-level tree of Papilionoidea.
160 We assembled a dataset of 994 taxa from the database VoSeq
161 (<http://www.nymphalidae.net/db.php>, Peña & Malm 2012), with each taxon
162 representing a genus. We chose to include gene fragments that were available across
163 the whole tree in order to avoid large clade-specific gaps in the molecular dataset. In
164 addition, Sahoo et al. (2016) pointed out a conflicting signal in the family Hesperidae
165 between nuclear and mitochondrial markers. Thus, we chose to primarily focus on
166 nuclear markers. Our final dataset included nine gene fragments: ArgKin (596bp),
167 CAD (850bp), EFI- α (1240 bp), GAPDH (691bp), IDH (710 bp), MDH (733 bp),
168 RPS2 (411 bp), RPS5 (617 bp) and wingless (412 bp) for a total length of 6260 base
169 pairs. The list of taxa and Genbank accession codes are available in the
170 Supplementary Material S1.

171 *Set of Time-Calibrations for Timing Analyses*

172 Fossil calibrations – Previous studies estimating times of divergence of butterfly
173 lineages have largely relied on unverified fossil calibrations. The identifications of
174 these calibrations were often based on overall similarity with extant taxa, not
175 apomorphies. In the present study, we initially chose 14 fossil butterflies that were

176 recently critically reviewed by de Jong (2017) and displayed apomorphic characters
177 or character combinations diagnostic of extant clades, thereby allowing reliable
178 allocation of fossils on the phylogenetic tree to provide minimum ages to the
179 corresponding nodes. These fossils included three inclusions in Dominican Amber
180 and 11 compression/impression fossils. For the age of these fossils we have relied on
181 the most recent dates established from recent advances in Cenozoic
182 chronostratigraphy, geochronology, chemostratigraphy and the geomagnetic polarity
183 time scale (Walker et al., 2013). These improvements by geologists and specialists in
184 allied disciplines have provided an increased precision in age dates of stratigraphic
185 record (International Commission on Stratigraphy, 2012). The list of fossils and their
186 positions in the tree is given in Table 1 and Figure 1. For more detailed information
187 on the identification of these fossils, localities, preservation type and current
188 depositories, see de Jong (2017).

189 When a fossil was assigned to a clade, we calibrated the stem age of this clade,
190 specifically the time of divergence from its sister clade, instead of the crown age or
191 the first divergence event recorded in the phylogeny. As a consequence of this choice,
192 we removed two of the 14 fossils. We did not use *Praepapilio colorado* Durden &
193 Rose, 1978 (Papilionidae, 48.4 Ma) nor the less well-preserved *Praepapilio gracilis*
194 Durden & Rose, 1978 (Papilionidae) of the same age because its position at the root
195 of the tree was uninformative given the presence of the 55.6 million years old
196 *Protoeliades kristenseni* de Jong, 2016 placed at the crown of the Hesperidae. For
197 similar reasons, we did not use *Doxocopa wilmattae* Cockerell, 1907
198 (Nymphalidae+Biblididae+Limenitidae+Apaturidae, 33.8 Ma) because its position
199 was uninformative given the presence of *Vanessa amerindica* Miller & Brown, 1989
200 of the same age but placed lower in the tree.

201 Host plant calibrations – Butterflies are well known for their strict relationships with
 202 specific groups of plants used by their larvae. Such associations have previously been
 203 suggested as evidence for coevolution (Ehrlich & Raven, 1964, Janz & Nylin 1998,
 204 Nylin & Janz 1999). In the present study, we selected nine calibration points based on
 205 known information of host plant specificity by butterflies since the large revision of
 206 Ackery (1988) (see also Beccalloni et al. 2008 for Neotropical species), and revised
 207 for those host plant records listed as having spurious or occasional records (AVLF
 208 unpublished data). Host plant clades used by single genera or a small group of
 209 recently-derived genera were discarded, such as the use of Aristolochiaceae by
 210 Troidini. In these cases the butterflies clearly are much more recent than their
 211 associated plant clades, and consequently do not contribute relevant time information
 212 to the tree. The ages of each plant group were defined as maximum ages for the
 213 respective nodes (Table 1). For all host plant maximum constraints we used the
 214 estimate from Magallón et al. (2015) using the upper boundary of the 95% credibility
 215 interval of the stem age of the host plant clade. We also constrained the root of the
 216 Papilionoidea with a maximum age corresponding to the crown age of angiosperms
 217 from Magallón et al. (2015). The host plant calibrations were placed at the crown of
 218 the butterfly clades as a conservative approach since we do not know when the host
 219 plant shift occurred on the stem branch. However, we assume that the diversification
 220 of the clade could not have begun earlier than the origin of the host plant family.

221

222 *Analyses Overview*

223 Given computational limitations for such a large dataset, we adopted the following
 224 procedure (details given below). We ran PartitionFinder v. 1.1 (Lanfear et al. 2012) to

225 identify the best partition scheme. Using this result we performed a maximum
 226 likelihood analysis to obtain a tree topology. This tree topology was transformed into
 227 a time-calibrated ultrametric tree and used thereafter as a fixed topology and starting
 228 tree in all our dating analyses. Branch lengths were estimated using BEAST v. 1.8.3
 229 (Drummond et al. 2012) with a simpler partitioning scheme, a birth-death tree prior,
 230 lognormal relaxed molecular clocks, and a combination of minimum (fossils) and
 231 maximum (host plants) constraints for which all were set with uniform priors. This
 232 constituted the core analysis. We then performed additional analyses to test the
 233 robustness of our results to (i) different subsets of fossil constraints, (ii) the prior
 234 distribution of fossil constraints, (iii) a different estimate for host-plant ages, (iv) a
 235 Yule tree prior, (v) a reduced taxon sampling, and (vi) the addition of a mitochondrial
 236 gene fragment.

237

238 *Core Analysis*

239 Tree topology – We started by running PartitionFinder v. 1.1 (Lanfear et al. 2012) on
 240 the concatenated dataset, allowing all possible combinations of codon positions of all
 241 genes. Substitution models were restricted to a GTR+G model and branch lengths
 242 were linked. We then performed a maximum likelihood analysis using RAxML v8
 243 (Stamatakis 2006) using the best partitioning scheme identified by PartitionFinder and
 244 1000 ultrafast bootstraps. The resulting tree was set as a fixed topology for the dating
 245 analyses. To do so, the tree was transformed into a time-calibrated ultrametric tree
 246 using the package *ape* (Paradis et al. 2004) and all minimum and maximum
 247 constraints in order to obtain a starting tree suitable for BEAST analyses.

248 Time tree – We used BEAST v. 1.8.3 (Drummond et al. 2012) to perform our time-
 249 calibration analysis. Given the size of our dataset, we reduced the number of
 250 partitions in our dating analysis to three partitions, each partition being one codon
 251 position of all genes pooled together. Substitution rate for each partition was modeled
 252 by GTR+G and an uncorrelated lognormal relaxed molecular clock. We used a Birth-
 253 Death process as branching process prior. In order to have a fixed topology we turned
 254 off the topology operators in BEAUTi and we specified the topology obtained with
 255 RAxML made ultrametric with the *ape* package.

256 Setting the priors for calibration points is always an important matter of discussion.
 257 Non-uniform priors are often used, yet in the majority of studies the choice of
 258 parameters defining the shape of the prior distribution is not justified (Warnock et al.
 259 2012). For the core analysis we followed a conservative approach – considering that
 260 fossils only provide a minimum age, while host plant calibrations only provide a
 261 maximum age for the nodes they were assigned to – and we used uniform prior
 262 distributions for all calibration points (Table 1). When a node was calibrated with
 263 fossil information, the distribution ranged from the estimated age of the fossil to the
 264 age of angiosperm origin (extracted from Magallón et al. 2015). When a node was
 265 calibrated using host-plant age, the prior distribution ranged from 0 (present) to the
 266 age of the host plant clade origin. When a node was calibrated with both types of
 267 information, the distribution ranged for the age of the fossil to the age of host plant
 268 clade origin. We also used a uniform prior for the tree root height, ranging between
 269 the oldest fossil used in the analysis and the age of angiosperm origin. Host plant
 270 calibrations, as well as the origin of angiosperms were extracted from Magallón et al.
 271 (2015), using the upper boundary of the 95% credibility interval of the stem age of the
 272 host plant clade. Our choice of combining (1) uniform prior distributions, (2) fossil

273 calibration of stem nodes, (3) the oldest stem age of the host plant clades and (4) host
274 plant calibration of crown nodes has important implications. On the one hand these
275 choices are the most conservative options, cautiously using the information given by
276 each type of calibration point and taking into account uncertainty surrounding the
277 information used. On the other hand, they are also the least informative.

278 We performed four independent runs of 30 million generations, sampling every 30
279 000 generations. We checked for a satisfactory convergence of the different runs
280 using Tracer v. 1.6.0 (Rambaut et al. 2014) and the effective sample size values in
281 combination. Using LogCombiner v. 1.8.3 (Drummond et al. 2012), we combined the
282 posterior distributions of trees from the three runs, discarding the first 100 trees (10%
283 burn-in) of each run. Using TreeAnnotator v. 1.8.3 (Drummond et al. 2012) we
284 extracted the median and the 95% credibility interval of the posterior distribution of
285 node ages.

286

287 *Alternative Analyses*

288 We tested the effect of making alternative choices along the core analysis on our
289 estimates of divergence times. Unless stated otherwise, we made only one
290 modification at a time; all other parameters remained identical to that described for
291 the core analysis. We performed at least two independent runs of 30 million
292 generations per alternative parameter set and more if convergence was not reached.

293 Different subset of fossils – We aimed at testing whether using only a fraction
294 of the fossil information affected the estimation of divergence times and whether the
295 position of calibrations (close to the root or close to the tips) also changed the results.

296 Thus, we divided our set of fossil constraints into two subsets depending on their
297 position in the tree. One subset included fossil calibration points assigned at a deep
298 level in tree (hereafter: higher-level fossils): *Lethe*, *Mylothrites*, *Neorinella*,
299 *Pamphilites*, *Prolibythea*, *Protochoeliades* and *Vanessa* (Table 1). The other subset
300 included fossil calibration points close to the tips of our phylogeny (hereafter: lower-
301 level fossils): *Doritites*, *Thaites*, *Dynamine*, *Theope* and *Voltinia* (Table 1). In both
302 cases the full set of maximum constraints was used. We performed one analysis for
303 each subset.

304 Exponential fossil priors – In the core analysis we used uniform distributions
305 for calibration points, which is a conservative option but also the least informative. As
306 an alternative, we designed exponential priors for fossil calibration points.
307 Exponential priors use the age of a fossil as minimum age for the node it has been
308 assigned to, but also assume that the probability for the age of the node decreases
309 exponentially as time increases. In BEAUTi, we set the offset of exponential
310 distributions with the age of the fossil. The distribution was truncated at the maximum
311 age used in the uniform priors. The shape of the exponential distribution is controlled
312 by a mean parameter, which has to be arbitrarily chosen by the users. The choice of
313 mean parameter can be found in Table 1. Priors for host plant calibration points were
314 not changed (i.e., uniform priors).

315 Alternative host plant ages –The origin and timing of diversification of
316 angiosperms is controversial. While the oldest undisputed fossil of Angiospermae is
317 from the early Cretaceous (136 Ma, Brenner 1996), most divergence time estimations
318 based on molecular clocks have inferred a much older origin. In the core analysis, we
319 chose to use host plant ages derived from the tree of angiosperms time-calibrated by

320 Magallón et al. (2015), who imposed a constraint on the origin of angiosperms based
 321 on this fossil information. They found a crown age for angiosperms of ~ 140 Ma. As
 322 an alternative consistent with an older origin of angiosperms we used ages recently
 323 inferred by Foster et al. (2017), who recovered a crown age of angiosperms of ~ 209
 324 Ma. All maximum constraints were replaced by those inferred by Foster et al. (2017).
 325 The origin of angiosperms was used as a maximum constraint was set to the upper
 326 boundary of the 95% credibility interval of the crown age of the angiosperms i.e.,
 327 252.8 Ma. Because the posterior distributions of node ages for this analysis were very
 328 skewed, we extracted the median of the distribution, the 95 % credibility interval and
 329 the mode of the kernel density estimate of nodes using the R package *hdrnde*. For
 330 comparison, we also estimated the mode of posterior distributions for the core
 331 analysis and all alternative tests.

332 Yule branching process prior – Condamine et al. (2015) showed that the prior
 333 for the tree growth can have a great impact on the estimated divergence times. In the
 334 core analysis we used a Birth–Death prior, which models the tree formation with a
 335 constant rate of lineage speciation and a constant rate of lineage extinction. As an
 336 alternative, we used a Yule prior, which involved a constant rate of speciation and no
 337 extinction to assess whether age estimates changed or not.

338 Reduced dataset – In our core analysis, we chose to maximize the taxon
 339 sampling – increasing the number of lineages – which increased the fraction of
 340 missing data in the molecular dataset. We tested whether increasing the molecular
 341 dataset completion to the detriment of taxon sampling changed the results. In this
 342 reduced dataset, we included all the genera for which a specific minimum number of
 343 genes were available. The missing data in the molecular dataset are not uniformly

distributed across the tree; for example, Lycaenidae have more missing data than the Nymphalidae. Therefore, a different cut-off value was chosen for each family in order to keep a good representation of the major groups (Papilionidae: 5 genes, Hedyliidae: 8 genes, Hesperidae: 9 genes, Pieridae: 8 genes, Lycaenidae: 4 genes, Riodinidae: 8 genes, Nymphalidae: 9 genes). In order to allow assignment of all fossils to the same place as in the core analysis, nine taxa having a number of genes below the cut-off value had to be added. We ended up with a dataset reduced to only 364 taxa instead of 994 in the core analysis. Accordingly, the fraction of missing data decreased from 39.5% in the core analysis to 21.4%. Given this important modification of the dataset we generated a new topology with RAxML, which was then calibrated identically to the core analysis.

Mitochondrial gene fragment – We tested whether adding mitochondrial information in the dataset would affect our results. To do so, we added the cytochrome-oxidase-subunit 1 gene to the molecular dataset. Given the conflicting signal in Hesperidae between nuclear and mitochondrial information (Sahoo et al. 2016), the COI was not added to the Hesperidae. We performed a new RAxML analysis in order to obtain a new topology. This new tree was calibrated with BEAST identically to the core analysis, with one difference. The mitochondrial gene was added as two partitions separated from the nuclear partitions: the first and second positions of COI were pulled together and the third position had its own partition. Therefore this analysis had five partitions.

Comparing Prior and Posterior Distributions

367 When performing a Bayesian analysis, comparing prior and posterior parameter
368 distributions can be informative about the amount of information contained by our
369 data compared to the influence of prior information. As exemplified by Brown &
370 Smith (2017), such a comparison can shed light on the discrepancies observed in the
371 fossil record and the divergence times estimated from a time-calibrated molecular
372 clock. It may also help to disentangle the effect of interaction among calibration
373 points. For each calibrated node we can compare the user-designed prior distribution
374 (e.g., uniform distributions in the case of the core analysis), the marginal prior
375 distribution that is the result of the interaction between the user priors and the tree
376 prior, and the posterior distribution that is the distribution after observing the data.

377 For the core analysis, the two different subsets of fossils and the alternative host plant
378 ages analyses were re-run without any data to sample from the marginal prior. In each
379 case we performed two independent runs of 50 million generations, sampling every
380 50 000 generations. The results were visualized with Tracer. When necessary, we
381 performed an additional run. Using LogCombiner, the runs were combined after
382 deleting the first 10% as burn-in. The results of the analyses with and without the
383 molecular dataset were imported into R (R Development Core Team 2008) and for
384 each calibrated node as well as the root height we compared the kernel density
385 estimates of the marginal prior and the posterior distributions (R package *hdcrcde*).

386 *Comparison with Previous Studies*

387 For the root of all Papilionoidea and the seven families we compared the estimates
388 obtained in the core analysis to previous studies that also used fossil information.

389

390 RESULTS

391 *Core Analysis*

392 The core analysis performed with BEAST used the full set of fossils and host plant
 393 constraints from Magallón et al. (2015) on the topology found with RAxML. This
 394 analysis resulted in a root estimate for all Papilionoidea of 107.6 Ma (Fig.1,
 395 Supplementary Material S2). The 95% credibility interval of the posterior distribution
 396 ranged from 88.5 to 129.5 Ma. The lineage leading to Papilionidae diverged first at
 397 the root of Papilionoidea and the crown age of Papilionidae was inferred to be 68.4
 398 Ma (95%CI=53.5–84.3). Hedyliidae and Hesperiiidae diverged from Pieridae–
 399 Lycaenidae–Riodinidae–Nymphalidae at 106.5 Ma (95%CI=88.0–127.2) and
 400 diverged from each other at 99.2 Ma (95%CI=80.7–119.2). The crown age of the
 401 sampled Hedyliidae was 32.8 Ma (95%CI=23.4–43.6) and crown age of Hesperiiidae
 402 was 65.2 Ma (95%CI=55.8–78.1). Pieridae diverged from Lycaenidae–Riodinidae–
 403 Nymphalidae at 101.1 Ma (95%CI=83.0–120.3) and extant lineages started
 404 diversifying around 76.9 Ma (95%CI=63.1–92.4). Lycaenidae and Riodinidae
 405 diverged from Nymphalidae at 97.4 Ma (95%CI=80.4–116.5) and diverged from each
 406 other at 87.8 Ma (95%CI=73.2–106.1). The crown age of Lycaenidae was 71.0 Ma
 407 (95%CI=57.2–85.2) and crown age of Riodinidae was 73.4 Ma (95%CI=60.3–88.1).
 408 Finally, the crown age of Nymphalidae was inferred to be 82.0 Ma (95%CI=68.1–
 409 98.3).

410

411 *Alternative Analyses*

412 In most cases the seven alternative parameters tested yielded very similar results (Fig.
 413 2, Supplementary Material S3-S8). Reducing the number of taxa in order to decrease
 414 the fraction of missing data, using higher-level calibration points only, or using a Yule
 415 process tree prior (instead of a Birth–Death prior), gave virtually identical results as
 416 the core analysis above. Using only lower-level fossil constraints (close to the tips of
 417 the phylogeny) resulted in the youngest estimates of all alternative runs, with a crown
 418 age of Papilionoidea of 94.5 Ma (mode=83.8, 95%CI=67.8–126.6). Using exponential
 419 fossil priors mainly resulted in a narrower credibility interval, while the mode and
 420 median age estimates were only 7–8 million years younger than the core analysis
 421 mode estimate (Fig. 2, Supplementary Material S6). Adding mitochondrial
 422 information also lead to a 7–8 million-year younger estimate for the crown age of
 423 Papilionoidea, but the credibility interval remained comparable to the core analysis
 424 (Supplementary Material S7). Finally, using a hypothesis of older host plant ages
 425 extracted from Foster et al. (2017), we obtained the greatest difference. The upper
 426 boundary of the credibility interval largely shifted toward much older ages (95%CI=88.5–167.2) as well as the median (119.5 Ma). The posterior distribution was,
 427 however, very skewed, with a mode of 101.0 Ma, and converged to the same age as
 428 the core analysis (Fig. 2, Supplementary Material S8).

430 These variations for the root age among different alternative analyses were recovered
 431 for the ages of the different subfamilies. For example, all lower-level fossils always
 432 led to younger estimates while older ages from Foster et al. (2017) always led to older
 433 estimates (Fig. 2).

434 *Comparing Prior and Posterior Distributions*

435 We compared the posterior distributions to the marginal prior distributions for the
436 different calibrated nodes in the core analysis. We set all fossil and host plant
437 constraints with uniform prior distributions as we considered this as the most
438 conservative approach. However, it is important to note that the marginal prior
439 distributions at these nodes, which result from the interactions between all calibration
440 priors and tree prior, are not uniform (Fig. 3).

441 Across all calibrated node points, many of them showed shifts of posterior
442 distributions from the marginal priors, indicating that the results of the core analysis
443 was not a simple outcome of our set of priors (Fig. 3). Interestingly, the nodes
444 calibrated by *Doritites*, *Dynamine*, *Thaites*, *Theope* and *Voltinia*, which are all the
445 fossils placed close to the tips of our phylogeny, tended to shift away from the
446 minimum boundary, toward older ages than the marginal prior distribution.
447 Alternative analyses performed with only these lower-level fossils yielded the
448 youngest tree for butterflies. This suggests that higher-level fossils bring important
449 additional information, leading posterior distributions of lower-level nodes to shift
450 away from the prior distributions in the core analysis.

451 The nodes calibrated with the higher-level fossils *Mylothrites*, *Prolibythea*, *Neorinella*
452 and *Vanessa* showed posterior distributions largely overlapping with their marginal
453 prior distributions. Many host plant calibrated points showed a shift from the marginal
454 prior distribution (Fig. 3). In all cases, except the node also calibrated with the fossil
455 *Lethe* the crown age of the butterfly clade inferred was much younger than the age of
456 the corresponding host plant clade.

457 For the root of Papilionoidea, the marginal prior and posterior distributions largely
458 overlapped in the core analysis, therefore not indicating whether our molecular

dataset contained significant information about the root age or not. We also compared the posterior and the marginal prior distributions for alternative analyses performed with different subsets of fossil calibrations (Fig. 4). When using only higher-level fossils, the posterior distribution was almost identical to the core analysis, but the marginal prior slightly shifted from the marginal prior of the core analysis toward a younger age. The use of only lower-level fossils had more profound effects. In such a case, prior distributions of the core analysis and the lower-level fossils alternative completely overlapped. The posterior distribution, however, shifted toward younger ages, yielding the most recent estimate for the root age among all analyses (mean=94.5, mode=83.8, 95%CI=67.8–126.5). We also looked at the effect of using relaxed maximum ages (based on Foster et al. 2017). In this case, marginal prior distribution for the root age shifted to a mean of ~148 Ma (Fig. 4) and a credibility interval spanning 100 Ma (95%CI=99.9–205.8). The posterior distribution was very skewed, retaining a wider credibility interval than the core analysis (95%CI=88.5–167.5), but significantly shifted from the prior distribution toward the posterior distribution of the core analysis (median=119.5, mode=101.0).

475 *Comparison with Previous Studies*

For the root of Papilionoidea, our estimate in the core analysis using the mode age of the distribution was very similar to Wahlberg et al. (2013) and Heikkilä et al. (2012), with a mean age estimate of 104.6 and 110.8 Ma, respectively (107.6 Ma in the core analysis, Fig. 5). For the crown age of families our estimates were often consistent with most of previous studies. For Papilionidae, our crown age estimate (68.4, 95%CI=53.5–84.3) was very similar to Wahlberg et al. (2013) and Heikkilä et al. (2012), while Condamine et al. (2012) in a study focusing primarily on this family

found younger ages of about 15 million years. For Hedylidae, only the study by Heikkilä et al. (2012) had an estimate for the crown age, whose mean age was 45.3 Ma, which is older than our result (32.8, 95%CI=23.4–43.6). The age of Hesperidae (65.2, 95%CI=55.8–78.1) was similar to Wahlberg et al. (2013) and Heikkilä et al. (2012), but much younger than Sahoo et al. (2017) with an estimate of 82 Ma. Pieridae is the family that showed highest variation in age estimates among different studies. Our estimate (76.9 Ma, 95%CI=63.1–92.4 Ma) falls in between the youngest estimate from Wahlberg et al. (2013), in which the credibility interval goes down to 39 Ma, and the oldest estimate from Braby et al. (2006), in which the oldest boundary of the credibility interval was 111.6 Ma. For Lycaenidae, which contain no fossils calibrations, the results between our core analysis (73.4, 95%CI=60.3–88.1), Wahlberg et al. (2013) and Heikkilä et al. (2012) were virtually identical. For the crown age of Riodinidae, our core analysis (70.9, 95%CI=57.2–85.2) gave identical results to Heikkilä et al. (2012). Espeland et al. (2015), in a study focusing specifically on this family found about 10 million-year older ages. Wahlberg et al. (2013), however, found a much younger estimate, about 20 Ma younger. For Nymphalidae, we have the greatest number of time calibrations, but they all tend to find very similar results. Our estimation (82.0, 95%CI=68.1–98.3) was very close to Wahlberg et al. (2013) and Heikkilä et al. (2012). This estimation was about 12 million years younger than the study by Wahlberg et al. (2009) focusing on Nymphalidae.

DISCUSSION

We generated a genus-level phylogeny of the superfamily Papilionoidea, including 994 taxa. Taking advantage of a recent revision of the lepidopteran fossil record we

507 established a new set of 12 fossil calibration points, which were combined to 10
508 secondary calibrations from host plant ages.

509 *Fossils and Minimum Ages*

510 In the core analysis we adopted a very conservative approach. This choice involves
511 taking into account the uncertainty surrounding the information available for each
512 calibration point, although at the expense of the amount of useful information
513 available. For fossil constraints, this had two consequences. First, we calibrated the
514 stem of the focal clade a fossil was assigned to, by calibrating the divergence from its
515 sister group instead of the first divergence recorded in the phylogeny within the focal
516 clade itself. Calibrating the crown age of the focal clade – meaning that we assume
517 that the fossil is “nested” within the clade – may lead to an overestimation of the
518 crown age. Such would be the case if lineages are undersampled at the root, or if
519 extinction occurred, or if the fossil belongs to a lineage that actually diverges
520 somewhere along the stem. Calibrating a higher node with the age of the fossil, which
521 involves loss of some information, is considered to be the best way to avoid these
522 problems. Second, we used uniform prior distributions bounded by the age of the
523 fossil and the age of angiosperms. We considered that fossils provide only a minimum
524 age for a node, a condition that is especially exacerbated by the exceptionally poor
525 fossil record of Lepidoptera in general (Labandeira and Sepkoski, 1993) and
526 Papilionoidea in particular (Sohn et al. 2015) when compared to the four other major
527 hyperdiverse insect lineages (Coleoptera, Hymenoptera, Diptera and Hemiptera).
528 Prior expectation on the age of the node cannot be modeled more accurately without
529 additional information. However, the marginal priors resulting from the interactions
530 among the different priors actually strongly differ from this assumption.

531

532 *Higher- versus Lower-Level Calibrations*

533 Generally, favoring multiple calibrations placed at various positions in a tree instead
 534 of a single or few calibrations, seem to produce more reliable estimates of molecular
 535 clocks (Conroy & Van Tuinen 2003, Smith & Peterson 2002, Soltis et al. 2002,
 536 Duchêne et al. 2014). Calibrations distributed across a tree may allow a better
 537 estimation of substitution rates and their pattern of variation among lineages
 538 (Duchêne et al. 2014), and improve age estimates in cases of taxon undersampling
 539 (Linder et al. 2005).

540 Calibrations placed at deep levels in the tree are usually favored (Sauquet 2012, Hug
 541 & Roger 2007) over calibrations at lower levels for better capturing the overall
 542 genetic variation (Duchêne et al. 2014). Yet, deep calibrations also tend to
 543 underestimate the mean substitution rate and lead to an overestimation of shallow
 544 nodes, referred to as “tree extension” by Phillips (2009). For the butterflies, we
 545 investigated the consequences of using different subsets of fossil calibrations
 546 according to their positions in the tree (higher *versus* lower-level calibrations),
 547 compared to the full set of fossil constraints. With a subset of fossils placed only at
 548 higher levels in the phylogeny we obtained results similar to the full set of fossils in
 549 the core analysis, either at the deep nodes or shallow nodes, indicating no tree
 550 extension effect. This effect may also indicate that the lower level calibration points
 551 that are close to the tips are uninformative, and when included in the core analysis, do
 552 not affect the timescale but clearly affected the priors (see below).

553 Alternatively, lower-level calibrations can lead to an overestimation of the mean
 554 substitution rate across the tree, thereby underestimating the timescale (Phillips 2009).

555 Interestingly, when only a subset of fossils were used and placed close to the tips, it
556 led to the youngest estimates, including the credibility intervals. This potentially
557 indicates an effect of mean substitution rate overestimation. Also, we noticed in the
558 core analysis that the nodes calibrated by *Protoceiliades* and *Vanessa* (two deep node
559 constraints) showed posterior distributions abutting against the minimum boundaries
560 defined by the age of the fossils, therefore preventing the tree (or at least these nodes)
561 to be younger.

562

563 *Host Plants and Maximum Ages*

564 For calibration points constrained by the age of the host plant group, we considered
565 that only the crown of the focal clade could be assigned confidently to the host plant
566 group, as the stem or part of the stem could be older than the host plant (the host plant
567 shift would be happening somewhere along the stem). Support arises from molecular
568 biological and paleobiological evidence that the establishment of specialized insect-
569 herbivore associations can considerably postdate the origins of their hosts, as in a
570 Bayesian analysis of 100 species of leaf-mining *Phyllonorycter* moths (Lepidoptera:
571 Gracillariidae) and their dicot angiosperm hosts (Lopez-Vaamonde et al. 2006).
572 Relying on host plant ages for calibrating a butterfly tree is questionable while the
573 timing of the divergence of angiosperms is still highly controversial (e.g. Magallón et
574 al. 2015, Foster et al. 2017). As a result, first we calibrated our tree using the oldest
575 boundary of 95% CI of the stem age of a host plant clade. This allowed us to take into
576 account the uncertainty surrounding the timing of the first appearance of the host
577 plant but consequently, it also relaxed the prior hypothesis for the calibrations.
578 Secondly, we compared two alternative timescales for the angiosperms: a

579 paleontological estimate, which infers an Early Cretaceous origin of angiosperms
580 (Magallón et al. 2015), and a molecular clock estimate that we extracted from Foster
581 et al. (2017), which infers a stem age for angiosperms during the Early Triassic about
582 100 million years older. These two alternative scenarios affected the size of the
583 credibility intervals and the shape of the posterior distributions. For the crown of
584 Papilionoidea, the upper boundary of the 95%CI was ~37 million years older when
585 using the molecular clock estimate. However, the shape of the distribution was very
586 asymmetrical, with a mode of the distribution very close to the core analysis (101.0
587 Ma), suggesting that the estimation for the age of the root still concentrated around
588 the same ages. Using the hypothesis of an Early Triassic origin of angiosperms
589 implied very permissive priors toward old ages, which are most likely responsible for
590 the very wide credibility intervals and asymmetrical posterior distributions recovered
591 in the alternative analysis of using ages from Foster et al. (2017). Therefore, it is
592 tempting to use the time-scale inferred using Magallón et al. (2015)'s ages of
593 angiosperms, as it greatly narrows down the uncertainty surrounding butterfly ages,
594 and aligns more realistically with the fossil angiosperm record. However, as long as
595 there is no consensus on the timing of angiosperm diversification there is no reason to
596 favor one or the other.

597

598 *Priors and Posterior Distributions*

599 We compared the marginal priors to the posterior distributions for different analyses
600 for the root of Papilionoidea and for the different calibration points in the core
601 analysis. We found several calibration points showing a substantial shift of posterior
602 distribution. This indicates that our age estimates are not entirely driven by the set of

603 constraints, but instead that the molecular dataset brings additional information about
604 the age of the calibrated nodes. An interesting pattern we found in the core analysis is
605 the consistent trend of posterior distributions of the lower-level calibrated nodes to
606 shift toward older ages than the priors. Meanwhile, some higher-level node
607 calibrations shifted toward younger ages than the prior but most of them largely
608 overlapped with their prior distribution. Consequently, posterior estimates tend to
609 contract the middle part of tree compared to the prior estimates.

610 There are at least three reasons for the anomalous gap between the earliest fossil
611 papilionoid occurring at 55.6 Ma and its corresponding Bayesian median age of 110
612 Ma, that represents a doubling of the lineage duration. First, it long has been known
613 that the lepidopteran fossil record is extremely poor when compared to the far more
614 densely and abundantly occurring fossils of the four other hyperdiverse, major insect
615 lineages of Hemiptera, Coleoptera, Diptera and Hymenoptera (Labandeira and
616 Sepkoski, 1993). Second, particularly large-bodied apoditrysians such as
617 Papilionoidea, have even a poorer fossil record than other Lepidoptera in general,
618 particularly as they bear a fragile body habitus not amenable to preservation.
619 Additionally, as external feeders papilionoids lack a distinctive, identifiable trace
620 fossil record such as leaf mines, galls and cases (Sohn et al. 2015). Third, there are
621 very few productive terrestrial compression or amber deposits spanning the Upper
622 Cretaceous, from 100 Ma to the Cretaceous–Paleogene boundary of 66.0 Ma, and the
623 part of the Paleogene Period from 66.0 Ma to the earliest papilionoid fossil of 55.6
624 Ma (Labandeira, 2014; Sohn et al., 2015). Some of these deposits have recorded very
625 rare small moth fossils, but to date no papilionoid, or for that matter, other large
626 lepidopteran taxa such as saturniids or pyraloids have been found.

627 The root of the tree was only calibrated with the oldest fossil in our dataset, a 55.6
628 million-year-old papilionoid, and the crown age of the angiosperms. However, the
629 prior distribution for the root in the core analysis clearly excluded an origin of
630 butterflies close to 55.6 Ma, but rather a distribution centered on a median of 110 and
631 ranging between 86.4 and 136.2 Ma. The posterior distribution for the root in the core
632 analysis largely overlapped with the prior. However, when we used alternative ages
633 for the angiosperms (older ages), the marginal prior for the root shifted to
634 substantially older ages. Nevertheless, the posterior distribution showed a significant
635 shift toward younger ages, albeit highly skewed, toward ages similar to the core
636 analysis. This suggests that our estimate of the root age in the core analysis is not
637 simply driven by our set of priors, even if we do not actually observe a shift between
638 marginal prior and posterior distributions.

639 We observed some differences in prior and posterior distributions at the root when
640 considering only subsets of fossils. When using only the subset of higher-level fossils,
641 the marginal prior for the root showed very little difference from the core analysis
642 prior and the posterior distributions completely overlapped. When using the subset of
643 lower-level fossils the marginal prior remained similar to the core analysis but the
644 posterior distribution showed a substantial shift toward younger ages, yielding the
645 youngest estimation of the age of Papilionoidea among all our analyses. As such, it
646 seems that the choice of fossils did not change the prior estimation of the root, but the
647 posterior distribution was largely influenced by higher-level fossils. As we suggested
648 earlier, lower-level fossils only may be overestimating the mean substitution rate
649 across the tree, and therefore underestimating the time scale, while the
650 implementation of higher-level fossils seems to be correcting for this.

651 *Timescale of Butterflies Revisited*

652 We propose a new estimate for the timing of diversification of butterflies, based on an
653 unprecedented set of fossil and host-plant calibrations. We estimated the origin of
654 butterflies between 89.5 and 129.5 Ma, the median of this posterior distribution being
655 107.6 Ma, which corresponds to the Early Cretaceous–Late Cretaceous boundary
656 interval. The result of our core analysis for the root is very close to previous estimates
657 by Wahlberg et al. (2013) and Heikkilä et al. (2012). The comparisons of alternative
658 analyses, the prior and posterior distributions showed that this result is robust to
659 almost all the choices made throughout the core analysis and that our molecular
660 dataset contains significant information in addition to the time constraints. This
661 estimation means that there is a 45 million-year-long gap between the oldest known
662 butterfly fossil and the molecular clock estimate. Accordingly, as Brown & Smith
663 (2017) stated for the case of angiosperms, we do not know whether a larger molecular
664 dataset – implying potentially more information for estimating the molecular clock–
665 would allow the root to become younger. Alternatively, the fossil record for
666 butterflies is so sparse that an intervening fossil gap is very likely. Besides, the fossil
667 *Protocoeliades kristenseni*, which is 55.6 Ma can be assigned confidently to the
668 crown of the family Hesperidae and the stem of Coeliadinae well within the
669 Papilionoidea. For angiosperms, a very rich fossil record is available compared to
670 butterflies (e.g., Magallón et al. (2015), which used 137 fossils to calibrate a
671 phylogeny of angiosperms), rendering the absence of angiosperms, either as pollen or
672 macrofossils, that are older than 136 Ma much more puzzling.

673

674 **ACKNOWLEDGMENTS**

675 This is contribution 366 of the Evolution of Terrestrial Ecosystems consortium at the
 676 National Museum of Natural History, in Washington D.C. NW acknowledges funding
 677 form the Swedish Research Council (Grant No. 2015-04441) and from the
 678 Department of Biology, Lund University. AVLF thanks CNPq (grant 303834/2015-3),
 679 National Science Foundation (DEB-1256742) and FAPESP (grant 2011/50225-3).
 680 This publication is part of the RedeLep (Rede Nacional de Pesquisa e Conservação de
 681 Lepidópteros) SISBIOTABrasil/CNPq (563332/2010-7). MH gratefully
 682 acknowledges funding from a Peter Buck Postdoctoral Stipend, Smithsonian
 683 Institution National Museum of Natural History.

684

685 **SUPPORTING INFORMATION**

686 S1. List of taxa and Genbank accession codes.

687 S2. Tree obtained from the core analysis. Node ages are the median of node age
 688 posterior distributions.

689 S3. Tree obtained from the reduced dataset. Node ages are the median of node age
 690 posterior distributions.

691 S4. Tree obtained when using only higher-level fossil calibrations. Node ages are the
 692 median of node age posterior distributions.

693 S5. Tree obtained when using only lower-level fossil calibrations. Node ages are the
 694 median of node age posterior distributions.

695 S6. Tree obtained when using exponential fossil calibration priors. Node ages are the
 696 median of node age posterior distributions.

697 S7. Tree obtained when adding a mitochondrial gene fragment. Node ages are the
698 median of node age posterior distributions.

699 S8. Tree obtained when using the host-plant ages obtained from Foster et al. (2017).
700 In S8a node ages are the median of node age posterior distributions, while in S8b the
701 node ages are the mode the mode of the kernel density estimate of the posterior
702 distribution.

703

704 **LITERATURE CITED**

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

708 Brenner, G.J. 1996. Evidence for the earliest stage of angiosperm pollen evolution: a
709 paleoequatorial section from Israel. In: Taylor, D.W., Hickey, L.J., editors, Flowering
710 plant origin, evolution & phylogeny. Boston, MA: Springer. P. 91–115.

711 Brown, F.M. 1976. †*Oligodonta florissantensis*, gen. n., sp. nov. (Lepidoptera:
712 Pieridae). Bull. Allyn Mus., 37:1–4.

713 Brown, J.W., and Smith, S.A. 2017. The past sure is tense: On interpreting
714 phylogenetic divergence time estimates. Syst. Biol., 66, doi: 10.1093/sysbio/syx074

715 Chazot, N., Wilmott, K.R., Freitas, A.V.L., de Silva, D.L., Pellens, R., Elias, M. 2016.
716 Patterns of species, phylogenetic and mimicry diversity of clearwing butterflies in the
717 Neotropics. In: Pellens, R., Grandcolas, P., editors, Biodiversity conservation and
718 phylogenetic systematics. Topics in Biodiversity and Conservation, 14:333–353

719 Cockerell, T.D.A. 1907. A fossil butterfly of the genus *Chlorippe*. *Canad. Entomol.*,
720 39:361–363,

721 Condamine, F.L., Toussaint, E.F.A., Clamens, A.L., Genson, G., Sperling, F.A.,
722 Kergoat, G.J. 2015. Deciphering the evolution of birdwing butterflies 150 years after
723 Alfred Russel Wallace. *Sci. Rept.*, 5:11860.

724 Conroy, C.J., and van Tuinen, M. 2003. Extracting time from phylogenies: positive
725 interplay between fossil and genetic data. *J. Mammal.* 84:444–455.

726 de Jong, R. 2016. Reconstructing a 55-million-year-old butterfly (Lepidoptera:
727 Hesperiiidae). *Eur. J. Entomol.* 113:423–428.

728 de Jong, R. 2017. Fossil butterflies, calibration points and the molecular clock
729 (Lepidoptera: Papilionoidea). *Zootaxa* 4270:1–63.

730 dos Reis, M. and Yang, Z., 2013. The unbearable uncertainty of Bayesian divergence
731 time estimation. *J. Syst. Evol.* 51:30–43.

732 Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A. 2012. Bayesian phylogenetics
733 with BEAUTi and BEAST 1.7. *Mol. Biol. Evol.* 29:1969–1973.

734 Duchêne, S., Lanfear, Ho, S.Y.W., 2014. The impact of calibration and clock-model
735 choice on molecular estimates of divergence times. *Mol. Phylogen. Evol.*, 78:277–289.

736 Durden, C.J., and Rose, H. 1978. Butterflies from the Middle Eocene: the earliest
737 occurrence of fossil Papilionoidea (Lepidoptera). *Pearce-Sellards Ser.* 29:1–25,

738 Foster, C.S.P., Sauquet H., van der Merwe, M., McPherson, H., Rosette, M., Ho,
739 S.Y.W. 2017. Evaluating the impact of genomic data and priors on Bayesian
740 estimates of the angiosperm evolutionary timescale. *Syst. Biol.*, 66:338–351.

741 Garzón-Orduña, I.J., Silva-Brandão, K.L., Wilmott, K.R., Freitas, A.V.L., Brower,
742 A.V.Z. 2015. Incompatible ages for clearwing butterflies based on alternative
743 secondary calibrations. *Syst. Biol.*, 64:752–767.

744 Hall, J.P.W., Robins, R.K., Harvey, D.J. 2004. Extinction and biogeography in the
745 Caribbean: new evidence from a fossil riodinid butterfly in Dominican amber. *Proc. R.*
746 *Soc. Lond. B*, 271:797–801.

747 Heer, O. 1849. Die Insektenfauna der Tertiärgelände von Oeningen und von Radoboj
748 in Croatien. Vol. 2. Wilhelm Engelmann, Leipzig, 264 pp.

749 Heikkilä, M., Kaila, L., Mutanen, M., Peña, C., Wahlberg, N. 2012, Cretaceous origin
750 and repeated Tertiary diversification of the redefined butterflies. *Proc. R. Soc. Lond.*
751 *B*, 279:1093–1099.

752 Ho, S.Y., and Phillips, M.J. 2009. Accounting for calibration uncertainty in
753 phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58: 367–380.

754 Hug, L.A., and Roger, A.J. 2007. The impact of fossils and taxon sampling on ancient
755 molecular dating analyses. *Mol. Biol. Evol.* 24:1889–1897.

756 International Commission on Stratigraphy. 2012. International Chronostratigraphic
757 Chart: <http://www.stratigraphy.org/ICSchart/ChronostratChart2012.pdf>

758 Kozak, K.M., Wahlberg, N., Neild, A.F.E., Dasmahapatra, K.J.K., Jiggins, C.D. 2015.
759 Multilocus species trees show the recent adaptive radiation of the mimetic *Heliconius*
760 butterflies. *Syst. Biol.*, 64:505–524,

761 Labandeira, C.C. 2014. Amber. In: Laflamme, M., Schiffbauer, J.D., Darroch, S.A.F.,
762 editors, Reading and writing of the fossil record: Preservational pathways to
763 exceptional fossilization. *Paleontol. Soc. Pap.*, 20:163–216.

764 Labandeira, C.C. and Sepkoski, J.J., Jr. 1993. Insect diversity in the fossil record.
765 *Science*, 261:310–315.

766 Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S. 2012. PartitionFinder: combined
767 selection of partitioning schemes and substitution models for phylogenetic analysis.
768 *Mol. Biol. Evol.* 29:1695–1701.

769 Linder, H.P., Hardy, C.R., Rutschmann, F. 2005. Taxon sampling effects in molecular
770 clock dating: An example from the African Restionaceae. *Mol. Phylogen. Evol.*
771 35:569–582.

772 Lopez-Vaamonde, C., Wikström, N., Labandeira, C., Godfray, H.C.J., Goodman, S.J.,
773 Cook, J.M. 2006. Fossil-calibrated molecular phylogenies reveal that leaf-mining
774 moths radiated millions of years after their host plants. *J. Evol. Biol.*, 19:1314–1326,

775 Magallón, S. Gómez-Acevedo, S., Sánchez-Reyes, L.L., Hernández-Hernández, T.
776 2015. A metacalibrated time-tree documents the early rise of flowering plant
777 phylogenetic diversity. *New Phytol.*, 207:437–453.

778 Martins-Neto, R.G., Kucera-Santos, J.C., Vieira, F.R. de M., Fragoso, L.M. de C.
779 1993. Nova espécie de borboleta (Lepidoptera: Nymphalidae: Satyrinae) da Formação
780 Tremembé, Oligoceno do Estado de São Paulo. Acta Geol. Leopold., 37:5–16.

781 Matos-Maravi, P.F., Peña, c., Wilmott, K.R., Freitas, A.V.L., Whalberg, N. 2013.
782 Systematics and evolutionary history of butterflies in the “*Taygetis* clade”
783 (Nymphalidae: Satyrinae: Euptychiina): towards a better understanding of
784 Neotropical biogeography. Mol. Phylogenet. Evol. 66:54–68.

785 Miller, J.Y. and Brown, F.M. 1989. A new Oligocene fossil butterfly, *Vanessa*
786 †*amerindica* (Lepidoptera: Nymphalidae), from the Florissant Formation, Colorado.
787 Bull. Allyn Mus., 126:1–9.

788 Nel, A. and Descimon, H. 1984. Une nouvelle espèce de Lépidoptère fossile du
789 Stampien de Cereste (Lepidoptera Satyridae). Géol. Méditerran. 11:287–293.

790 Nel, A., Nel, J, Balme, C. 1993. Un nouveau Lépidoptère Satiyrinae fossile de
791 l’Oligocene du Sud-Est de la France (Insecta, Lepidoptera, Nymphalidae). Linn. Belg.,
792 14:20–36.

793 Paradis, E., Claude, J., & Strimmer, K. 2004. APE: analyses of phylogenetics and
794 evolution in R language. *Bioinformatics*, 20(2), 289-290.

795 Parham, J.F., Donoghue, P., Bell, C.J., Calway, T.D., Head, J.J., Holroyd, P.A., Inoue,
796 J.G., Irmis, R.B., Joyce, W.G., Ksepka, D.T., Patané, J.S.L., Smith, N.D., Tarver, J.E.,
797 van Tuinen, M., Yang, Z., Angielczyk, K.D., Greenwood, J.M., Hipsley, C.A., Jacobs,
798 L., Mackovicky, P.J., Miller, J., Smith, K.T., Theodor, J.M., Warnock, R.C.M.,
799 Benton, M.J. 2012. Best practices for justifying fossil calibrations. Syst. Biol.,
800 61:346–359.

801 Peña, C., and Malm, T. 2012. VoSeq: A voucher and DNA sequence web application.
802 PLoS ONE, 7(6):e39071.

803 Peñalver, E., and Grimaldi, D.A. 2006. New data on Miiocene butterflies in
804 Dominican Amber (Lepidoptera: Riodinidae and Nymphalidae) with the description
805 of a new nymphalid. Am. Mus. Novit. 3591:1–17.

806 Phillips, M.J. 2009. Branch-length estimation bias misleads molecular dating for a
807 vertebrate mitochondrial phylogeny. Gene, 441:132–140.

808 R Development Core Team (2008). R: A language and environment for statistical
809 computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-
810 900051-07-0, URL <http://www.R-project.org>.

811 Rambaut, A., Suchard, M. A., Xie, D., & Drummond, A. J. 2014. Tracer v1.6,
812 Available from <http://tree.bio.ed.ac.uk/software/tracer/>.

813 Rannala, B. and Yang, Z. 2007. Inferring speciation times under an episodic
814 molecular clock. Syst. Biol., 56:453–466.

815 Rebel, H. 1898. Fossile Lepidopteren aus der Miocän-Formation von Gabbro.
816 Sitzungsber. Akad. Wiss. Wien, 107:731–745.

817 Sahoo, R. K., Warren, A. D., Collins, S. C., & Kodandaramaiah, U. 2017. Hostplant
818 change and paleoclimatic events explain diversification shifts in skipper butterflies
819 (Family: Hesperiidae). BMC Evolutionary Biology, 17: 174. Doi: 10.1186/s12862-
820 017-1016-x

821 Sahoo, R.K., Warren, A.D., Wahlberg, N., Brower, A.V.Z., Lukhtanov, V.A.,
822 Kodandaramaiah, U. 2016. Ten genes and two topologies: an exploration of higher

relationships in skipper butterflies (Hesperiidae). Peer J, 4:e2653, doi:
10.7717/peerj.2653

Sauquet, H., Ho, S.Y.W., Gandolfo, M.A., Jordan, G.I., Wilf, P., Cantrill, D.J., Bayly, M.J., Bromham, L., Brown, G.K., Carpenter, R.J., Lee, D.M., Murphy, D.J., Sniderman, J.M.K., Udovice, F. 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: The case of *Nothofagus* (Fagales). Syst. Biol., 61:289–313.

Scudder, S.H. 1875. Fossil butterflies. Mem. Am. Assoc. Adv. Sci. 1:1–99.

Scudder, S.H. 1889. The fossil butterflies of Florissant. United States Geological Survey, 8th Annual Report, pp. 439–472.

Scudder, S.H. 1892. Some insects of special interest from Florissant, Colorado, and other points in the Tertiaries of Colorado and Utah. Bull. U.S. Geol. Surv., 93:1–35.

Smith, A.B., and Peterson, K.J. 2002. Dating the time of origin of major clades: Molecular clocks and the fossil record. Annu. Rev. Earth Planet. Sci. 30:65–88.

Sohn, J.C., Labandeira, C., Davis, D., Mitter, C. 2012. An annotated catalog of fossil and subfossil Lepidoptera (Insecta: Holometabola) of the world. Zootaxa 3286:1–132.

Sohn, J.C., Labandeira, C.C., Davis, D.R. 2015. The fossil record and taphonomy of butterflies and moths (Insecta, Lepidoptera) and implications for evolutionary diversity and divergence-time estimates. BMC Evol. Biol. 15:12.

Soltis, D.E., Soltis, P.S., Zanis, M.J. 2002. Phylogeny of seed plants based on evidence from eight genes. Am. J. Bot. 89:1670–1681.

844 Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic
845 analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

846 Toussaint, E.F.A. and Balke, M. 2016. Historical biogeography of *Polyura* butterflies
847 in the oriental Palaeotropics: trans-archipelagic routes and South Pacific island
848 hopping. *J. Biogeogr.* 43:1560–1572.

849 Wahlberg, N., Leneveu, J., Kodandaramaiah, U., Peña, C., Nylin, S., Freitas, A.V.L.,
850 Brower, A.V.Z. 2009. Nymphalid butterflies diversify following near demise at the
851 Cretaceous/Tertiary boundary. *Proc. R. Soc. Lond. B*, 276:4295–4302.

852 Wahlberg, N., Wheat, C.W., Peña, C. 2013. Timing and patterns in the taxonomic
853 diversification of Lepidoptera (butterflies and moths). *PLoS ONE*, 8(11):e80875.

854 Walker, J.D., Geissman, J.W., Bowring, S.A., Babcock, L.E. 2013. The Geological
855 Society of America geologic time scale. *GSA Bull.*, 125:259–272.

856 Warnock, R.C.M., Yang, Z., Donoghue, P.C.J. 2012. Exploring uncertainty in the
857 calibration of the molecular clock. *Biol. Lett.*, 9:156–159.

858 Warnock, R.C.M., Parham, J.F., Joyce, W.G., Lyson, T.R., Donoghue, P.C.J. 2015.
859 Calibration uncertainty in molecular dating analyses: there is no substitute for the
860 prior evaluation of time priors. *Proc. R. Soc. Lond. B*, 282:20141013.

861 Yang, Z., and Rannala, B. 2006. Bayesian estimation of species divergence times
862 under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol.*
863 *Evol.*, 23:212–226.

864

865 TABLE 1. (a) Fossil calibration points used to calibrate the tree as a minimum age for
866 the *Clade calibrated*. Unless stated otherwise, the fossil calibrations were placed at
867 the stem of the clade calibrated. *Lower* and *upper* values indicate the prior truncation
868 for both the uniform and exponential priors. The 140 Ma year upper truncation
869 corresponds to the age of Angiosperms from Magallón et al. 2015. A different upper
870 truncation value results from a fossil prior interacting with a host plant prior placed at
871 the same node or a lower node. *Mean* and *offset* are parameter values for the
872 exponential prior distribution. (b) Host-plant clades used to calibrate the tree as a
873 maximum age for the *Calibrated node*. Host plant calibrations were placed at the
874 crown of the clade calibrated. Ages from both Magallón et al. (2015) and Foster et al.
875 (2017) are indicated.

876 a)

Fossils	Clade calibrated	lower	upper	mean	offset
<i>Doritis bosniaskii</i> Rebel, 1898	Papilionidae: Parnassiinae: Luehdorfiini	5.3	140	25	5.3
<i>Dynamine alexae</i> Peñalver & Grimaldi, 2006	Nymphalidae: Biblidinae: <i>Dynamine</i>	15.9	89	20	15.9
<i>Lethe corbieri</i> Nel, Nel & Balme, 1993	Nymphalidae: Satyrinae: Satyrini	28.3	65	25	28.3
<i>Mylothrites pluto</i> Heer, 1849	Pieridae: Coliadinae+Pierinae	15.9	100	50	15.9
<i>Neorinella garciae</i> Martins-Neto et al., 1993	Crown of Amathusiini	23.0	65	20	23.0
<i>Pamphilites abdita</i> Scudder, 1875	Hesperiidae: Hesperinae	23.0	140	30	23.0
<i>Prolibythea vagabunda</i> Scudder, 1889	Nymphalidae: Libytheinae	33.8	140	40.0	33.8
<i>Protochalcids kristenseni</i> de Jong, 2016	Hesperiidae: Coeliadinae	55.6	140	35	55.6
<i>Thaites ruminiana</i> Scudder, 1875	Papilionidae: Parnassiinae: Parnassiini	23.0	140	25	23.0
<i>Theope</i> sp	Riodinidae: Riodininae: Nymphidiini: <i>Theope</i>	15.9	140	25	15.9
<i>Voltinia damba</i> Hall, Robinson & Harvey, 2004	Riodinidae: Riodininae: Mesosemiini: <i>Voltina</i>	15.9	140	30	15.9
<i>Vanessa amerindica</i> Miller & Brown, 1989	Nymphalidae: Nymphalinae: Nymphalini	33.8	140	30	33.8
<i>Doxocopa wilmattae</i> Cockerell, 1907	Nymphalidae: Nymphalinae+Biblidinae+ Limenitidinae+Apaturinae		Not used		
<i>Praepapilio colorado</i> Durdin & Rose, 1978	Papilionidae		Not used		

877

878

879 b)

Host plant clade	Clade calibrated	Magallón et al. 2015	Foster et al. 2017
Angiospermae	root	140	252
Poaceae	Hesperiidae: Hesperinae	65	112
Poaceae	Nymphalidae: Satyrinae	65	112
Fabaceae	Pieridae	100	123
Brassicaceae	Pieridae: Pierinae	103	97
Rubiaceae	Riodinidae: Leucochimona+Mesophthalma+ Mesosemia+Perophthalma +Semomesia	87	85
Apocynaceae	Nymphalidae: Danainae	69	85
Solanaceae	Nymphalidae: Ithomiini	87	68
Euphorbiaceae	Nymphalidae: Biblidinae	89	104
Sapindaceae	Nymphalidae: Biblidinae: Epiphilini+ Callicorini	87	91

880

881

882

883 FIGURE 1. Time-calibrated tree obtained from the core analysis. a) The relationships
884 and age estimates among the subfamilies of Papilionoidea. b) The relationships and
885 age estimates among the genera across the different families. Age estimates are
886 indicated at the nodes (Ma). Node bars represent the 95% credibility intervals.

887

888 FIGURE 2. Comparison of node age estimates for the root of Papilionoidea and the
889 seven families between the core analysis and the seven alternative analyses. Mode,
890 median and 95% credibility interval are presented.

891

892 FIGURE 3. Marginal prior (grey) and posterior distributions (orange) for the nodes
893 calibrated in the core analysis. Blue dashed lines represent minimum boundaries;
894 green dashed lines represent maximum boundaries.

895

896 FIGURE 4. Marginal prior and posterior distributions for the root age in the core
897 analysis using either a) alternative host-plant ages or b) alternative subsets of fossil
898 calibrations.

899

900 FIGURE 5. Comparison of node age estimates for the root of Papilionoidea and the
901 seven families between this study (core analysis) and estimates from previous studies.
902 Mode and 95%CI for the core analysis are presented. For the other studies the values
903 reported in the original study are used.

Cretaceous

Paleogene

Neogene

SATYRINAE

Satyrini

Paleogene

Neogene

Density

a)

b)