

Mind the Outgroup and Bare Branches in Total-Evidence Dating: a Case Study
of Pimpliform Darwin Wasps (Hymenoptera, Ichneumonidae)

TAMARA SPASOJEVIC^{1,2,3}, GAVIN R. BROAD⁴, ILARI E. SÄÄKSJÄRVI⁵, MARTIN
SCHWARZ⁶, MASATO ITO⁷, STANISLAV KORENKO⁸, AND SERAINA
KLOPFSTEIN^{1,2,9}

¹ *Invertebrates, Naturhistorisches Museum der Burgergemeinde Bern, Bernastrasse 15, 3005
Bern, Switzerland*

² *Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, 3012 Bern,
Switzerland*

³ *Department of Entomology, National Museum of Natural History, 10th St. & Constitution
Ave. NW, Washington, DC 20560, USA*

⁴ *Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD,
UK*

⁵ *Biodiversity Unit, University of Turku, Turku, Finland*

⁶ *Eben 21, A-4202 Kirchsschlag, Austria*

⁷ *Graduate School of Agricultural Science, Kobe University, 657-8501 Japan*

⁸ *Department of Agroecology and Crop Production, Faculty of Agrobiolgy, Food and
Natural Resources, Czech University of Life Sciences Prague, Kamýcká 129, 165 21 Prague
6, Suchdol, Czech Republic*

⁹ *Biowissenschaften, Naturhistorisches Museum Basel, Augustinergasse 2, 4051 Basel,
Switzerland*

Correspondence to be sent to: Tamara Spasojevic, Department of Entomology, National
Museum of Natural History, 10th St. & Constitution Ave. NW, Washington, DC 20560,
USA, spasojevic.ta@gmail.com

Spasojevic et al.

ABSTRACT

Taxon sampling is a central aspect of phylogenetic study design, but it has received limited attention in the context of total-evidence dating, a widely used dating approach that directly integrates molecular and morphological information from extant and fossil taxa. We here assess the impact of commonly employed outgroup sampling schemes and missing morphological data in extant taxa on age estimates in a total-evidence dating analysis under the uniform tree prior. Our study group is Pimpliformes, a highly diverse, rapidly radiating group of parasitoid wasps of the family Ichneumonidae. We analyze a data set comprising 201 extant and 79 fossil taxa, including the oldest fossils of the family from the Early Cretaceous and the first unequivocal representatives of extant subfamilies from the mid Paleogene. Based on newly compiled molecular data from ten nuclear genes and a morphological matrix that includes 222 characters, we show that age estimates become both older and less precise with the inclusion of more distant and more poorly sampled outgroups. These outgroups not only lack morphological and temporal information, but also sit on long terminal branches and considerably increase the evolutionary rate heterogeneity. In addition, we discover an artefact that might be detrimental for total-evidence dating: “bare-branch attraction”, namely high attachment probabilities of certain fossils to terminal branches for which morphological data are missing. Using computer simulations, we confirm the generality of this phenomenon and show that a large phylogenetic distance to any of the extant taxa, rather than just older age, increases the risk of a fossil being misplaced due to bare-branch attraction. After restricting outgroup sampling and adding morphological data for the previously attracting, bare branches, we recover a Jurassic origin for Pimpliformes and Ichneumonidae. This first age estimate for the group not only suggests an older origin than previously thought, but also that diversification of the crown group happened well before the Cretaceous-Paleogene boundary. Our case study demonstrates that in order to obtain robust

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

age estimates, total-evidence dating studies need to be based on a thorough and balanced sampling of both extant and fossil taxa, with the aim of minimizing evolutionary rate heterogeneity and missing morphological information.

Keywords: ichneumonids, phylogeny, bare-branch attraction, fossils, RoguePlots, morphological matrix

Dating phylogenetic trees remains one of the most controversial issues in systematics. Until recently, the so-called “node dating” approach (ND) was the gold standard for dating phylogenetic trees, with fossils providing minimum ages for specific nodes in a phylogeny. ND requires the prior assessment of fossil placement, which is usually far from straightforward, and it can only incorporate the oldest fossil assignable to a particular node in the tree (Donoghue and Benton 2007; Donoghue and Yang 2016). In addition, one must decide on a probability distribution of the node’s age, because minima are insufficient to date molecular trees; although largely arbitrary, these settings determine the outcome of any ND analysis (Warnock et al. 2011; Klopstein 2020). These issues led to the development of the total-evidence dating approach (TED), which allows inclusion of all available fossils as tips, while accounting for uncertainty in their age and placement in a tree (Pyron 2011; Ronquist et al. 2012a). The downside of TED is that it requires extensive morphological matrices to be compiled, which inform fossil placements and associated branch lengths. Some concerns were also raised about the potential lack of clock-likeness of morphological data (O’Reilly et al. 2015), and some authors reported what they deemed unrealistically old ages from their TED analyses when compared to the oldest known fossils of the group (Beck and Lee 2014; Arcila et al. 2015).

Several methodological modifications to the TED approach have been suggested, such

Spasojevic et al.

as replacing the initially introduced uniform tree prior (Ronquist et al. 2012a) with the fossilized birth–death tree prior, which models speciation, extinction and fossilization rates to reconstruct branching events, and thus does not necessarily require morphological data (Heath et al. 2014; Ronquist et al. 2016; Zhang et al. 2016). It has also become possible to account for fossils possibly being sampled ancestors (Gavryushkina et al. 2014) and for a diversified sampling strategy of extant taxa (Höhna et al. 2011). Finally, some recent studies combined tip and node dating approaches (O’Reilly and Donoghue 2016; Kealy and Beck 2017; O’Hanlon et al. 2018; Travouillon and Phillips 2018). Unfortunately, these different implementations of TED have often reported disagreeing age estimates (Grimm et al. 2015; Herrera and Davalos 2016; Harrington and Reeder 2017; Kealy and Back 2017; Gustafson et al. 2017), revealing that our understanding of the inner workings of this method is still regrettably patchy (Parins-Fukuchi and Brown 2017).

Incorporating more complex models often requires more and better data for improved estimation of model parameters. With the development of next generation sequencing technologies, acquiring large amounts of molecular data is no longer a problem (Delsuc et al. 2005; McCormack et al. 2013; Misof et al. 2014), but taxon sampling is still a major limiting factor in phylogenetic study design. Besides the sampling of the focal, ingroup taxa, the choice of outgroup also requires special attention: it should ideally include a sufficient sample of taxa that are closely related to, but clearly different from the ingroup (Wheeler 1990; Nixon and Carpenter 1993; Giribet and Ribera 1998; Graham et al. 2002; Philippe et al. 2011). A poorly chosen outgroup can significantly affect topology estimates in non-clock analyses by introducing or at least exacerbating long branch attraction (Graham et al. 2002; Holland et al. 2003; Philippe et al. 2011) and/or by increasing compositional heterogeneity of sequences and among-lineage rate variation (Tarrío et al. 2000; Rota-Stabelli and Telford 2008; Borowiec et al. 2019). Numerous simulations and empirical studies have demonstrated

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

positive effects of improved taxon sampling on the estimation of topology (Graybeal 1998; Dunn et al. 2008; Heath et al. 2008; Klopstein et al. 2017), branch lengths (Fitch and Bruschi 1987; Pick et al. 2010), and parameters of evolutionary models (Zwickl and Hillis 2002; Heath et al. 2008). Nevertheless, thorough taxon sampling is not always easy to achieve, especially in very species-rich groups.

Several studies have investigated the impact of poor taxon sampling on age estimates in molecular dating analysis and demonstrated its severe negative effect. The effect was most pronounced when taxon sampling strategy led to strongly imbalanced phylogenetic trees and/or high among-lineage rate variation, which could not be adequately accommodated by existing rate smoothing algorithms or relaxed molecular-clock models (Milne 2009; Soares and Schrago 2012, 2015; Wertheim et al. 2012; Duchêne et al. 2014, 2015). Duchêne et al. (2015) have shown that the effect of tree imbalance on age estimates is even larger when heterochronous sequences are included, as is the case in molecular tip-dating with ancient DNA or in virus studies; a similar effect can be expected when fossil taxa of different ages are included as tips. Both the tree imbalance and high among-lineage rate variation can be introduced through a poorly chosen outgroup, as demonstrated for non-clock analyses (see above), but the outgroup choice notably has received little attention in molecular dating analyses, where a molecular-clock model can infer the root of the tree and thus makes the addition of an outgroup unnecessary. Accordingly, the outgroups then can be excluded from dating analysis to decrease the heterogeneity of evolutionary rates across the tree (Welch and Bromham 2005).

In practice, however, most of the recent TED studies included one to a handful of outgroup taxa, which are thus severely underrepresented compared to the ingroup, both in terms of morphological characters and fossils (Ronquist et al. 2012a; Arcila et al. 2015; Dornburg et al. 2015; Close et al. 2016; Lee 2016; Kittel et al. 2016; Herrera and Davalos

Spasojevic et al.

2016; Bannikov et al. 2017; Wang et al. 2018; Paterson et al. 2020). The reasons for including these outgroups, despite their potentially negative effects, is seldom reported in these studies and might vary between authors. They might not trust the clock model enough to rely on it for proper rooting, especially when rate heterogeneity is large, which has led to topological artefacts in the past (Ronquist et al. 2012a). Or they might feel uncomfortable with estimating the root age of their ingroup without providing additional taxa that branch off earlier than that node of main interest. Habit might be another incentive to stick to outgroups, which are standard for rooting non-clock trees. And finally, systematists might simply assume that adding more taxa and thus more data to an analysis will overall improve the results. We here aim to systematically test the influence of outgroup sampling on age estimates in TED, using parasitoid wasps of the family Ichneumonidae as a case study.

The Ichneumonidae, or Darwin wasps (Klopfstein et al. 2019b), is the most species-rich family of parasitoid wasps, with more than 25,000 described species (Yu et al. 2016), and at the same time one of the most severely understudied taxa. We focus on Pimpliformes, a monophyletic group comprising nine subfamilies: Acaenitinae, Collyriinae, Cylocheriinae, Diacritinae, Diplazontinae, Orthocentrinae, Pimplinae, Poemeniinae and Rhyssinae (Wahl and Gauld 1998; Quicke 2014; Klopfstein et al. 2019a). Pimpliformes are especially interesting from a biological perspective, since they cover nearly the entire diversity of hosts and parasitoid strategies known from ichneumonids (Broad et al. 2018). They oviposit into (endoparasitoids) or onto (ectoparasitoids) their host, which they either permanently paralyze (idiobionts) or allow to continue developing (koinobionts). Recorded hosts span almost all holometabolous insect orders, as well as spiders (Araneae). Several attempts have been made in the past to reconstruct the evolution of important biological traits in Pimpliformes (Wahl and Gauld 1998; Gauld et al. 2002; Quicke 2014), but their conclusions were highly

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

dependent on the stability and resolution of the pimpliform phylogeny, which is still in part unresolved (Klopfstein et al. 2019a).

Even less-well understood than the sequence of events during the radiation of Pimpliformes is their timing and thus ecological context. The fossil record of ichneumonids is very poorly studied, and most described species come from just a handful of localities (Menier et al. 2004). It starts in the Early Cretaceous with the extinct subfamily Tanychorinae (Kopylov 2010a), but affiliation of this subfamily with Ichneumonidae is somewhat unclear, as its wing venation is intermediate between Ichneumonidae and their sister family Braconidae (Sharkey and Wahl 1992). The Palaeoichneumoninae (Kopylov 2009), also extinct and of a similar age, are thus usually referred to as the oldest ichneumonids. All remaining ichneumonid fossils from the Cretaceous period have been classified in the extinct subfamilies Labenopimplinae and Novichneumoninae (Kopylov 2010b; Kopylov et al. 2010; Li et al. 2017), except for a single amber fossil that was tentatively placed in the extant subfamily Labeninae (McKellar et al. 2013). The oldest fossil associated with an extant pimpliform subfamily (and an extant genus) is an acaenitine, *Phaenolobus arvenus* Piton, from the latest Paleocene. However, its placement is questionable due to poor preservation, and more reliable records come from two Early Eocene localities, the Green River Formation and Messel Pit (Spasojevic et al. 2018a, 2018b).

No studies have to date attempted to infer the age of Ichneumonidae as a whole or of Pimpliformes in particular. The only previous studies with some bearing on the question are either concerned with the sister family Braconidae (Whitfield 2002) or with the entire order Hymenoptera (Peters et al. 2017); both studies included a very sparse sample of ichneumonids. They report very different age estimates for the divergence between Ichneumonidae and Braconidae, approx. 138 Ma (Whitfield 2002) and 155–224 (mean 188) Ma (Peters et al. 2017), respectively. In addition to obtaining the first age estimates for

Spasojevic et al.

Ichneumonidae and Pimpliformes and testing the influence of outgroup sampling on these estimates, we investigate the impact of fossil sampling and phylogenetic placement on the accuracy of age estimates. Finally, we discuss the implications of our findings for taxon sampling in dating studies in general.

MATERIAL AND METHODS

Taxon Sampling

We have included 289 species in our datasets, comprising 210 extant and 79 fossil taxa, with a focus on the pimpliform subfamilies within Ichneumonidae. Among the extant taxa, nine were outside of Ichneumonidae, including seven Braconidae (which together with Ichneumonidae belong to the superfamily Ichneumonoidea) and two more distantly related parasitoid wasps from the superfamilies Chalcidoidea and Evanioidea. The remaining extant taxa consisted of 30 non-pimpliform ichneumonids belonging to 19 subfamilies and an extensive sampling of Pimpliformes, for which we included 142 of the 188 known genera. For most genera, we included a single representative, but additional species were included in some morphologically heterogeneous genera. The complete list of extant taxa is given in Supplementary File S1 (available on Dryad at <https://dx.doi.org/10.5061/dryad.m0cfxpnm>).

We aimed to get a good representation of fossil taxa from different time periods: from the oldest ichneumonids from the Early Cretaceous to fossils from the latest Oligocene period (Supplementary File S2). Due to the controversial position of the extinct subfamily Tanychorinae, which is somewhat intermediate in morphology between Ichneumonidae and Braconidae, and poor sampling of the morphological diversity in Braconidae, we excluded the two Tanychorinae fossils from most analyses. However, we also ran two analyses under different outgroup sampling schemes with Tanychorinae included to assess their placement and the impact of their inclusion on age estimates (see below).

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

Morphological and Molecular Data

We used the morphological matrix from Klopstein and Spasojevic (2019), but with a strongly expanded taxon sampling. Numerous additional character states were defined to capture the added specimen diversity. The complete morphological matrix consists of 222 morphological characters coded initially for 150 extant and 79 fossil taxa. After observing attraction of fossil taxa to some non-pimpliform ichneumonids for which we had not yet coded any morphological data (“bare-branch attraction”, see Results section), we added another 20 extant taxa to the morphological matrix, increasing the total number of scored extant taxa to 170 (Supplementary File S3, also available at MorphoBank <http://morphobank.org/permalink/?P3821>). Morphology was scored for the same species from which we obtained molecular data, with a few exceptions where we scored a closely related, congeneric species for morphology (Supplementary File S1).

Our molecular dataset includes nine nuclear protein coding genes, which for 57 taxa were extracted from a previously compiled hybrid-capture dataset (Klopstein et al. 2019). For another 89 taxa, we newly obtain sequences of these genes using standard PCR and Sanger sequencing (primers and protocols in Klopstein et al. 2019). For all taxa, we added the D2/D3 portion of the nuclear rRNA gene 28S and the mitochondrial cytochrome c oxidase (COI; but see below) by Sanger sequencing. We achieved good coverage for these 11 genes for the 146 taxa, with an average of seven genes successfully sequenced per taxon (Supplementary File S4). The sequences were edited and aligned in Geneious 7.1.3 (<https://www.geneious.com>, Kearse et al. 2012) using the MAFFT v.7.017 plug-in and the algorithm “E-INS-I” (Katoh and Standley 2013). The “translation alignment” option was used for protein coding genes. The complete alignment contained 6213 base pairs (bp) (Supplementary File S5). The newly generated sequences are deposited in GenBank under accession numbers MW048210–MW048331, [NNNN] and [NNNN].

Spasojevic et al.

Non-Clock Analysis

To assess branch length heterogeneity and to examine the power of our combined molecular and morphological dataset to resolve ichneumonid relationships, we first ran a non-clock analysis. As the analysis showed convergence issues with parameter estimation for most of the 3rd codon partitions of the nine nuclear protein-coding genes and for the entire COI gene, we excluded those from all further analyses. The final alignment thus contained ten genes and 4,011 bp. We partitioned the dataset by gene and codon position and used PartitionFinder 2 (Lanfear et al. 2017) to identify partitions that could be combined (settings: branch lengths = linked, models =all, model_selection = aicc, search = recluster). In addition, we combined all 1st and all 2nd codon position partitions, respectively, that contained fewer than 30 parsimony informative sites, assuming that substitution model parameter estimates would be very poor for those. The preferred evolutionary model for all partitions was GTR+G+I according to PartitionFinder 2. As MrBayes allows model jumping over the entire GTR subspace, the model parameters for the molecular partitions were set as nst = mixed and rates = invgamma, with all substitution model parameters unlinked across partitions. Morphological characters were analyzed under the Mk model (Lewis 2001), accounting for ascertainment bias (“Mkv”, i.e., only variable characters coded), allowing gamma-distributed rate variation across characters, and ordering all the characters where transition only between neighboring states could be assumed (Supplementary File S6). This model has been identified as the preferred model for the morphological partition in a previous analysis (Klopfstein and Spasojevic 2018). The non-clock analysis included the full set of outgroup taxa (Chalcidoidea, Evanioidea, Braconidae and non-pimpliform Ichneumonidae), with *Gasteruption* (Evanioidea) chosen as the functional outgroup.

We ran four independent runs with four Metropolis-coupled chains each for 150 million generations with a sampling frequency of 1000. The heating coefficient was decreased from

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

the default value 0.1 to 0.05 in order to increase chain swap probabilities. To summarize the result, we used a conservative burn-in of 50%, while convergence of runs was assessed using typical Markov Chain Monte Carlo (MCMC) diagnostics: the average standard deviation of split frequencies (ASDSF), effective sample size (ESS), and the potential scale reduction factor (PSRF). We also visually inspected the trace plots of the likelihoods and of all parameters for all four runs using Tracer v1.7 (Rambaut et al. 2018). All phylogenetic analyses in this study were carried out using Bayesian inference in MrBayes 3.2.6 (Ronquist et al. 2012b) on the HPC cluster UBELIX of the University of Bern, Switzerland (<http://www.id.unibe.ch/hpc>). The final data matrices and resulting consensus trees are available at TreeBASE (<http://purl.org/phylo/treebase/phyloids/study/TB2:S26969>).

Total-Evidence Dating (TED) Analysis

In addition to the settings above, in the TED analysis, we used the uniform tree prior and a relaxed clock model with independent gamma rates (IGR). To set priors on the relaxed-clock model parameters, we relied on the calculations from Ronquist et al. (2012a), putting an exponential prior on the IGR variance with a rate of 37.12, and a lognormal prior on the clock rate with a mean on the log scale of -7.08069 and standard deviation of 1; the standard deviation was decreased compared to Ronquist et al. (2012) to put increased weight on lower clock rates and thus older age estimates, assuming that our dataset contains enough information from the data to correctly estimate the posterior. The prior on the tree age was set to $\text{offsetexp}(126, 309)$, with the offset based on the minimum age of the oldest Evanioidea fossil (Deans et al. 2004), while the mean corresponds to the mean age estimate for Hymenoptera from Ronquist et al. (2012a). In the analyses where non-ichneumonid outgroups were excluded (see below), we used the minimum age of the oldest unequivocal ichneumonid fossils (Palaeoichneumoninae, 112.6 Ma) as an offset. We set hard bounds using uniform priors on the age of the fossils according to the range of age estimates for the

Spasojevic et al.

fossil stratum (Barido-Sottani et al. 2019; Püschel et al. 2020). See Supplementary File S2 for the full list of included fossils and their age intervals with corresponding references.

To obtain the effective prior implied by our settings, including the clock rate (which is only effective when running with data), we performed an analysis with molecular and morphological data for extant taxa, but without the fossils and thus without temporal information. We ran two independent runs for 100 million generations each and under three different clock rate settings, and always obtained a rather flat age distribution for crown group ichneumonids (Supplementary File S7). Our preferred setting for the clock rate (lognormal with mean of log values = -7.08069 and standard deviation = 1), for instance, resulted in a median age estimate of 229.3 Ma and a 95% credibility interval (CI) of 51.3 – 738.6 Ma (Supplementary File S7). We can thus assume that any more precise age estimates resulting from the analyses with fossils will indeed be informed by the data and not the prior (Parins-Fukuchi and Brown 2017).

It has been shown that relaxed-clock models can lead to topology artefacts, especially close to the root (Ronquist et al. 2012a). We thus set hard constraints on the monophyly of Braconidae, Ichneumonidae and Ichneumonoidea (Braconidae + Ichneumonidae), each of which are widely accepted as being monophyletic and have been recovered in previous analyses (Sharkey and Wahl 1992; Downton and Austin 1994; Peters et al. 2017), as well as in our non-clock analysis. To improve convergence on the clock rate and tree length parameters, we increased the probabilities of the respective MCMC moves (MrBayes command blocks are provided as Supplementary File S8). MCMC convergence proved much more difficult to attain than in the non-clock analysis and was thus deemed satisfactory when the ASDSF value was below 0.03, ESS values of all scalar parameters were above 100, and PSRF values were below 1.01. The ESS values between 50 and 100 and PSRF values above 1.01 were still

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

accepted for the tree length, tree height and clock rate parameters in some runs, as it was difficult to get convergence on those even after 150 million generations.

TED Outgroup Settings

We tested five different outgroup sampling strategies (Table 1): i) “full outgroup” (with all outgroup taxa as in the non-clock analysis), ii) “Braconidae” (all braconids and non-pimpliform Ichneumonidae, but excluding the two non-ichneumonoid taxa), iii) “Braconidae (1)” (a single braconid taxon, *Homolobus*, the only braconid with both molecular and morphological data, and all non-pimpliform Ichneumonidae), iv) “Xoridinae” (only members of Ichneumonidae as outgroups), and v) “Xorides-only” (a single non-pimpliform Ichneumonidae, *Xorides*, as outgroup to Pimpliformes). To set the functional outgroup in each of these cases (i.e., non-ichneumonoids, braconids, one braconid, Xoridinae or *Xorides*), we enforced monophyly of the remaining taxa through a topology constraint. The fossils were in all cases included in the “Ichneumonidae” monophyly constraint, but excluded from any additional topology constraints (see below). Instead, their placement within Ichneumonidae was estimated entirely from the morphological data. In the “Xoridinae” analysis, we applied a partial constraint to enforce the sister relationship of xoridines and the remaining extant ichneumonids, for which there is strong evidence from previous phylogenetic studies (Klopfstein et al. 2019a), and at the same time to allow fossils to attach freely to any branch of the tree.

To assess potential bias when excluding the extinct Tanychorinae, we also performed the “full outgroup” and the “Xoridinae” analyses with two members of this subfamily included. In the “full outgroup” case, we constrained Tanychorinae within Ichneumonoidea, but applied a partial constraint on Braconidae and Ichneumonidae, which allowed Tanychorinae to attach to any crown or stem branch of the two families. In the “Xoridinae”

Spasojevic et al.

analysis, Tanychorinae were constrained to be the sister group to the remaining taxa, thus effectively rooting the tree with this extinct subfamily.

Fossil Placement

We assumed that an erroneous placement of the oldest included fossils would have the greatest influence, if any, on age estimates. We thus used “RoguePlots” as described in Klopstein and Spasojevic (2018) to examine the placement of the Cretaceous fossils on 1,000 evenly sampled trees from the four runs in relevant analyses (R package available at <https://github.com/seraklop/RoguePlots>). In the “Xoridinae” outgroup setting, the Cretaceous impression fossils were predominantly placed on some terminal branches leading to non-pimpliform taxa without morphological data (see Results section). We thus ran an additional analysis under the “Xoridinae” outgroup sampling scheme with an improved morphological matrix, which now did not contain any outgroup taxa without morphological data.

Simulations to Study Bare-Branch Attraction

To ascertain that the bare-branch attraction artefact that we observed was not due to some irregularity of our dataset, but instead represents a general phenomenon, we conducted simulations. As a simulation tree, we randomly chose one tree from the post-burnin sample under the “Xoridinae” outgroup scheme. Because the full tree with 278 taxa would have been too large to obtain a sufficient number of replicates in a reasonable timeframe, we reduced it to 61 extant and 61 fossil taxa that represent all major groups within Ichneumonidae (Supplementary File S9). To ensure that bare-branch attraction does not only occur in highly unbalanced trees, as with the tree we had produced through our outgroup sampling strategy, we removed mostly pimpliform ingroup taxa, which led to a more balanced tree of ichneumonids.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

On this simulation tree, which had all the fossils locked in place, we simulated 50 replicate datasets, each containing 300 bp of DNA data simulated under a Jukes-Cantor model and 100 binary characters simulated under equal rates. As an evolutionary rate, we chose 0.25 expected substitutions between root and any tip for both the molecular and the morphological data partitions. This value has been shown to be near-optimal for phylogenetic inference (Klopfstein et al. 2017), thus potentially producing informative data for TED analyses also. For all fossils, we removed the entire molecular partition, plus the same proportion of morphological data that was missing in the original dataset. Fossils thus had on average 27 characters coded (range: 8–61 characters). For the extant taxa, we included all characters, but in each replicate randomly chose 12 of the 61 extant taxa to become “bare” taxa – taxa with only molecular and no morphological data included. Analyses were run in MrBayes with settings identical to the original analyses, except only performing 10 million generations, which was sufficient to obtain ASDSF values below 0.025 while strongly reducing the computational burden.

To assess the impact of bare branches on fossil placement, we calculated for each of the 61 terminal branches leading to extant taxa the cumulative probability that they attracted a fossil, depending on whether they were bare in a given replicate or not. Since the placement of older fossils might be more unstable, e.g., due to poorer preservation or weakening of overall morphological signal from tips to root, we tested whether fossil age was the primary determinant for a fossil to be attracted to bare branches by calculating the same statistic separately for the Cretaceous and the Cenozoic fossils. The number of characters coded for a fossil might be another factor, so we also obtained the statistic separately for those fossils that had fewer than 20 characters coded. And finally, we examined the distance of each fossil from its closest extant taxon on the simulation tree, by adding up the relevant branch lengths. We then denoted one half of the fossils as “close” and one half as “distant”, depending on

Spasojevic et al.

whether their distance from any extant taxon was below or above the median distance. The R and shell scripts used to conduct these simulations and summarize the results, along with the corresponding simulated datasets, are available in Supplementary File S9.

RESULTS

Tree Resolution and Topology

The backbone of the majority-rule consensus trees from both the non-clock and TED analyses was unresolved when fossils were included, which was also reflected in the node support values for higher-level relationships. However, when the fossils were excluded from the sampled trees before summarizing them, resolution at the backbone was strongly improved (Fig. 1) and all the basal pimpliform nodes were highly supported (posterior probability >0.95). The exception was the “*Xorides*-only” TED analysis, where outgroup choice negatively affected the topological resolution of the backbone (Supplementary File S8). There were only a few topological differences between the non-clock and TED consensus trees, all concerning weakly supported nodes in both analyses (Supplementary files S6 and S8). Most of the expected higher-level relationships outside of Pimpliformes were recovered, such as Xoridae as the sister group to all other ichneumonids (Fig. 1). Within Pimpliformes, Diplazontinae were recovered as the sister group of the remaining subfamilies. The majority of pimpliform subfamilies were recovered as monophyletic (if we disregard a few taxa with alternative placements), the exceptions being Cylloceriinae, Diacritinae and Pimplinae. Details on alternative placements, taxonomic implications and proposed changes in classification are provided in Supplementary File S10.

Impact of Outgroup Sampling on Age Estimates

The median age estimates for Pimpliformes varied widely across the different outgroup analyses, ranging from 167 Ma to 263 Ma (Fig. 2, Supplementary File S8). Age estimates

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

were consistently older and typically less precise when outgroup taxa other than ichneumonids were included (i.e., more distant and more poorly sampled outgroups) and when only a single non-pimpliform ichneumonid was included as the outgroup (Fig. 2, Supplementary files S8 and S11). Age estimates were oldest when all seven braconid taxa were used as the outgroup, followed by the “Braconidae (1)” and “*Xorides*-only” sampling scheme; the latter resulted in the least precise age estimates (Fig. 2, Supplementary File S11). The inclusion of two representatives of the extinct Tanychorinae in the “full outgroup” and “Xoridinae” analyses also decreased the precision of the age estimates, while the impact on median age estimate was close to zero in the former and only minor in the latter analysis (Fig. 2).

Due to these large differences in age estimates, we assessed the impact of the outgroup setting on the variance of the clock rate to detect any pronounced rate heterogeneities between outgroups and ingroups. The estimated variance for the relaxed clock (IGR) varied considerably across the different outgroup settings (Fig. 3) and was nearly twice as high in the analyses with distant outgroups (“full outgroup” and “Braconidae”) compared to the analysis with only close outgroups (“Xoridinae” and “*Xorides*-only”). The exception was when a single braconid was included, in which case the terminal branch leading to the single braconid species also obtained a much higher clock rate than the remaining branches, but this was not yet sufficient to increase the average clock variance very much (results not shown). This finding confirms the pronounced rate variation among outgroups as observed already in the non-clock tree (Fig. 1), where especially some Braconidae species showed very long terminal branches compared to the ingroup.

From here onwards, we suppose that the age estimates were biased when more distant outgroups were included, as these outgroups introduced high rate variation, were poorly sampled, and no fossils were included that might have provided time information for the

Spasojevic et al.

outgroup branches to accurately estimate evolutionary rate in this part of the tree. This notion is supported by the smaller clock-rate variance and higher consistency in the age estimates when only close and well sampled outgroups were included (“Xoridinae” outgroup setting). We also consider age estimates in the “*Xorides*-only” analysis to be biased, given that it resulted in topological artefacts, poor precision of age estimates, and most importantly, numerous fossils were missing their closest extant relatives under this outgroup setting (see next section). Our preferred analysis is thus the “Xoridinae” outgroup sampling scheme.

Placement of Cretaceous Fossils

Placements of most of the Cretaceous compression fossils in our initial analyses with different outgroup settings were rather similar: they clustered predominantly within crown-groups of non-pimpliform ichneumonid subfamilies (Fig. 4b-c, Supplementary File S12). The highest placement probabilities were centered around extant Banchinae (*Banchus* and *Apophua*) and Tersilochinae (in the case of Labenopimplinae and *Tryphopimpla xoridoptera*) and/or on the branch leading to *Orthopelma* (especially in Palaeoichneumoninae), with only a few weakly supported placements in other parts of the tree (less than 10% attachment probability). All Labenopimplinae were placed with the highest probability (34–44%, respectively) on the branch leading to *Banchus*. The overall support for Labenopimplinae belonging to crown group Banchinae was quite high (46–53%), and it was even higher if we also considered stem Banchinae (63–73%). *Tryphopimpla xoridoptera*, a fossil with highly uncertain taxonomic affinities, was mostly associated with *Apophua* (34%), with a total probability of attachment within crown Banchinae of 55%, which increased to 71% when adding stem placements. In contrast, the small undescribed ichneumonid (3311_856b) from Late Cretaceous Yantardakh amber (Rasnitsyn et al. 2016) was attached with very high probability (97%) to the branch leading to the extant Phygadeuontinae genus *Gelis*. Interestingly, most of the tips to which the Cretaceous fossils attached contained no or only

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

sparse morphological information (Fig. 4b-c), leading us to postulate a “bare-branch attraction” phenomenon – a tendency especially for fossils with uncertain affinities to be attracted to branches leading to extant taxa for which no or very sparse morphological data has been obtained.

We thus repeated the analysis with the “Xoridinae” outgroup setting, after completing the scoring of morphological characters for all ichneumonid outgroup taxa, including those on “attracting”, bare branches. The placement of the Cretaceous impression fossils then changed considerably and mostly shifted towards the root of the tree compared to the previous analysis (Fig. 4d). Most of the Cretaceous fossils were now placed on stem branches of both non-pimpliform and pimpliform lineages, with only small placement probabilities for long branches of outgroup taxa (e.g., *Orthopelma*, *Brachycyrtus*). All Labenopimplinae now attached with the highest probability to the stem branch of the pimpliform subfamily Diplazontinae (23-39%) and to a nearby long branch leading to the extant tryphonine genera *Zagryphus* and *Thymaris* (22-32%). The probabilities of a placement of the Cretaceous impression fossils with crown Banchinae was now close to zero (Fig. 4d, Supplementary File S12).

Impact of fossil placement on age estimates

Although the placement of the Cretaceous fossils changed when the improved morphological matrix was employed, the median age estimates remained relatively stable (Fig. 5), but with a consistent improvement in the precision of the age estimates for all but one of the examined nodes. For example, the width of the 95% CI decreased from 53 Myr to 49 Myr for the ancestral node of Pimpliformes and from 63 Myr to 53 Myr for the ancestor of Diplazontinae (Supplementary Files S8 and S11).

Bare-branch attraction in simulated datasets

Spasojevic et al.

Our 50 simulated datasets overall behaved similarly to the original data, in that most nodes of the tree of extant taxa were recovered well, but a few nodes that were subtended by very short branches in the simulation tree were not resolved. Most of the fossils behaved as rogues – taxa with uncertain placement in phylogenetic analyses, as in the original analyses, and including them removed any resolution from the consensus tree. RoguePlots of the fossils showed a heterogeneous picture, with some fossils placed rather firmly, while others shifted around a lot, again reflecting rather well the situation with the empirical data. Analyses of fossil attachment probabilities for the terminal branches leading to extant taxa confirm that bare-branch attraction occurs in simulated datasets as well (Fig. 6). On average, this probability was higher on bare branches than on non-bare ones, although with a lot of variation. The effect was strongest for the one half of fossils that were more distant from any extant taxon than the others, which were preferentially attaching to bare branches, while the opposite was true for the fossils with rather close ties to at least one extant taxon. Age also seemed to have an impact, but it was far less strong, as was the influence of the number of characters coded for a fossil (Fig. 6).

Age of Pimpliformes

The age estimates resulting from our preferred analysis, with multiple but only closely related and well-sampled outgroup taxa (“Xoridinae”, Table 2 and Fig. 7), and after filling in the attracting bare branches, suggest that Ichneumonidae originated during the Early Jurassic (95% CI spans most of the Jurassic: 154.0–204.7 Ma) and Pimpliformes during the Middle (95% CI spans most of the Jurassic 144.0–193.4 Ma). The start of the radiation of most of the pimpliform subfamilies is estimated as having occurred in the Early Cretaceous.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

DISCUSSION

Impact of Outgroup Sampling on Age Estimates

We show here that with the inclusion of more distantly related and/or poorly sampled outgroups, age estimates for pimpliform parasitoid wasps become older and often less precise. Besides creating a highly unbalanced tree in terms of sampling of both extant and fossil taxa, including distant outgroups in our case introduced large variation in the clock rate, which is most likely the main reason for the observed bias in age estimates (Magallón et al. 2013; Magallón 2014; Beck and Lee 2014; Beaulieu et al. 2015; King et al. 2017a). In addition, in the absence of time information in the form of fossils along the branches of distant outgroups together, the age estimates might have been driven by our clock rate prior, resulting in older estimates for crown group Ichneumonidae and Pimpliformes in those analyses (see analysis of effective prior in Material and Methods). Multiple studies have shown that outgroup choice can greatly affect tree topology estimates, especially in cases with large heterogeneity of branch lengths and uneven taxon sampling (Puslednik and Serb 2008; Ware et al. 2008; Hayes et al. 2009; Thomas et al. 2013; Kirchberger et al. 2014; Wilberg 2015). However, the influence of outgroup choice on divergence time estimates, especially in the context of total-evidence dating, has scarcely been studied, and only indirectly (Linder et al. 2005; Soares and Schrago 2012; Duchêne et al. 2015; Matschiner 2019).

In the last eight years since total-evidence dating became established, most of studies have included outgroups in their taxon sampling (Ronquist et al. 2012a; Arcila et al. 2015; Dornburg et al. 2015; Close et al. 2016; Herrera and Davalos 2016; Kittel et al. 2016; Lee 2016; Bannikov et al. 2017). We here covered most of the outgroup-sampling schemes found in these studies, from including a single closely related taxon (“*Xorides*-only” outgroup

Spasojevic et al.

setting), over a few relatively closely related taxa (“Braconidae” and “Braconidae (1)”), to the inclusion of a series of more to less closely related outgroup taxa (“full outgroup”). As in most previous studies, our non-ichneumonid outgroups were not only sparsely sampled, but were also missing fossils and morphological data for most of the taxa. All these outgroup settings recovered older and usually less precise age estimates than when we restricted our dataset to more closely related outgroup taxa, for which greater and more even effort in sampling both extant and fossil taxa had been applied (“Xoridinae” outgroup setting). Interestingly, when two potentially transitional fossils between Braconidae and Ichneumonidae (the Tanychorinae) were included, the median age estimates and clock variance were only slightly affected, but their precision dropped significantly. This could result from the rather large uncertainty in the age of these fossils themselves, which results from a disagreement over the age of the Khasurty Formation (Kopylov and Rasnitsyn 2017; PaleoBioDB 2020), and the fact that only very few morphological characters were scored for them (9% and 17%, respectively).

Our results might suggest that many of the previous TED analyses would recover younger and likely more accurate age estimates without outgroups or with either more detailed or more restricted, but more even outgroup sampling. Presumably, the effect would be most pronounced for datasets where there are long, unbroken outgroup branches which introduce large rate heterogeneity across the tree. A similar effect has been demonstrated for node dating in the simulation study by Soares and Schrago (2015), where age estimates were significantly biased when there was a combination of large among-lineage rate variation and poor taxon sampling. As in our case, both accuracy and precision were affected in their simulations: the mean age of the node in question was constantly overestimated and precision severely decreased. Some TED studies already conjectured that high evolutionary rate heterogeneity can lead to biased age estimates (Beck and Lee 2014; Lee 2016; King et al.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

2017b; Bagley et al. 2018; Luo et al. 2020). Among-lineage rate heterogeneity can already be identified on a non-clock tree by comparing root-to-tip distances of extant taxa and problematic taxa can accordingly be excluded, as we demonstrated here; however this has rarely been employed in TED studies (but see Grimm et al. 2015).

In addition, more realistic models of morphological evolution might eventually improve the clock-likeness of morphological datasets by improving their match to the underlying evolutionary mechanisms. Some of the aspects for which attempts have already been made to increase realism are asymmetry in transition rates among character states (Lewis 2001; Wright et al. 2016; Pyron 2017), non-stationarity of the evolutionary process (Klopfstein et al. 2015) or even the modelling of undetected character states through hidden Markov models (Tarasov 2019). These models have not been routinely used in TED studies and, at least at present, many of them are either restricted to specific character types or are computationally challenging for large morphological data sets. The latter was also true for our dataset when we tried to account for asymmetric transition rates among character states (Lewis 2001; Wright et al. 2016)

Finally, incompletely sampled phylogenies, with respect to fossil and/or extant taxa, could potentially be modelled more adequately with the fossilized birth-death tree prior (FBD) instead of the uniform tree prior used here, but only when model assumptions are satisfied (Zhang et al. 2016). When model assumptions are violated, such as assumptions related to the sampling density of fossil and/or extant taxa, age estimates are quickly compromised (Matschiner 2019; Luo et al. 2020; Püschel et al. 2020). This makes it difficult to apply the FBD prior to the ichneumonid dataset at present, as species richness through time even today is still poorly known for this group.

Bare-branch attraction hampering correct fossil placement in TED

Spasojevic et al.

In TED analyses, the placement of fossils is solely dependent on the available morphological information for both fossil and extant taxa. The quality of fossil placement based on morphological matrices is thus primarily limited by imperfect preservation of fossils, but also by high levels of morphological homoplasy, which has been reported for ichneumonids (Gauld and Mound 1982; Klopstein and Spasojevic 2019). We here identified another potentially major issue in TED studies, which we called “bare-branch attraction”: the tendency of fossils which are only distantly related to any of the included extant taxa to attach to terminal branches of extant taxa for which no morphological data have been collected. This artefact exposes the dangers of insufficient sampling of extant taxa for morphology, which can distort fossil placement and consequently age estimates in TED analyses.

Many of the previous TED studies included from a few to more than half of the extant taxa without morphological data or at least with high amounts of missing data (Ronquist et al. 2012b; Arcila et al. 2015; Dornburg et al. 2015; Harrington and Reeder 2017). Guillaume and Cooper (2016) addressed this issue in the context of topology reconstruction in TED analyses. In their simulations, topology estimates were more negatively affected by a large percentage of extant taxa with missing morphological data than by any other analyzed parameter. It remains unclear to what extent their results were influenced by bare-branch attraction, as they did not analyze individual fossil placements, but it is likely that the artefact played a role under their scenario as well.

In our study, bare-branch attraction was most obvious in the compression fossils from the Cretaceous. With the exception of *T. xoridoptera*, these fossils are all classified in two extinct subfamilies, Palaeoichneumoninae and Labenopimplinae. The phylogenetic position of these subfamilies is unclear, but two options have been suggested: a transitional position between Tanychorinae and extant Ichneumonidae, which would mean they represent stem ichneumonids, or some rather basal position as crown ichneumonids (Kopylov 2009, 2010b).

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

In fact, the name “Labenopimplinae” reflects their similarity to the extant subfamilies Labeninae and Pimplinae (Kopylov 2010b).

We showed that the predominant placement in our initial analysis, of most Labenopimplinae and Paleoichneumoninae with crown group Banchinae and Tersilochinae, was a bare-branch attraction artefact. Among the branches where these fossils attached on the tree, only a single tip (*Apophua*) contained morphological data, while the remaining taxa were only sampled for molecular characters. Some other, younger fossils with rather labile placement also often attached to these branches, which might suggest that when morphological information for the placement of a fossil is limited, TED analyses tend to place a fossil on “bare-branches”, especially if those branches are long. Filling in the missing information among extant outgroup taxa was enough to reverse the initially erroneous placements of the Cretaceous fossils. They then instead ended up on rather basal branches of crown group ichneumonids, in accordance with one of the hypotheses suggested at the time of their original description (Kopylov 2009, 2010b).

Simulations confirm the bare-branch attraction phenomenon

Our simulations have confirmed bare-branch attraction as a general phenomenon, which could be reproduced with simulated molecular and morphological data. The distance in branch lengths between a fossil and its closest extant relative emerged as the strongest factor determining the strength of the artefact, ahead of fossil age and number of morphological characters coded. This makes intuitive sense: fossils that are close to an extant taxon will on average have high information content in their morphological data, even if it is incomplete, while fossils that are distant to any extant taxon are notoriously difficult to place and might thus be more prone to falling victim to bare-branch attraction.

Spasojevic et al.

More extensive simulations are needed to further assess the types of biases arising from bare-branch attraction. In a recent simulation study concerning the FBD tree prior, Luo et al. (2020) reported difficulties in recovering correct fossil placement when rate variation across branches was high and when morphological data were not included. This might indicate the bare-branch attraction and it remains to be shown how detrimental this artefact is in the contexts of FBD prior. In any case, this phenomenon has the potential to mislead TED analyses and should be accounted for in future studies. Fortunately, the remedy is very simple: morphological datasets of extant taxa should be as complete as possible.

The age of Pimpliformes and the biological context of their diversification

Our most credible analysis estimated the median age of the family Ichneumonidae at 181 Ma (95% CI: 154.0–204.7 Ma) and of Pimpliformes to 168 Ma (95% CI: 144.0–193.4 Ma), with 95% credibility intervals in both cases spanning most of the Jurassic. The estimated age for Ichneumonidae is thus 60–70 Ma older than the oldest certain ichneumonid fossils, which implies a rather long ghost lineage. However, a similar gap exists between the oldest and the second oldest ichneumonid fossils, which is between 26 and 53 Ma, depending on the age of the geological formations in question (Kopylov and Rasnitsyn 2017; PaleoBioDB 2020). This suggests that the ghost range implied by our analysis is not that long after all. Such a gap is even more acceptable if we consider the paucity of Jurassic Hymenoptera fossils in general (Rasnitsyn and Quicke 2002). Furthermore, the last few years have seen unexpected discoveries of fossils that have closed large gaps between much older (molecular) age estimates and the previously known fossil record, for instance in Lepidoptera (Eldijk et al. 2018).

Further insights into the age of ichneumonids come from age estimates for their hosts (nearly exclusively holometabolous insects), which had to originate before their parasitoids

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

could radiate. Initially, the radiations of the most species-rich orders of holometabolous insects, Hymenoptera, Coleoptera, Diptera and Lepidoptera, were believed to have been associated with the radiation of flowering plants and were dated to the Early Cretaceous (Grimaldi 1999; Misof et al. 2014). However, these estimates were later deemed too young, and a Late Permian origin was suggested based on a re-analysis of a large phylogenomic dataset (Misof et al. 2014) with more appropriate calibration points (Tong et al. 2015). This later study implies that the major host groups for ichneumonids were already present during the Jurassic, when both Ichneumonidae and Pimpliformes originated, suggesting that the radiation of these parasitoids might have happened only shortly after the radiation of their host groups.

CONCLUSIONS

We have demonstrated that poor outgroup sampling can negatively affect both accuracy and precision of age estimates in total-evidence dating analyses. Even though the exact mechanisms leading to this decrease in performance with the inclusion of increasingly more distant outgroups remain somewhat unclear, the introduction of large among-lineage rate variation and tree imbalance probably play major roles. Thus, to achieve more reliable age estimates, one should consider excluding outgroups altogether from dating analyses, unless they can be sampled adequately. Correct positioning of the root then can be through topology constraints instead, that are based on prior knowledge, or on a non-clock analysis of the data. In case an outgroup needs to be included in the analysis, for whatever reasons, one should sufficiently sample not only molecular, but also morphological and fossil data.

We also illustrated the importance of careful consideration of fossil placement in total-evidence dating analyses in order to identify artefacts or biases. The bare-branch attraction artefact that we have discovered here is likely universally problematic for TED. Most

Spasojevic et al.

concerning is that the artefact most severely affects fossils distant from extant taxa in a phylogeny, such as older and stem-lineage fossils, whose correct placement has proven crucial for reliable time-calibration of phylogenies. Thus, the bare-branch attraction artefact deserves further assessment in the future, through more detailed simulation studies. Nevertheless, it can easily be circumvented by a more complete sampling of morphological data for extant taxa.

Finally, we provided the first age estimate for the extremely diverse group of ichneumonid parasitoid wasps. The Jurassic origin for the family and for Pimpliformes agrees with the timing of the radiation of their major host groups. It remains to be seen how the age estimate for the family will change when more taxa and especially more fossils are included in a TED analysis. As new fossil ichneumonids are being described at a regular pace (Khalaim 2008; Kopylov 2009, 2010b; McKellar et al. 2013; Antropov et al. 2014; Kopylov et al. 2018; Spasojevic et al. 2018a, 2018b), the coming years will certainly provide further evidence for an older age of many groups within this species-rich family. Our insights on taxon sampling will help provide guidance to design increasingly efficient dating studies, not only in Darwin wasps, but across the entire tree of life.

FUNDING

This work was supported by the Swiss National Science Foundation (SNSF, grant number PZ00P3_154791 to Seraina Klopstein). During the submission and revision process, Tamara Spasojevic was supported by the SNSF (grant number P2BEP3_1882252). Stanislav Korenko was supported by the grant LTAUSA 19084 provided by the Ministry of Education Youth and Sports of Czech Republic (program Inter-excellence, Inter-action).

ACKNOWLEDGEMENTS

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

For access to fossil specimens, we are grateful to Marsh Finnegan and Alan M. Rulis (Smithsonian National Museum of Natural History, Washington, USA), Talia Karim and David Zelagin (Museum of Natural History, University of Colorado, Boulder, USA) and Ricardo Pérez-de la Fuente (Museum of Comparative Zoology, Harvard University, USA). We also thank Dmitry Kopylov (Borissiak Paleontological Institute, Russian Academy of Sciences, Moscow, Russia) and Sonja Wedmann (Forschungsstation Grube Messel, Senckenberg, Germany) for providing access to additional fossil specimens and for commenting on derived fossil ages. We are grateful to Eric Chapman and Mike Sharkey (Kentucky, USA), Rikio Matsumoto (Osaka, Japan), and the Swedish Malaise Trap project for extensive material for morphological and molecular studies. Remo Ryser (German Centre for Integrative Biodiversity Research, Leipzig, Germany) and Djordje Markovic (Institute of Zoology, University of Belgrade, Belgrade, Serbia) contributed to the lab work. Finally, we thank Simon Ho (editor), Alexander Pyron and two anonymous reviewers for critical and insightful comments which significantly improved the quality of the manuscript.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository:

<https://dx.doi.org/10.5061/dryad.m0cfxpnm>

REFERENCES

Antropov A.V., Belokobylskij S.A., Compton S.G., Dlussky G.M., Khalaim A.I., Kolyada V.A., Kozlov M.A., Perfilieva K.S., Rasnitsyn A.P. 2014. The wasps, bees and ants (Insecta: Vespida=Hymenoptera) from the Insect Limestone (Late Eocene) of the Isle of Wight, UK. *Earth Environ. Sci. Trans. R. Soc. Edinb.* 104:335–446.

Spasojevic et al.

Arcila D., Alexander Pyron R., Tyler J.C., Ortí G., Betancur-R. R. 2015. An evaluation of fossil tip-dating versus node-age calibrations in tetraodontiform fishes (Teleostei: Percomorphaceae). *Mol. Phylogenet. Evol.* 82:131–145.

Bagley J.C., Mayden R.L., Harris P.M. 2018. Phylogeny and divergence times of suckers (Cypriniformes: Catostomidae) inferred from Bayesian total-evidence analyses of molecules, morphology, and fossils. *PeerJ.* 6:e5168.

Bannikov A.F., Tyler J.C., Arcila D., Carnevale G. 2017. A new family of gymnodont fish (Tetraodontiformes) from the earliest Eocene of the Peri-Tethys (Kabardino-Balkaria, northern Caucasus, Russia). *J. Syst. Palaeontol.* 15:129–146.

Barido-Sottani J., Aguirre-Fernández G., Hopkins M.J., Stadler T., Warnock R. 2019. Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence times under the fossilized birth–death process. *Proc. R. Soc. B Biol. Sci.* 286:20190685.

Beaulieu J.M., O’Meara B.C., Crane P., Donoghue M.J. 2015. Heterogeneous Rates of Molecular Evolution and Diversification Could Explain the Triassic Age Estimate for Angiosperms. *Syst. Biol.* 64:869–878.

Beck R.M.D., Lee M.S.Y. 2014. Ancient dates or accelerated rates? Morphological clocks and the antiquity of placental mammals. *Proc. R. Soc. B Biol. Sci.* 281:20141278.

Borowiec M.L., Rabeling C., Brady S.G., Fisher B.L., Schultz T.R., Ward P.S. 2019. Compositional heterogeneity and outgroup choice influence the internal phylogeny of the ants. *Mol. Phylogenet. Evol.* 134:111–121.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

- Broad G.R., Shaw M.R., Fitton M.G. 2018. Ichneumonid Wasps (Hymenoptera: Ichneumonidae): Their Classification and Biology. Field Studies Council and Royal Entomological Society.
- Close R.A., Johanson Z., Tyler J.C., Harrington R.C., Friedman M. 2016. Mosaicism in a New Eocene Pufferfish Highlights Rapid Morphological Innovation Near the Origin of Crown Tetraodontiforms. *Palaeontology*. 59:499–514.
- Deans A.R., Basibuyuk H.H., Azar D., Nel A. 2004. Descriptions of two new Early Cretaceous (Hauterivian) ensign wasp genera (Hymenoptera: Evaniidae) from Lebanese amber. *Cretac. Res.* 25:509–516.
- Delsuc F., Brinkmann H., Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6:361–375.
- Donoghue P.C.J., Benton M.J. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. *Trends Ecol. Evol.* 22:424–431.
- Donoghue P.C.J., Yang Z. 2016. The evolution of methods for establishing evolutionary timescales. *Phil Trans R Soc B.* 371:20160020.
- Dornburg A., Friedman M., Near T.J. 2015. Phylogenetic analysis of molecular and morphological data highlights uncertainty in the relationships of fossil and living species of Elopomorpha (Actinopterygii: Teleostei). *Mol. Phylogenet. Evol.* 89:205–218.
- Dowton M., Austin A.D. 1994. Molecular phylogeny of the insect order Hymenoptera: apocritan relationships. *Proc. Natl. Acad. Sci.* 91:9911–9915.

Spasojevic et al.

Duchêne D., Duchêne S., Ho S.Y.W. 2015. Tree imbalance causes a bias in phylogenetic estimation of evolutionary timescales using heterochronous sequences. *Mol. Ecol. Resour.* 15:785–794.

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

Dunn C.W., Hejnal A., Matus D.Q., Pang K., Browne W.E., Smith S.A., Seaver E., Rouse G.W., Obst M., Edgecombe G.D., Sørensen M.V., Haddock S.H.D., Schmidt-Rhaesa A., Okusu A., Kristensen R.M., Wheeler W.C., Martindale M.Q., Giribet G. 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature.* 452:745–749.

Eldijk T.J.B. van, Wappler T., Strother P.K., Weijst C.M.H. van der, Rajaei H., Visscher H., Schootbrugge B. van de. 2018. A Triassic-Jurassic window into the evolution of Lepidoptera. *Sci. Adv.* 4:e1701568.

Fitch W.M., Bruschi M. 1987. The evolution of prokaryotic ferredoxins--with a general method correcting for unobserved substitutions in less branched lineages. *Mol. Biol. Evol.* 4:381–394.

Gauld I.D., Mound L.A. 1982. Homoplasy and the delineation of holophyletic genera in some insect groups. *Syst. Entomol.* 7:73–86.

Gauld I.D., Wahl D.B., Broad G.R. 2002. The suprageneric groups of the Pimplinae (Hymenoptera: Ichneumonidae): a cladistic re-evaluation and evolutionary biological study. *Zool. J. Linn. Soc.* 136:421–485.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

- Gavryushkina A., Welch D., Stadler T., Drummond A.J. 2014. Bayesian Inference of Sampled Ancestor Trees for Epidemiology and Fossil Calibration. *PLOS Comput. Biol.* 10:e1003919.
- Giribet G., Ribera C. 1998. The Position of Arthropods in the Animal Kingdom: A Search for a Reliable Outgroup for Internal Arthropod Phylogeny. *Mol. Phylogenet. Evol.* 9:481–488.
- Graham S.W., Olmstead R.G., Barrett S.C.H. 2002. Rooting Phylogenetic Trees with Distant Outgroups: A Case Study from the Commelinoid Monocots. *Mol. Biol. Evol.* 19:1769–1781.
- Graybeal A. 1998. Is It Better to Add Taxa or Characters to a Difficult Phylogenetic Problem? *Syst. Biol.* 47:9–17.
- Grimaldi D. 1999. The Co-Radiations of Pollinating Insects and Angiosperms in the Cretaceous. *Ann. Mo. Bot. Gard.* 86:373–406.
- Grimm G.W., Kapli P., Bomfleur B., McLoughlin S., Renner S.S. 2015. Using More Than the Oldest Fossils: Dating Osmundaceae with Three Bayesian Clock Approaches. *Syst. Biol.* 64:396–405.
- Guillerme T., Cooper N. 2016. Effects of missing data on topological inference using a Total Evidence approach. *Mol. Phylogenet. Evol.* 94, Part A:146–158.
- Harrington S.M., Reeder T.W. 2017. Phylogenetic inference and divergence dating of snakes using molecules, morphology and fossils: new insights into convergent evolution of feeding morphology and limb reduction. *Biol. J. Linn. Soc.* 121:379–394.

Spasojevic et al.

- Hayes K.A., Cowie R.H., Thiengo S.C. 2009. A global phylogeny of apple snails: Gondwanan origin, generic relationships, and the influence of outgroup choice (Caenogastropoda: Ampullariidae). *Biol. J. Linn. Soc.* 98:61–76.
- Heath T.A., Hedtke S.M., Hillis D.M. 2008. Taxon sampling and the accuracy of phylogenetic analyses. *J. Syst. Evol.*:19.
- Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proc. Natl. Acad. Sci.* 111:E2957–E2966.
- Herrera J.P., Davalos L.M. 2016. Phylogeny and Divergence Times of Lemurs Inferred with Recent and Ancient Fossils in the Tree. *Syst. Biol.* 65:772–791.
- Höhna S., Stadler T., Ronquist F., Britton T. 2011. Inferring Speciation and Extinction Rates under Different Sampling Schemes. *Mol. Biol. Evol.* 28:2577–2589.
- Holland B.R., Penny D., Henny M.D., Sullivan J. 2003. Outgroup Misplacement and Phylogenetic Inaccuracy Under a Molecular Clock—A Simulation Study. *Syst. Biol.* 52:229–238.
- Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.
- Kealy S., Beck R. 2017. Total evidence phylogeny and evolutionary timescale for Australian faunivorous marsupials (Dasyuromorphia). *BMC Evol. Biol.* 17:240.
- Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P., Drummond A. 2012.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28:1647–1649.

Khalaim A.I. 2008. Fossil ichneumon wasps (Hymenoptera: Ichneumonidae) from Biamo (Russia), Oligocene. *Alavesia*. 2:101–112.

King B., Qiao T., Lee M.S.Y., Zhu M., Long J.A. 2017a. Bayesian Morphological Clock Methods Resurrect Placoderm Monophyly and Reveal Rapid Early Evolution in Jawed Vertebrates. *Syst. Biol.* 66:499–516.

King B., Qiao T., Lee M.S.Y., Zhu M., Long J.A. 2017b. Bayesian Morphological Clock Methods Resurrect Placoderm Monophyly and Reveal Rapid Early Evolution in Jawed Vertebrates. *Syst. Biol.* 66:499–516.

Kirchberger P.C., Sefc K.M., Sturmbauer C., Koblmüller S. 2014. Outgroup effects on root position and tree topology in the AFLP phylogeny of a rapidly radiating lineage of cichlid fish. *Mol. Phylogenet. Evol.* 70:57–62.

Kittel R.N., Austin A.D., Klopfstein S. 2016. Molecular and morphological phylogenetics of chelonine parasitoid wasps (Hymenoptera: Braconidae), with a critical assessment of divergence time estimations. *Mol. Phylogenet. Evol.* 101:224–241.

Klopfstein S. 2020. The age of insects and the revival of minimum age trees. *Austral Entomology*. [in press].

Klopfstein S., Langille B., Spasojevic T., Broad G.R., Cooper S.J.B., Austin A.D., Niehuis O. 2019a. Hybrid capture data unravel a rapid radiation of pimpliform parasitoid wasps (Hymenoptera: Ichneumonidae: Pimpliformes). *Syst. Entomol.* 44:361–383.

Spasojevic et al.

Klopfstein S., Massingham T., Goldman N. 2017. More on the Best Evolutionary Rate for Phylogenetic Analysis. *Syst. Biol.* 66:769–785.

Klopfstein S., Santos B.F., Shaw M.R., Alvarado M., Bennett A.M.R., Pos D.D., Giannotta M., Florez A.F.H., Karlsson D., Khalaim A.I., Lima A.R., Mikó I., Sääksjärvi I.E., Shimizu S., Spasojevic T., Noort S. van, Vilhelmsen L., Broad G.R. 2019b. Darwin wasps: a new name heralds renewed efforts to unravel the evolutionary history of Ichneumonidae. *Entomol. Commun.* 1:ec01006–ec01006.

Klopfstein S., Spasojevic T. 2019. Illustrating phylogenetic placement of fossils using RoguePlots: An example from ichneumonid parasitoid wasps (Hymenoptera, Ichneumonidae) and an extensive morphological matrix. *PLOS ONE.*:425090.

Klopfstein S., Vilhelmsen L., Ronquist F. 2015. A Nonstationary Markov Model Detects Directional Evolution in Hymenopteran Morphology. *Syst. Biol.* 64:1089–1103.

Kopylov D.S. 2009. A new subfamily of ichneumonids from the Lower Cretaceous of Transbaikalia and Mongolia (Insecta: Hymenoptera: Ichneumonidae). *Paleontol. J.* 43:83–93.

Kopylov D.S. 2010a. Ichneumonids of the subfamily Tanychorinae (Insecta: Hymenoptera: Ichneumonidae) from the Lower Cretaceous of Transbaikalia and Mongolia. *Paleontol. J.* 44:180–187.

Kopylov D.S. 2010b. A new subfamily of ichneumon wasps (Insecta: Hymenoptera: Ichneumonidae) from the Upper Cretaceous of the Russian Far East. *Paleontol. J.* 44:422–433.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

- Kopylov D.S., Brothers D.J., Rasnitsyn A.P. 2010. Two New Labenopimpline Ichneumonids (Hymenoptera: Ichneumonidae) from the Upper Cretaceous of Southern Africa. *Afr. Invertebr.* 51:423–430.
- Kopylov D.S., Rasnitsyn A.P. 2017. New sepulcids (Hymenoptera: Sepulcidae) from the Lower Cretaceous of Asia: II. Ghilarellinae and Trematothoracinae. *Paleontol. J.* 51:291–303.
- Kopylov D.S., Spasojevic T., Klopstein S. 2018. New ichneumonids (Hymenoptera, Ichneumonidae) from the Eocene Tadushi Formation, Russian Far East. *Zootaxa.* 4442:319–330.
- Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Mol. Biol. Evol.* 34:772–773.
- Lee M.S.Y. 2016. Multiple morphological clocks and total-evidence tip-dating in mammals. *Biol. Lett.* 12:20160033.
- Lewis P.O. 2001. A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data. *Syst. Biol.* 50:913–925.
- Li L., Kopylov D.S., Shih C., Ren D. 2017. The first record of Ichneumonidae (Insecta: Hymenoptera) from the Upper Cretaceous of Myanmar. *Cretac. Res.* 70:152–162.
- Linder H.P., Hardy C.R., Rutschmann F. 2005. Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Mol. Phylogenet. Evol.* 35:569–582.

Spasojevic et al.

Luo A., Duchêne D.A., Zhang C., Zhu C.-D., Ho S.Y.W. 2020. A Simulation-Based

Evaluation of Tip-Dating Under the Fossilized Birth–Death Process. *Syst. Biol.*

69:325–344.

Magallón S. 2014. A Review of the Effect of Relaxed Clock Method, Long Branches, Genes, and Calibrations in the Estimation of Angiosperm Age. *Bot. Sci.* 92:1–22.

Magallón S., Hilu K.W., Quandt D. 2013. Land plant evolutionary timeline: Gene effects are secondary to fossil constraints in relaxed clock estimation of age and substitution rates. *Am. J. Bot.* 100:556–573.

Matschiner M. 2019. Selective Sampling of Species and Fossils Influences Age Estimates Under the Fossilized Birth–Death Model. *Front. Genet.* 10.

McCormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., Brumfield R.T. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol. Phylogenet. Evol.* 66:526–538.

McKellar R.C., Kopylov D.S., Engel M.S. 2013. Ichneumonidae (Insecta: Hymenoptera) in Canadian Late Cretaceous amber. *Foss. Rec.* 16:217–227.

Menier J.J., Nel A., Waller A., Ploëg G.D. 2004. A new fossil ichneumon wasp from the Lowermost Eocene amber of Paris Basin (France), with a checklist of fossil Ichneumonoidea s.l. (Insecta: Hymenoptera: Ichneumonidae: Metopiinae). *Geol. Acta Int. Earth Sci. J.* 2.

Milne R.I. 2009. Effects of taxon sampling on molecular dating for within-genus divergence events, when deep fossils are used for calibration. *J. Syst. Evol.* 47:383–401.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B., Ware J., Flouri T., Beutel R.G., Niehuis O., Petersen M., Izquierdo-Carrasco F., Wappler T., Rust J., Aberer A.J., Aspöck U., Aspöck H., Bartel D., Blanke A., Berger S., Böhm A., Buckley T.R., Calcott B., Chen J., Friedrich F., Fukui M., Fujita M., Greve C., Grobe P., Gu S., Huang Y., Jermin L.S., Kawahara A.Y., Krogmann L., Kubiak M., Lanfear R., Letsch H., Li Y., Li Z., Li J., Lu H., Machida R., Mashimo Y., Kapli P., McKenna D.D., Meng G., Nakagaki Y., Navarrete-Heredia J.L., Ott M., Ou Y., Pass G., Podsiadlowski L., Pohl H., Reumont B.M. von, Schütte K., Sekiya K., Shimizu S., Slipinski A., Stamatakis A., Song W., Su X., Szucsich N.U., Tan M., Tan X., Tang M., Tang J., Timelthaler G., Tomizuka S., Trautwein M., Tong X., Uchifune T., Walz M.G., Wiegmann B.M., Wilbrandt J., Wipfler B., Wong T.K.F., Wu Q., Wu G., Xie Y., Yang S., Yang Q., Yeates D.K., Yoshizawa K., Zhang Q., Zhang R., Zhang W., Zhang Y., Zhao J., Zhou C., Zhou L., Ziesmann T., Zou S., Li Y., Xu X., Zhang Y., Yang H., Wang J., Wang J., Kjer K.M., Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*. 346:763–767.

Nixon K.C., Carpenter J.M. 1993. On Outgroups. *Cladistics*. 9:413–426.

O’Hanlon S.J., Rieux A., Farrer R.A., Rosa G.M., Waldman B., Bataille A., Kosch T.A., Murray K.A., Brankovics B., Fumagalli M., Martin M.D., Wales N., Alvarado-Rybak M., Bates K.A., Berger L., Böll S., Brookes L., Clare F., Courtois E.A., Cunningham A.A., Doherty-Bone T.M., Ghosh P., Gower D.J., Hintz W.E., Höglund J., Jenkinson T.S., Lin C.-F., Laurila A., Loyau A., Martel A., Meurling S., Miaud C., Minting P., Pasmans F., Schmeller D.S., Schmidt B.R., Shelton J.M.G., Skerratt L.F., Smith F., Soto-Azat C., Spagnoletti M., Tessa G., Toledo L.F., Valenzuela-Sánchez A., Verster R., Vörös J., Webb R.J., Wierzbicki C., Wombwell E., Zamudio K.R., Aanensen

Spasojevic et al.

D.M., James T.Y., Gilbert M.T.P., Weldon C., Bosch J., Balloux F., Garner T.W.J., Fisher M.C. 2018. Recent Asian origin of chytrid fungi causing global amphibian declines. *Science*. 360:621–627.

O'Reilly J.E., Donoghue P.C.J. 2016. Tips and nodes are complementary not competing approaches to the calibration of molecular clocks. *Biol. Lett.* 12:20150975.

O'Reilly J.E., Reis M. dos, Donoghue P.C.J. 2015. Dating Tips for Divergence-Time Estimation. *Trends Genet.* 31:637–650.

PaleoBioDB. 2020. The Paleobiology Database. Available from <https://paleobiodb.org>.

Parins-Fukuchi C., Brown J.W. 2017. What drives results in Bayesian morphological clock analyses? *bioRxiv*.:219048.

Paterson R.S., Rybczynski N., Kohno N., Maddin H.C. 2020. A Total Evidence Phylogenetic Analysis of Pinniped Phylogeny and the Possibility of Parallel Evolution Within a Monophyletic Framework. *Front. Ecol. Evol.* 7.

Peters R.S., Krogmann L., Mayer C., Donath A., Gunkel S., Meusemann K., Kozlov A., Podsiadlowski L., Petersen M., Lanfear R., Diez P.A., Heraty J., Kjer K.M., Klopstein S., Meier R., Polidori C., Schmitt T., Liu S., Zhou X., Wappler T., Rust J., Misof B., Niehuis O. 2017. Evolutionary History of the Hymenoptera. *Curr. Biol.* CB. 27:1013–1018.

Philippe H., Brinkmann H., Lavrov D.V., Littlewood D.T.J., Manuel M., Wörheide G., Baurain D. 2011. Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough. *PLOS Biol.* 9:e1000602.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

- Pick K.S., Philippe H., Schreiber F., Erpenbeck D., Jackson D.J., Wrede P., Wiens M., Alié A., Morgenstern B., Manuel M., Wörheide G. 2010. Improved Phylogenomic Taxon Sampling Noticeably Affects Nonbilaterian Relationships. *Mol. Biol. Evol.* 27:1983–1987.
- Püschel H.P., O'Reilly J.E., Pisani D., Donoghue P.C.J. 2020. The impact of fossil stratigraphic ranges on tip-calibration, and the accuracy and precision of divergence time estimates. *Palaeontology*. 63:67–83.
- Puslednik L., Serb J.M. 2008. Molecular phylogenetics of the Pectinidae (Mollusca: Bivalvia) and effect of increased taxon sampling and outgroup selection on tree topology. *Mol. Phylogenet. Evol.* 48:1178–1188.
- Pyron R.A. 2011. Divergence Time Estimation Using Fossils as Terminal Taxa and the Origins of Lissamphibia. *Syst. Biol.* 60:466–481.
- Pyron R.A. 2017. Novel Approaches for Phylogenetic Inference from Morphological Data and Total-Evidence Dating in Squamate Reptiles (Lizards, Snakes, and Amphisbaenians). *Syst. Biol.* 66:38–56.
- Quicke D.L.J. 2014. *The Braconid and Ichneumonid Parasitoid Wasps: Biology, Systematics, Evolution and Ecology*. John Wiley & Sons.
- Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A., Susko E. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* 67:901–904.
- Rasnitsyn A.P., Bashkuev A.S., Kopylov D.S., Lukashevich E.D., Ponomarenko A.G., Popov Yu.A., Rasnitsyn D.A., Ryzhkova O.V., Sidorchuk E.A., Sukatsheva I.D., Vorontsov

Spasojevic et al.

D.D. 2016. Sequence and scale of changes in the terrestrial biota during the Cretaceous (based on materials from fossil resins). *Cretac. Res.* 61:234–255.

Rasnitsyn A.P., Quicke D.L. 2002. *History of Insects*. Springer Science & Business Media.

Ronquist F., Klopfstein S., Vilhelmsen L., Schulmeister S., Murray D.L., Rasnitsyn A.P. 2012a. A Total-Evidence Approach to Dating with Fossils, Applied to the Early Radiation of the Hymenoptera. *Syst. Biol.* 61:973–999.

Ronquist F., Lartillot N., Phillips M.J. 2016. Closing the gap between rocks and clocks using total-evidence dating. *Philos. Trans. R. Soc. B-Biol. Sci.* 371:20150136.

Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012b. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* 61:539–542.

Rota-Stabelli O., Telford M.J. 2008. A multi criterion approach for the selection of optimal outgroups in phylogeny: Recovering some support for Mandibulata over Myriochelata using mitogenomics. *Mol. Phylogenet. Evol.* 48:103–111.

Saladin B., Leslie A.B., Wueest R.O., Litsios G., Conti E., Salamin N., Zimmermann N.E. 2017. Fossils matter: improved estimates of divergence times in *Pinus* reveal older diversification. *Bmc Evol. Biol.* 17:95.

Sharkey M.J., Wahl D.B. 1992. Cladistics of the Ichneumonoidea (Hymenoptera). *J. Hymenopt. Res.* 1:15–24.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

- Soares A.E.R., Schrago C.G. 2012. The Influence of Taxon Sampling and Tree Shape on Molecular Dating: An Empirical Example from Mammalian Mitochondrial Genomes. *Bioinforma. Biol. Insights.* 6:BBIS9677.
- Soares A.E.R., Schrago C.G. 2015. The influence of taxon sampling on Bayesian divergence time inference under scenarios of rate heterogeneity among lineages. *J. Theor. Biol.* 364:31–39.
- Spasojevic T., Broad G.R., Bennett A.M.R., Klopstein S. 2018a. Ichneumonid parasitoid wasps from the Early Eocene Green River Formation: five new species and a revision of the known fauna (Hymenoptera, Ichneumonidae). *PalZ.* 92:35–63.
- Spasojevic T., Wedmann S., Klopstein S. 2018b. Seven remarkable new fossil species of parasitoid wasps (Hymenoptera, Ichneumonidae) from the Eocene Messel Pit. *PLOS ONE.* 13:e0197477.
- Tarrió R., Rodríguez-Trelles F., Ayala F.J. 2000. Tree Rooting with Outgroups When They Differ in Their Nucleotide Composition from the Ingroup: The *Drosophila saltans* and *willistoni* Groups, a Case Study. *Mol. Phylogenet. Evol.* 16:344–349.
- Thomas J.A., Trueman J.W.H., Rambaut A., Welch J.J. 2013. Relaxed Phylogenetics and the Palaeoptera Problem: Resolving Deep Ancestral Splits in the Insect Phylogeny. *Syst. Biol.* 62:285–297.
- Tong K.J., Duchêne S., Ho S.Y.W., Lo N. 2015. Comment on “Phylogenomics resolves the timing and pattern of insect evolution.” *Science.* 349:487–487.

Spasojevic et al.

- Travouillon K.J., Phillips M.J. 2018. Total evidence analysis of the phylogenetic relationships of bandicoots and bilbies (Marsupialia: Peramelemorphia): reassessment of two species and description of a new species. *Zootaxa*. 4378:224.
- Wahl D., Gauld I. 1998. The cladistics and higher classification of the Pimpliformes (Hymenoptera: Ichneumonidae). *Syst. Entomol.* 23:265–298.
- Wang X., Grohé C., Su D.F., White S.C., Ji X., Kelley J., Jablonski N.G., Deng T., You Y., Yang X. 2018. A new otter of giant size, *Siamogale melilutra* sp. nov. (Lutrinae: Mustelidae: Carnivora), from the latest Miocene Shuitangba site in north-eastern Yunnan, south-western China, and a total-evidence phylogeny of lutrines. *J. Syst. Palaeontol.* 16:39–65.
- Ware J.L., Litman J., Klass K.-D., Spearman L.A. 2008. Relationships among the major lineages of Dictyoptera: the effect of outgroup selection on dictyopteran tree topology. *Syst. Entomol.* 33:429–450.
- Warnock R.C.M., Yang Z., Donoghue P.C.J. 2011. Exploring uncertainty in the calibration of the molecular clock. *Biol. Lett.*:rsbl20110710.
- Welch J.J., Bromham L. 2005. Molecular dating when rates vary. *Trends Ecol. Evol.* 20:320–327.
- Wertheim J.O., Fourment M., Pond K., L S. 2012. Inconsistencies in Estimating the Age of HIV-1 Subtypes Due to Heterotachy. *Mol. Biol. Evol.* 29:451–456.
- Wheeler W.C. 1990. Nucleic Acid Sequence Phylogeny and Random Outgroups. *Cladistics.* 6:363–367.

TAXON SAMPLING IN TOTAL-EVIDENCE DATING

- Whitfield J.B. 2002. Estimating the age of the polydnavirus/braconid wasp symbiosis. *Proc. Natl. Acad. Sci.* 99:7508–7513.
- Wilberg E.W. 2015. What's in an Outgroup? The Impact of Outgroup Choice on the Phylogenetic Position of Thalattosuchia (Crocodylomorpha) and the Origin of Crocodyliformes. *Syst. Biol.* 64:621–637.
- Wright A.M., Lloyd G.T., Hillis D.M. 2016. Modeling Character Change Heterogeneity in Phylogenetic Analyses of Morphology through the Use of Priors. *Syst. Biol.* 65:602–611.
- Yu D.S.K., van Acherberg C., Horstmann K. 2016. *Taxapad 2016, Ichneumonoidea 2015*. Nepean, Ontario, Canada: .
- Zhang C., Stadler T., Klopstein S., Heath T.A., Ronquist F. 2016. Total-Evidence Dating under the Fossilized Birth–Death Process. *Syst. Biol.* 65:228–249.
- Zwickl D.J., Hillis D.M. 2002. Increased Taxon Sampling Greatly Reduces Phylogenetic Error. *Syst. Biol.* 51:588–598.

Table 1. Summary of outgroup sampling strategies.

Taxon	outgroup setting				
	“full outgroup”	“Braconidae”	“Braconidae (1)”	“Xoridinae”	“Xorides-only”
<i>Gasteruption</i> (Eva)	x				
<i>Eupelmophotismus</i> (Cha)	x				
<i>Aleiodes</i> (Bra)	x	x			
<i>Aphidius</i> (Bra)	x	x			
<i>Cotesia</i> (Bra)	x	x			
<i>Dancusa</i> (Bra)	x	x			
<i>Diaeretus</i> (Bra)	x	x			
<i>Macrocentrus</i> (Bra)	x	x			
<i>Homolobus</i> (Bra)	x	x	x		
<i>Tanychora</i> (Tan)	x ^a			x ^a	
<i>Kharsutella</i> (Tan)	x ^a			x ^a	
<i>Xorides</i> (Xor)	x	x	x	x	x
<i>Odontocolon</i> (Xor)	x	x	x	x	
<i>Aplomerus</i> (Xor)	x	x	x	x	
<i>Ischnoceros</i> (Xor)	x	x	x	x	
other non-pimpliform ichneumonids ^b	x	x	x	x	

Notes: Abbreviations in brackets stand for higher level classification of the outgroup taxa:

Eva–Evanioidea, Cha–Chalcidoidea, Bra–Braconidae, Tan–Tanychorinae, Xor–Xoridinae.

^a Included only in the “full outgroup + Tanychorinae” and “Xoridinae + Tanychorinae” analyses. ^b For a list of all included non-pimpliform ichneumonid taxa, see Supplementary File S1.

Table 2. Age estimates from the preferred analysis with credibility intervals for crown group Ichneumonidae, Pimpliformes and the pimpliform subfamilies.

Taxon Group	Median	Mean	95% Credibility Interval (millions of years)
Ichneumonidae	181.2	181.4	154.0–204.7
Pimpliformes	167.8	168.7	144.0–193.4
parasitoids of Diptera	126.2	127.8	107.3–150.0
Acaenitinae	127.6	129.0	105.1–156.7
Diplazontinae	123.5	124.8	98.9–152.0
Poemeniinae	82.7	83.5	59.1–113.4
Rhyssinae	96.8	97.7	69.6–129.9
Delomeristini	79.7	80.6	51.1–111.8
Ephialtini	103.7	105.4	82.9–128.9
Pimplini	146.7	147.4	122.1–175.1
Theroniini	80.4	80.8	54.3–106.7

Notes: As the subfamily Pimplinae was not recovered as monophyletic, age estimates for the tribes are given (Delomeristini, Ephialtini, Pimplini and Theroniini). The parasitoids of Diptera here comprise three subfamilies: Cylloceriinae, Diacritinae and Orthocentrinae (excluding Diplazontinae and the diacritine *Ortholaba*, which did not form a monophyletic group with them; note that Diacritinae are included in this clade, but there are no host records for this small subfamily).

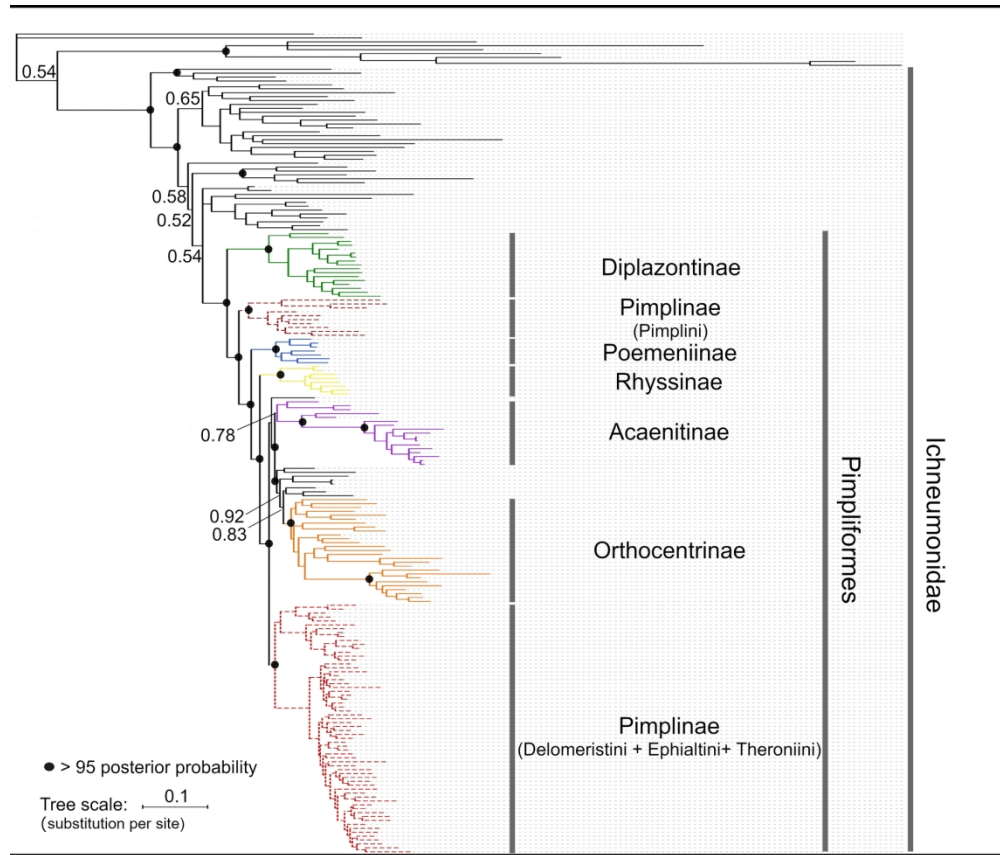


Figure 1. Majority rule consensus tree of the non-clock analysis of the "full outgroup" setting. The tree contains only extant tips, as fossils were excluded prior to summarizing the tree samples from the Bayesian analysis. Posterior probability values are given only for the nodes of interest.

165x140mm (300 x 300 DPI)

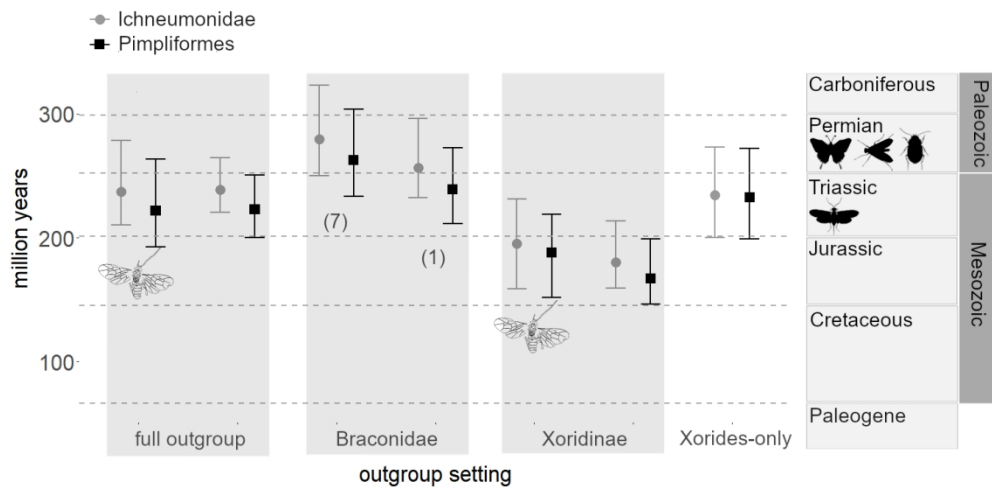


Figure 2. Age estimates in million years for Ichneumonidae and Pimpliformes across different analyses. The median and 95% credibility intervals for different outgroup settings are plotted on the y axis. Numbers in brackets indicate the number of included Braconidae species. Dashed horizontal lines indicate boundaries between geological periods. Analyses with Tanychorinae included are denoted by a fossil drawing (modified after Kopylov 2010a). The insect silhouettes denote the estimated time of radiation of the biggest orders of holometabolous insects (according to Tong et al. 2015), which are the main hosts of ichneumonids: Lepidoptera, Diptera, Coleoptera (Permian) and Hymenoptera (Triassic).

397x200mm (96 x 96 DPI)

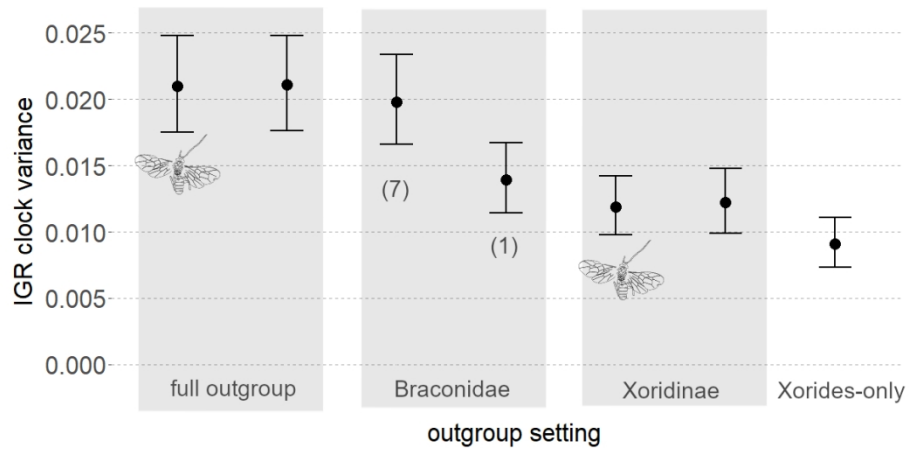


Figure 3. Variance of the relaxed-clock parameter as estimated under an independent gamma rates model (IGR). Median and 95% credibility intervals are plotted on the y axis across different outgroup settings. Numbers in brackets indicate number of included Braconidae taxa. Analyses with Tanychorinae included are denoted by a fossil drawing (modified after Kopylov 2010a).

350x159mm (96 x 96 DPI)

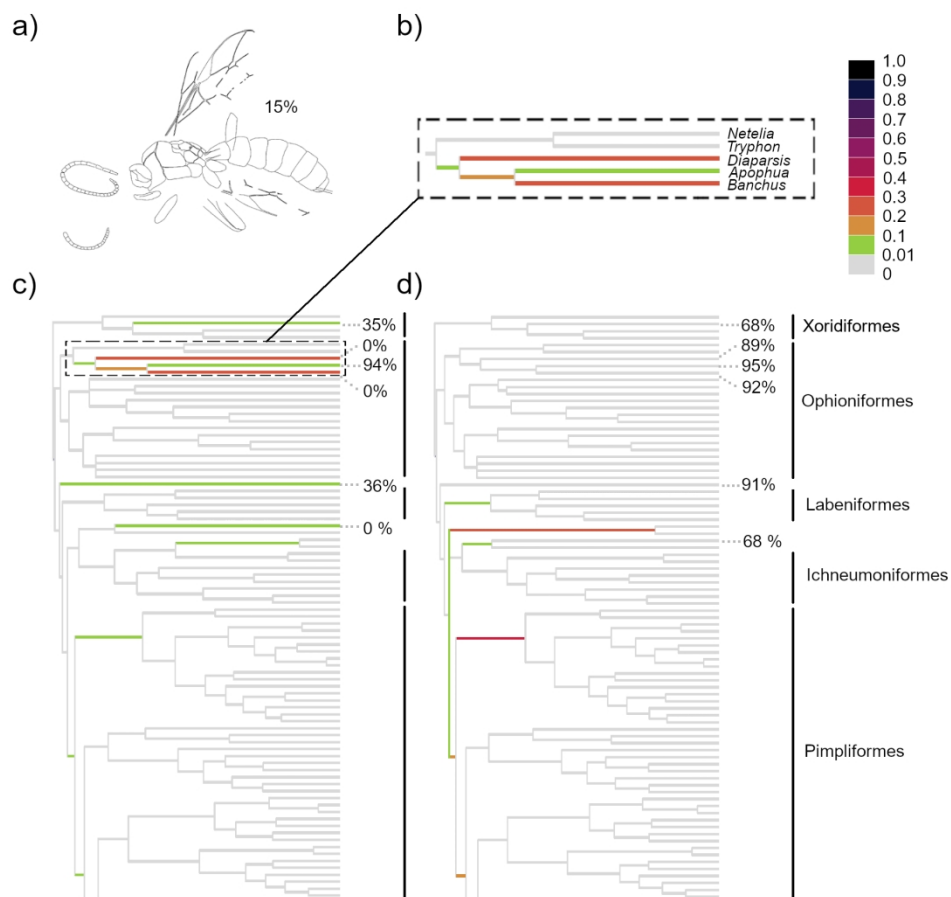


Figure 4. Placement of a) a representative of the Cretaceous fossils (*Labenopimpla kasparyan*; the image modified after Kopylov (2010b)) in the "Xoridinae" analysis. RoguePlots derived b-c) with and d) without some outgroup taxa missing morphological data ("bare taxa"). The trees represent the majority-rule consensus tree with fossils excluded. Branches are colored by probability of a fossil attaching to them. Percent values refer to the portion of scored morphological characters for a given extant taxon (c and d) and for the fossil in question (a). The remaining Cretaceous compression fossils had similar attachment patterns as the ones depicted here (Supplementary File S12).

170x159mm (300 x 300 DPI)

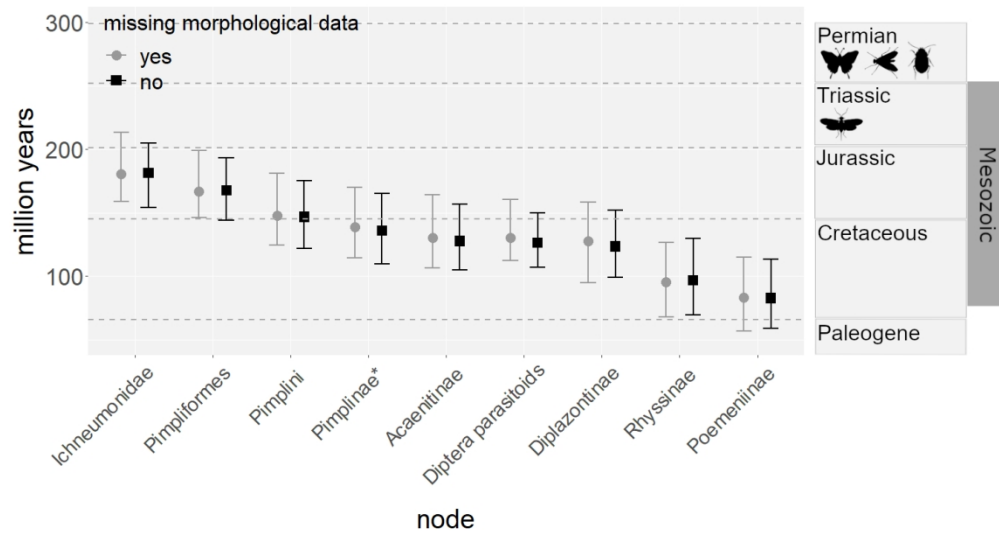


Figure 5. Age estimates in million years for the nodes of interests in the “Xoridinae” outgroup analyses with and without missing morphological data for some outgroup taxa. The median and 95% credibility intervals are plotted on the y axis. Dashed horizontal lines indicate transitions between major geological periods. The insect silhouettes denote the estimated time of radiation of the biggest orders of holometabolous insects (according to Tong et al. 2015), which are the main hosts of ichneumonids: Lepidoptera, Diptera, Coleoptera (Permian) and Hymenoptera (Triassic).* Pimplinae without the tribe Pimplini.

394x211mm (96 x 96 DPI)

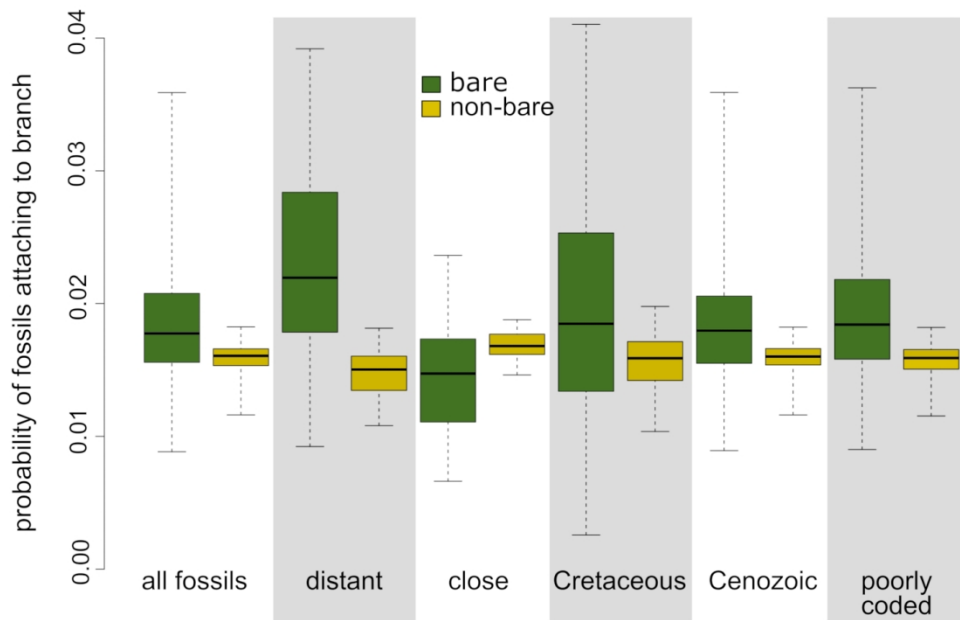


Figure 6. Bare-branch attraction in simulated data. Boxplots showing the cumulative probability of a branch leading to an extant taxon to attract a fossil, depending on whether it was bare in the simulation replicate (i.e., it was lacking morphological data) or not. The branches for which morphological data was deleted were chosen randomly for each simulation replicate. The graphs to the right show the same probability, but for subsets of the fossils: the half of the fossils that are more distant versus closer to extant taxa; Cretaceous versus Cenozoic fossils; and finally fossils with fewer than 20 characters coded plotted separately. Distance from any extant taxa thus appears as the most decisive factor here, but beware that we only obtained simulation replicates on a single tree and thus the factors concerning fossil placement are not independent.

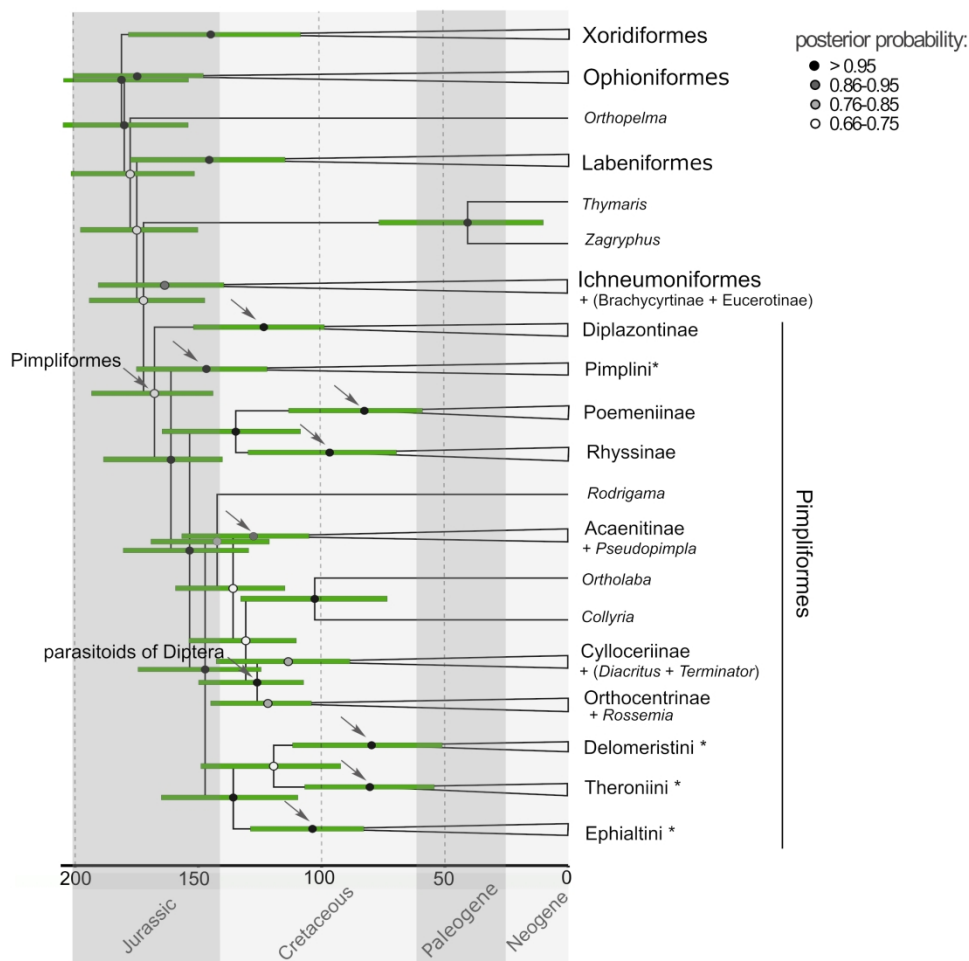


Figure 7. Dated majority-rule consensus tree from the total-evidence dating analysis under the “Xoridinae” outgroup setting with improved morphological matrix. The tree contains only extant tips, as the fossils were excluded prior to summarizing the tree samples from the MCMC analysis. Most of the clades are collapsed to depict subfamily-level relationships among Pimpliformes. Arrows indicate nodes for which age estimates are reported in Table 1. The names of the nodes are given if they do not correspond to the names of their tips. Horizontal bars represent 95% credibility intervals and corresponding circles mark median values for age estimates.