A transcriptome-based phylogeny of Scarabaeoidea confirms the sister group relationship of dung beetles and phytophagous pleurostict scarabs (Coleoptera)

Lars Dietz[1#], Matthias Seidel[2,3], Jonas Eberle[4], Bernhard Misof[1,5], Thaynara L. Pacheco[1], Lars Podsiadlowski[1], Sasanka Ranasinghe[1], Nicole L. Gunter[7], Oliver Niehuis[6], Christoph Mayer[1], Dirk Ahrens[1#]

[1]*Museum A. Koenig, Bonn, Leibniz Institute for the Analysis of Biodiversity Change (LIB), Germany*
[2]*CeNak Hamburg, Leibniz Institute for the Analysis of Biodiversity Change (LIB), Germany*
[3]*Naturhistorisches Museum Wien, Austria*
[4]*Paris-Lodron-University, Salzburg, Austria*
[5]*Rheinische Friedrich-Wilhelms-Universität Bonn, Germany*
[6]*Institute of Biology I (Zoology), University of Freiburg, Germany*
[7]*The Cleveland Museum of Natural History, United States*

[#] Authors equally contributed to this study

Corresponding author: Dirk Ahrens
Email: ahrens.dirk_col@gmx.de

Abstract

Scarab beetles (Scarabaeidae) are a diverse and ecologically important group of angiosperm-associated insects. As conventionally understood, scarab beetles comprise two major lineages: dung beetles and the phytophagous Pleurosticti. However, previous phylogenetic analyses have not been able to convincingly answer the question whether or not the two lineages form a monophyletic group. Here we report our results from phylogenetic analyses of more than 4,000 genes mined from transcriptomes of more than 50 species of Scarabaeidae and other Scarabaeoidea. Our results provide convincing support for the monophyly of Scarabaeidae, confirming the debated sister group relationship of dung beetles and phytophagous pleurostict scarabs. Supermatrix-based maximum likelihood and multispecies

coalescent phylogenetic analyses strongly imply the subfamily Melolonthinae as currently understood being paraphyletic. We consequently suggest various changes in the systematics of Melolonthinae: Sericinae Kirby, 1837 stat. rest. and sensu n. to include the tribes Sericini, Ablaberini and Diphucephalini, and Sericoidinae Erichson, 1847 stat. rest. and sensu n. to include the tribes Automoliini, Heteronychini, Liparetrini, Maechidiini, Scitalini, Sericoidini, and Phyllotocini. Both subfamilies appear to consistently form a monophyletic sister group to all remaining subfamilies so far included within pleurostict scarabs except Orphninae. Our results represent a major step towards understanding the diversification history of one of the largest angiosperm-associated radiations of beetles.

Keywords. Scarabaeidae, phylogeny, reclassification

Introduction

The evolution of large parts of extant terrestrial biodiversity has been driven by the evolutionary success of angiosperm plants; these radiations have been linked to increased productivity and growth rates of angiosperm vegetation (de Boer et al. 2012), the rise of ectomycorrhiza enhancing chemical weathering of soils (Taylor et al. 2011, 2012), and the promotion of soil nutrient release by angiosperm litter that is easily decomposed (Berendse & Scheffer 2009). While the diversification of many insects, and especially that of beetles, was directly or indirectly fostered by that of angiosperms (Hunt et al. 2007; Ahrens et al. 2014; McKenna et al. 2019), the evolutionary mechanisms and timescales of angiosperm-dependent radiations have remained poorly understood, as the phylogenetic relationships of many lineages remained insufficiently known. This is especially true for scarab beetles (Scarabaeidae), which represent a diverse lineage of beetles feeding predominantly on either angiosperm plants or mammal dung. Although traditionally grouped into a single family (Scholtz & Grebennikov 2005), it is divided into two major lineages: 1) the plant-feeding lineage Pleurosticti, which include, for example, rose chafers, rhinoceros beetles, and Christmas beetles, and 2) a clade of taxa feeding on mammal dung (Aphodiinae + Scarabaeinae).

Molecular phylogenetic analyses of the Scarabaeidae have been controversial, with only nine out of 21 recently published studies reporting Scarabaeidae being

monophyletic (Table 1). Scarabaeidae are part of a wider clade (Scarabaeoidea) which also includes Lucanidae (stag beetles), Geotrupidae (earth-boring dung beetles) and several other families (Scholtz & Grebennikov 2005), and the monophyly of this superfamily has been confirmed by all major molecular studies (e.g., Zhang et al 2018, McKenna et al. 2019). The current classification of Scarabaeoidea (Scholtz & Grebennikov 2005) is founded on morphological evidence (Lawrence & Newton 1995). Yet, despite extensive research on the morphology of Scarabaeoidea by Browne & Scholtz (1998, 1999), the monophyly of Scarabaeidae has yet to be adequately tested. The analysis of Browne & Scholtz (1998) included only lineages of “Scarabaeidae” which were rooted with a single outgroup, while the analysis of Browne & Scholtz (1999) coded “Scarabaeidae” as a single terminal taxon. A recent cladistic analysis based on morphology in a wider systematic framework (Lawrence et al. 2011) did not recover Scarabaeidae as a monophyletic group.

The uncertainty surrounding the monophyly of Scarabaeidae prompted discussions about the classification of the family (Kohlmann & Morón 2003). Apart from many unresolved issues of classification within the group, the question arose whether to split scarab dung beetles and pleurostict scarabs (Erichson 1848) in two or more families (Morón 1997; Cherman & Morón 2014). Prior to the era of molecular phylogenies, there has been an impressive array of classification schemes on Scarabaeidae and related families (Jansens 1949; Crowson 1955; Balthasar 1963; Endroedi 1966; Medvedev 1976; Iablokoff-Khnzorian 1977; Howden 1982; Lawrence & Newton 1982, 1995; Paulian & Baraud 1982; Morón 1984, 1997, 2003; Scholtz 1990; Browne & Scholtz 1995; Nikolaev 1995; Kohlmann & Morón 2003; Ratcliffe & Jameson 2004; Scholtz & Grebennikov 2005; Smith 2006; Smith et al. 2006). As robust and convincing evidence is lacking for most of these classification schemes, many systematic and taxonomic studies arbitrarily followed one of these classifications according to the author’s opinion or geographic provenience (e.g., Cherman & Morón 2014).

Here, we readdress the question about the controversial sister group relationship of scarab dung beetles and pleurostict scarabs (Ahrens et al. 2014), which is fundamental to gain a more complete understanding of the evolutionary impact of Angiosperms on the diversification of Scarabaeidae. If the monophyly of Scarabaeidae is confirmed, then it raises the question of why angiosperms and their

related follow-up radiations (e.g., that of large herbivore mammals; see Ahrens et al. 2014) seemingly had a more significant impact on the radiation of this lineage than on any other scarabaeoid beetle lineage. To this end, we expand the phylogenomic data set compiled by McKenna et al. (2019) within additional taxa of Scarabaeoidea. The expanded taxonomic sampling is used to assess which aspects of the data could result in the inference of incompatible topologies (e.g., Zhang et al. 2018; Cai et al. 2022).

Material and methods

Taxon sampling and new transcriptome data

We analyzed 57 transcriptomes that covered almost all major families of Scarabaeoidea and two outgroup taxa (Table S1). Fifteen of these transcriptomes had been published by McKenna et al. (2019). Two transcriptomes were sequenced in context of the 1KITE project but had not been analyzed and published before. For details relating to the extraction of total mRNA and fragmentation, construction of cDNA libraries, and tagging of these two datasets see Peters et al. (2017).
40 additional transcriptomes were specifically generated in context of this study (Table S1).
All mRNA libraries were sequenced with Illumina HiSeq 2000 sequencers (Illumina, San Diego, CA, USA), using paired-end 150-bp reads.
Extraction of RNA, complementary deoxyribonucleic acid (cDNA) library construction, library normalization, and Illumina sequencing were carried out by a commercial sequencing company (Starseq, Mainz, Germany).
All raw nucleotide sequences are deposited at the National Center for Biotechnology Information (NCBI), Sequence Read Archive (see Table S1 for accession numbers).
The raw data of previously published transcriptomes of 17 Scarabaeoidea (see above) and of two outgroup taxa (*Ocypus brunnipes* (Staphylinidae) and *Helophorus nanus* (Helophoridae)) (McKenna et al. 2019) were downloaded from NCBI. All raw nucleotide reads from both newly sequenced and published transcriptomes were trimmed with TrimGalore 0.6.6 (Krueger et al. 2021) and assembled with Trinity 2.11.0 (Grabherr et al. 2011) using the software's default settings. Transcriptome

assemblies (see Table S1 for accession numbers) are deposited at the Transcriptome Shotgun Assembly (TSA) Database, NCBI Bioproject ID PRJNA906571 for newly sequenced transcriptomes or PRJNA936991 for re-assemblies of published transcriptomes (http://www.ncbi.nlm.nih.gov/bioproject). The assemblies were filtered with a custom Perl script (trinity_longest_d.pl, see Supplement) to retain only the longest isoform per locus, as loci with multiple isoforms could otherwise be falsely discarded as paralogs in the gene orthology assessment step.

Data extraction and alignment

Tab-delimited files were downloaded from the OrthoDB10 database (www.orthodb.org) for all groups of orthologous genes (= ortholog groups) at the hierarchical level Coleoptera that were present in at least eight of the nine coleopteran genomes in the database and single-copy in all of them. This included a total of 4,296 genes. This is similar to the principle of universal single-copy orthologs (USCOs) as used by the program BUSCO (Simão et al. 2015). However, USCOs have to be present and single-copy in at least 90% of all known genomes of a taxonomic group. In this case, this would mean that the genes would have to be present in all nine annotated coleopteran genomes available at the time. To avoid excluding genes that may be absent simply due to the incompleteness of one of the nine genome assemblies, we decided to include genes present in only eight of the nine coleopteran genomes.

Furthermore, we downloaded the official gene sets (OGS) for all nine available coleopteran genomes from OrthoDB. Tab-delimited files were modified for use in Orthograph 0.7.1 (Petersen et al. 2017) and used together with the OGS, to create a SQLite database of the genetic information with that program. Hidden Markov models (HMMs) were created with Orthograph from the available amino acid sequences of each ortholog group. HMMs were used then to extract the target genes from the filtered Trinity contigs of each specimen with Orthograph, using the software's default settings.

USCO nucleotide and corresponding amino acid sequences that were identified and inferred with Orthograph were aligned in two different ways. We first aligned the inferred amino acid sequences against the HMMs from Orthograph with hmmalign

(part of the HMMER 3.3 package; Eddy 2011; http://eddylab.org/software/hmmer/hmmer.org). The amino acid alignment was then used as a blueprint to align the corresponding nucleotide sequences with pal2nal 14.1 (Suyama et al. 2006). Alignment regions not covered by the HMMs were removed with a custom Perl script (hmmalign_cut2_d.pl, Supplement Files). We additionally aligned the amino acid sequences with MAFFT 7.305b (Katoh & Standley 2013) using the L-INS-i algorithm. The corresponding nucleotide sequence alignments were inferred using again the software pal2nal. Poorly aligned regions in the amino acid sequence alignments were identified with ALISCORE 2.0 (Misof & Misof 2009; Kück et al. 2010) and removed from the amino acid sequence alignments and corresponding nucleotide sequence alignments with the software ALICUT 2.31 (available from: https://github.com/PatrickKueck/AliCUT). Outlier sequences were identified and removed with the software OliInSeq 0.9.3 (https://github.com/cmayer/OliInSeq) using the software's default parameters. As the third codon position is typically hypervariable and often exhibits inhomogeneous nucleotide frequencies, we also used a custom Perl script (extract_codpos_d.pl, see Supplement) to generate nucleotide sequence alignments in which the third codon position was removed.

To test the monophyly of Scarabaeidae with data from another dataset, we downloaded the OrthoDB10 Endopterygota data from BUSCO 4.0.6 (Simão et al. 2015; Manni et al. 2021), including Hidden Markov Models (HMMs) and information files, from the BUSCO website (busco.ezlab.org). This set comprises 2,120 genes that are present in single copy in at least 90% of all known genomes of Endopterygota (hereafter referred to as Endopterygota USCOs). The OGS for all 56 species in that dataset were downloaded from OrthoDB and used to create an SQLite database with Orthograph. Together with the HMMs from BUSCO, the information was used to extract the Endopterygota USCO genes of each specimen with Orthograph as described above.

Phylogenetic tree inference

We inferred phylogenetic trees from the Coleoptera and Endopterygota data sets, analyzing the amino acid sequence alignments (AA) generated with hmmalign and MAFFT and the corresponding nucleotide sequence alignments (NT) inferred with pal2nal. We then considered two sets of nucleotide sequence alignments: those that

included all three codon positions (NT123) and those that include only first and second codon positions (NT12). The multiple sequence alignments were analyzed using coalescent-based and concatenation-based tree inference methods. For conducting the concatenation-based phylogenetic analyses, the multiple sequence alignments of a given type (i.e., amino acid, nucleotide) of all genes were concatenated with a custom Perl script (concat_eogs_part_d.pl, see Supplement). The resulting super-alignments were then analyzed with IQ-TREE 2.1.2 (Minh et al. 2020). The datasets were partitioned by gene. The best-fitting model and partitioning scheme were inferred with ModelFinder, using the IQ-TREE option -m MFP+MERGE (Kalyaanamoorthy et al. 2017; Chernomor et al. 2016). Branch support was assessed with approximate likelihood ratio tests (aLRT) and via ultrafast bootstrapping (Hoang et al. 2018) applying 1,000 replicates and nearest neighbor interchange (NNI) as tree rearrangement method. For conducting coalescent-based phylogenetic analyses, we first calculated phylogenetic trees of each gene with IQ-TREE, determining the best-fitting model with ModelFinder. The resulting gene trees were used for analyses with ASTRAL 5.6.1 (Zhang et al. 2018) that we used to conduct the coalescent-based phylogenetic analysis. All trees were rooted with *Helophorus* (Helophoridae) and *Ocypus* (Staphylinidae) as outgroups.

To test the effect of missing data, we generated reduced versions of all nucleotide and amino-acid datasets in which positions with a taxon coverage of less than 70% were removed with a custom Perl script (removegaps_d.pl, Suppl. File 5). The previously described phylogenetic analyses were repeated using these reduced datasets.

Furthermore, we examined the effect of varying substitution rates between different genes. For this test, we divided datasets containing at least 70% complete positions into sets of fast and of slowly evolving genes. Using a custom script (pairwise_id2.pl, see Supplement Files) we calculated the pairwise sequence identity within each gene alignment according to Sharma et al. (2014). We then divided both the hmmalign- and MAFFT-inferred multiple sequence alignments of individual genes into two sets with high and with low pairwise sequence identity, each including 50% of the genes. We then conducted concatenation- and coalescent-based analyses on these sets using the same methods as described above.

Topology tests

To assess support for the monophyly of Scarabaeidae in our datasets, we conducted a number of topology tests on the six different datasets (i.e., NT12, NT123 and AA, inferred using hmmalign or MAFFT). First, we computed the likelihood scores of maximum likelihood trees constrained to support different topologies for each dataset with IQ-TREE. The constrained trees covered all possible phylogenetic relationships of the following four taxa: Glaphyridae, Hybosoridae, Scarabaeinae + Aphodiinae, Pleurosticti. We compared the likelihood of the constraint trees using a variety of resampling tests in IQ-TREE using RELL approximation (Kishino et al. 1990) with 10,000 iterations, including bootstrap proportion (BP), Kishino-Hasegawa test (KH; Kishino & Hasegawa 1989), Shimodaira-Hasegawa test (SH; Shimodaira & Hasegawa 1999), expected likelihood weights (ELW; Strimmer and Rambaut 2002), and the approximately unbiased (AU) test (Shimodaira 2002).

We used the same strategy to assess support for all possible phylogenetic relationships between Dynastinae and the genera *Anomala* and *Adoretus* (both part of the potentially paraphyletic Rutelinae) as well as between Melolonthinae s. str., Hopliini, and Cetoniinae + Rutelinae + Dynastinae (hereafter referred to as CRD).

We further tested the monophyly of Scarabaeidae using four-cluster likelihood mapping (Strimmer and von Haeseler 1997) in IQ-TREE. For this purpose, we divided the taxon set into Hybosoridae, Scarabaeinae + Aphodiinae, Pleurosticti, and all others. All 4,410 unique quartets containing one taxon of each quartet were tested for their support for the three possible four-taxon trees.

Results

Data completeness

Nucleotide sequences from all but one of 4,296 Coleoptera-specific genes and from all 2,120 Endopterygota USCO genes were successfully recovered. In total, the dataset of the Coleoptera-specific genes aligned with hmmalign comprised 6,960,192 nucleotide positions, of which 3,670,744 were parsimony informative. The overall alignment completeness was 62.7%. The corresponding MAFFT-aligned dataset comprised 3,929,520 nucleotides, of which 2,186,562 were parsimony informative. The alignment completeness was 79.6%. The dataset of Endopterygota USCO

genes comprised 2,337,876 nucleotides aligned with hmmalign, of which 1,126,272 were parsimony informative. The alignment completeness was 62.8%. The corresponding dataset aligned with MAFFT was 1,649,496 nucleotide positions long, of which 893,640 were parsimony informative. The alignment completeness was 82.2%.

Phylogenetic analyses

The choice of the alignment software (hmmalign vs. MAFFT) had little impact on the results of the phylogenetic analyses (Fig. 1; Suppl. Figs 1 and 2), and neither did the presence of incomplete alignment sites. However, the choice of the alignment software had an impact on the inferred position of Passalidae (Suppl. Figs 3 and 4). Both sets of genes yielded very similar phylogenies, except that in trees based on the smaller Endopterygota USCO dataset, the single species of passalid included in our analyses was consistently placed as sister to Glaphyridae + Hybosoridae + Scarabaeidae, while in the larger Coleoptera-specific set, its position differed in the various phylogenetic analyses. Including only fast or slowly evolving genes did not lead to noteworthy consistent differences in topology, although support values were generally lower than with the complete gene set. The most significant topological differences were found between trees inferred from datasets that included all three codon positions (NT123) and between trees inferred from datasets which included only nucleotides of the first and second codon position and those that were analyzed on the amino acid level. Whether a concatenation-based and coalescent-based approach was used to analyze the data had little impact on the phylogenetic results, except on the placement of the outgroup. All datasets yield well-resolved trees, in which most splits received maximal support.

Scarabaeidae are strongly supported as being monophyletic by all datasets except the nucleotide datasets that include the 3rd codon position (NT123) (Fig. 1). Analysis of the latter suggested Hybosoridae as the sister group of pleurostict Scarabaeidae. All remaining scarabaeoid families that were represented by more than one species in our datasets were consistently supported as being monophyletic.

Phylogenetic analysis of the NT12 and AA supermatrices suggested Lucanidae is the sister group of Glaresidae + Trogidae. They furthermore suggested Geotrupidae + Bolboceratidae is the sister group of a clade that comprised Glaphyridae,

Hybosoridae, and Scarabaeidae. Passalidae was placed in various positions within that group depending on the specific analysis. Hybosoridae was found as the sister group of Scarabaeidae. The monophyly of Scarabaeidae was always maximally supported by our data (Fig. 1; Suppl. Figs 1-2). Coalescence-based summary trees of all gene trees differed from the supermatrix-based trees in suggesting Geotrupidae + Bolboceratidae being the sister group of Lucanidae, Glaresidae, and Trogidae. However, some the coalescence-based summary trees inferred from exclusively slowly evolving genes showed the same topology as the concatenation-based trees. The two subgroups of Scarabaeidae, Aphodiinae + Scarabaeinae and Pleurosticti were always strongly supported as being monophyletic, as were the two subfamilies, Aphodiinae and Scarabeinae. Within Pleurosticti, *Orphnus* (Orphninae) was consistently found as sister group of the remaining pleurostict scarabaeid lineages, that were divided into two major clades: the first clade comprised 1) Sericoidinae — Australasian and Neotropical taxa referred to by some authors as Southern World Melolonthinae (Ahrens & Vogler 2008; Ahrens et al. 2011, 2014; Šípek et al. 2016) or Liparetrinae (Lacroix 2007, 2014; Eberle et al. 2019; Pacheco et al. 2022) — and 2) Ablaberini + Sericini + the Australian genus *Diphucephala* (i.e., Diphucephalini), which represented the sister group of the two former tribes. The second clade comprised *Pachypus* (Pachypodini) and three major lineages, which *Pachypus* was found to be sister group of. The phylogenetic relationships among the latter three lineages differed among the inferred trees. These three lineages were: 1) Hopliini, 2) Melolonthini, including the genus *Sparrmannia* (currently placed within the probably polyphyletic Tanyproctini; Eberle et al. 2019), and 3) Cetoniinae + (Dynastinae + Rutelinae). Supermatrix-based phylogenetic analyses consistently provided support for Melolonthini being the sister group of the remaining two lineages. Coalescence-based analysis were inconsistent in regard of the phylogenetic arrangement of the three lineages. While the monophyly of Dynastinae + Rutelinae was consistently well supported, a monophyly of Rutelinae was strongly supported only in concatenation-based trees. In coalescence-based trees, the grouping was not consistently found, and if so, it typically received only low support — a pattern also confirmed by results of previous studies in which monophyly of Rutelinae did not result (e.g., Ahrens & Vogler 2008; Ahrens et al. 2014; Šipek et al. 2016; Neita-Moreno et al. 2019). With Cetoniinae + (Dynastinae + Rutelinae) consistently found nested within the lineages so far classified as Melolonthinae, our study confirmed the paraphyly of

Melolonthinae (Ahrens & Vogler 2008; Ahrens et al. 2014; Gunter et al. 2016; Šipek et al. 2016; Neita-Moreno et al. 2019; McKenna et al. 2019).

Topology tests

In tests assessing the support for monophyly of Scarabaeidae, the originally inferred topology (i.e., monophyly of Scarabaeidae) consistently received the highest support when analyzing the NT12 and AA supermatrices. Likewise, the monophyly of Hybosoridae + Pleurosticti consistently received the highest support when analyzing the NT123 datasets. Alternative trees were rejected with $p < 0.001$ (Table 2). The topology tests supported the sister group relationship of Hopliini and the CRD clade, rejecting alternative topologies with $p < 0.001$. Likelihood tests testing the monophyly of Rutelinae supported the latter irrespective of what dataset we analyzed, but often only with p between 0.01 and 0.05 (Tables 3, 4).

Four-cluster likelihood mapping revealed support for a monophyly of Scarabaeidae when analyzing the NT12 and AA datasets (Fig. 3). When analyzing the NT12 datasets, a monophyly of Scarabaeidae was supported by 55–60% of the quartets. When analyzing the AA datasets, the support was > 80%. However, it should be noted that in all analyses almost all quartets that include the most remotely related outgroups *Helophorus* and *Ocypus* supported Hybosoridae + Pleurosticti, while among those containing the closest outgroup *Eulasia* (Glaphyridae), more than 90% supported monophyly of Scarabaeidae, even when analyzing the NT12 datasets. When analyzing the NT123 datasets, more than 70% of quartets supported Hybosoridae + Pleurosticti.

Discussion

The interfamilial results found in our study agree well with those of other multi-gene phylogenetic studies with less taxon and/or gene sampling (i.e., McKenna et al. 2019; Zhang et al. 2018; Cai et al. 2022). One exception is the position of Passalidae, which other studies found to be phylogenetically closely related to Geotrupidae + Bolboceratidae (McKenna et al. 2019; Beza-Beza et al. 2020). We found this phylogenetic position only in a few of our trees. In the majority of the inferred trees, we found Passalidae to be closely related to Scarabaeidae. As Passalidae had been

represented by only few taxa on relatively long branches in all phylogenetic studies (including ours) conducted so far, a definitive statement on the phylogenetic position of Passalidae must await broader taxonomic sampling of the family itself and of related clades.

Our results confirm the early divergence of Orphninae within Pleurosticti, which had remained unresolved in some earlier molecular studies (e.g., Ahrens et al. 2014). However, an early divergence had been suspected based on morphological evidence (e.g., Browne & Scholtz 1998; Ahrens 2006). For example, Orphninae lack the derived conformation of spiracles after which Pleurosticti is named (Erichson 1848).

The results of our study confirmed the sister group relationship of dung feeding scarabs and phytophagous pleurostict scarabs (i.e., monophyly of Scarabaeidae), which was not always found in previous studies with limited gene sampling (see Table 1). We found monophyly of Scarabaeidae with both supermatrix and coalescent-based tree reconstruction approaches. However, tree reconstruction results based on the nucleotide sequence data heavily depended on whether or not the 3rd codon position was included — a phenomenon frequently observed in phylogenomic studies (Li et al. 2014).

The results of the likelihood mapping analyses strongly suggested that the non-monophyly of Scarabaeidae in the analyses of the NT123 data is an artifact. As this result is suggested primarily by quartets that include distantly related outgroup taxa, it can possibly be explained by long-branch attraction between the outgroup and the relatively long-branched dung-feeding scarabs. Likewise, long-branch attraction between the CRD clade and outgroups may have led to an artificial grouping of Melolonthinae and Hopliini.

The inferred monophyly of Scarabaeidae supports the initial hypothesis that the radiation of angiosperms had primarily affected species diversity, diversity of feeding habits, and morphological disparity of only a single lineage of Scarabaeoidea. This sheds new light on possible causes for their successful diversification, which might be highly lineage related, possibly also in regard to genome-driven events (e.g., McKenna et al. 2019). The successful diversification of other major herbivorous beetle lineages, such as the Phytophaga (Curculionoidea + Chrysomeloidea), was attributed to the genomic presence of plant cell wall-degrading enzymes obtained from bacteria and fungi (McKenna et al. 2019). However, the presence of PCWDEs in Scarabaeoidea was only limited to GH1 and GH9 that were expected to occur in

most beetle species (McKenna et al. 2019), therefore other mechanisms are likely responsible for the successful radiation of Scarabaeidae. Another important factor that promoted diversification with angiosperm plants could be the efficient management of endosymbionts by the species of Scarabaeidae. As one key innovation in this regard can be seen the development of female accessory glands at the end of the digestive and genital duct, which are present in Aphodiinae and pleurostict scarabs (Ahrens 2006).

Accessory glands are known to have an important function in the transmission of endosymbiont bacteria that are known, beyond their active part in cellulose digestion (Martin 1983; Martin et al. 1991), to assist the production of pheromones (Hoyt et al. 1971). The latter could have had an important impact on the improvement of chemical communication in these groups (Pacheco et al. 2022). It should be noted that accessory glands are absent in Scarabaeinae, but other mechanisms for transmission of endosymbiotic bacteria have been documented (Estes et al. 2013). Nevertheless, the confirmed monophyly of Scarabaeidae allows us to interpret the reduction of accessory glands in Scarabaeinae dung beetles as a true loss and not as a parallelism between pleurostict scarabs and Aphodiinae.

The diverse Mesozoic fossil record of Hybosoridae (Lu et al. 2022) still provides important information for dating the onset of scarab divergences, given the poor fossil record of the Mesozoic Scarabaeidae (Krell 2007), particularly in amber which often allows a more accurate classification and a more robust systematic placement of its fossils. While we confirm the sister-group relationship between Scarabaeidae and Hybosoridae, the limited taxonomic sampling of this study restricts our ability to accurately place most of the known fossils into the inferred phylogenetic tree. Therefore, we explicitly refrained from inferring divergence time estimates.

Implications on the classification of Scarabaeidae

Our phylogenetic analyses on Scarabaeidae revealed that there is — from a phylogenetic point of view — no necessity to split Scarabaeidae into two families. Given the monophyly of dung beetles and phytophagous pleurostict scarabs, such a splitting (Cherman & Morón 2014) would be rather arbitrary and in contrary to the aim of maintaining a stable nomenclature and classification, which are important

backbones of all biodiversity-related databases (e.g., NCBI, GBIF). It should be noted, though, that our study did not include some taxonomic lineages (e.g., Ochodaeidae, Eremazinae) that earlier studies found (although with poor support) within a clade which included Glaphyridae, Hybosoridae, scarab dung beetles, and pleurostict Scarabaeidae (Ahrens et al. 2014; Neita-Moreno et al. 2019). Interpretations for classification and evolution of the Scarabaeidae will be thus more robust when the position of these groups is better known.

The phylogenetic tree hypothesis supported by our study (Fig. 1) shows some examples of non-homogenous lineage classifications, in which sister taxa, according to the current classification, are either classified as tribes or as subfamilies (e.g., Pachypodini vs. Melolonthinae vs. Rutelinae/ Dynastinae/ Cetoniinae vs. Hopliini). Given the non-monophyly of Melolonthinae as currently understood (Smith 2006; Bouchard et al. 2011), the question arises to which clade the name “Melolonthinae” should be referred, with respective modifications to the current classification. Many unanswered questions remain due to the limited sampling here and contradictory tree topologies compared to and between previous studies (e.g., Ahrens et al. 2014; Eberle et al. 2019).

The clear phylogenetic separation of the monophyletic “Southern World” Melolonthinae and the clade containing Sericini, Ablaberini, and Diphucephalini from the remaining pleurostict scarabs (i.e., Cetoniinae, Rutelinae, Dynastinae, Melolonthinae etc.) makes it reasonable to treat both lineages as separate subfamilies. Both lineages are currently classified as a series of tribes within Melolonthinae (Smith 2006). A revised classification that elevates these clades to subfamilies would alleviate the problem, at least in part, of rendering Melolonthinae polyphyletic. It would also help to focus on the problem of whether Melolonthinae are monophyletic under inclusion of Hopliini and Macrodactylini (the latter not included in taxonomic sampling of this study). In regard to the sister group relationship Hopliini + Melolonthini, based on the current sampling there is good support that they do not form a monophyletic group (Table 4) although due to the limited sampling our results should not be regarded as fully conclusive.

Currently, the clade Melolonthini, which we refer here to a restricted interpretation of the subfamily “Melolonthinae”, includes several lineages currently circumscribed as subtribes, such as Rhizotrogina, Pyglina, Melolonthina, Enariina, Schizonychina, Leucopholina. It also includes several other minor lineages (Eberle et al. 2019) and

could at least contain the genus *Sparrmannia*. The latter is so far assigned to Tanyproctini but the position of other genera of the polyphyletic Tanyproctini remains yet uncertain (Eberle et al. 2019). A restricted Melolonthinae (see Fig. 1) would be a starting point for a re-classification that would allow retaining well-established subfamily names (e.g., Dynastinae, Rutelinae, Cetoniinae). However, the exact extent of this clade Melolonthinae is yet to be identified, particularly with reference to the other lineages so far classified as "Melolonthinae" (e.g., Diplotaxini, Hopliini, Macrodactylini; e.g., Ahrens et al. 2014). To further address this topic, the taxonomic sampling needs to be extended, also to allow more robust statistical topology testing (Table 2-4; Fig. 3). The same applies to Rutelinae + Dynastinae: the monophyly of Rutelinae with respect to Dynastinae recovered by our results was often not supported in other studies (e.g., Ahrens & Vogler 2008; Ahrens et al. 2014; Šipek et al. 2016; Neita-Moreno et al. 2019), although Guo et al. (2022) also found a clade of Adoretini + Anomalini to the exclusion of Dynastinae using mitochondrial genomes, similar to our results. However, as our datasets contained only one representative each of two of the seven currently recognized tribes of Rutelinae, a decision on the classification of these two subfamilies should await further studies.

The following formal classification changes are proposed. Sericinae Kirby, 1837 **stat. rest.** and **sensu n.** is re-elevated to subfamily and revised to include the tribes Sericini, Ablaberini and Diphucephalini. Sericoidinae Erichson, 1847 **stat. rest.** and **sensu n.** is re-elevated to subfamily and revised to include the tribes Sericoidini, Liparetrini, Maechidiini, Heteronychini, Scitalini and Phyllotocini. Sericoidinae Erichson, 1847 has formally priority over the younger name Liparetrinae Burmeister, 1855 (and other tribal names within the lineage) and also includes the tribe Automoliini not included in this analysis but being confirmed in previous molecular phylogenies to be part of the same lineage (Ahrens & Vogler 2008; Ahrens et al. 2014). Based on morphological characters, other candidate members of this subfamily are Colymbomorphini, Comophorinini, Pachytrichini, and Phyllotocidiini, however, their phylogenetic placement has not been confirmed yet with molecular data.

Apart of some of our robust results regarding the monophyly of Scarabaeidae, we consider this work also as a primer and starting point for further and more detailed phylogenomic research in Scarabaeoidea. In this, the generated transcriptomic data will serve as backbone for other approaches such as DNA target enrichment

approaches, which would possibly allow to considerably extend the taxon sampling. Because also dry museum specimens could be analyzed this way, we expect that even yet entirely obscure lineages (e.g., Dynamopodinae, Phaenomeridinae, Belohinidae), which never have been considered in any phylogenetic analysis, will find their place in a phylogenetic tree.

Supplementary Material

Table S1. Overview on samples included in this study, referring to taxon name, current systematic placement, collection site, voucher number, accession numbers of sequence data repositories. The table also includes information on the number of recovered genes (Endopterygota USCOs vs. Coleoptera-specific genes), alignment completeness [in %] (Endopterygota USCOs, hmmalign vs. MAFFT vs. Coleoptera-specific genes, hmmalign vs. MAFFT).

Table S2. Likelihood tests regarding the monophyly of different groupings using the Endopterygota USCO dataset, with constraint trees using a variety of resampling tests in IQ-TREE using the RELL approximation including bootstrap proportion (BP), Kishino-Hasegawa test (KH), Shimodaira-Hasegawa test (SH), expected likelihood weights (ELW), and the approximately unbiased (AU) test.

Supplement Files 1-12 (see Dryad, link: xxx; to be provided)

Supplement File 1: Perl script to filter Trinity contigs for the longest variant. (trinity_longest_d.pl).

Supplement File 2: Perl script to remove unaligned regions from hmmalign output. (hmmalign_cut2_d.pl).

Supplement File 3: Perl script to remove third codon positions from an alignment. (extract_codpos_d.pl).

Supplement File 4: Perl script to concatenate gene alignments into a partitioned superalignment. (concat_eogs_part_d.pl).

Supplement File 5: Perl script to remove alignment positions present in less than a given number of taxa. (removegaps_d.pl).

Supplement File 6: Perl script to calculate average pairwise identities between sequences within alignments. (pairwise_id2.pl).

Supplement File 7: Coalescent-based phylogenetic trees from ASTRAL analysis (astral_trees.zip) (see astral-trees-descriptions.txt for descriptions).

Supplement File 8: Concatenated sequence alignments of coding sequences of all loci in a dataset in FASTA format (concat_alignments.zip) (see concat-alignments-descriptions.txt for descriptions).

Supplement File 9: Maximum-likelihood phylogenetic trees created with IQ-TREE based on concatenated alignments of coding sequences of USCO loci (concat_trees.zip) (see concat-trees-descriptions.txt for descriptions).

Supplement File 10: Sequence alignments of coding sequences of individual loci in FASTA format (gene_alignments.zip) (see gene-alignments-descriptions.txt for descriptions).

Supplement File 11: Maximum-likelihood phylogenetic trees created with IQ-TREE for individual loci (gene_trees.zip) (see gene-trees-descriptions.txt for descriptions).

Supplement File 12: Partition files for concatenated gene alignments in NEXUS format (partition_files.zip) (see partition-files-descriptions.txt for descriptions).

Acknowledgements

The study was funded by the German Research Foundation grants AH 175/6-1, AH175/6-2 (D.A.), MI 649/18-1 (B.M.), NI 1387/6-1 and 6-2 as well as by institutional funding of the Museum Koenig Bonn and the A. Koenig Stiftung Bonn.

References

Ahrens, D. (2006) The phylogeny of Sericini and their position within the Scarabaeidae based on morphological characters (Coleoptera: Scarabaeidae). *Systematic Entomology*, 31, 113–144.

Ahrens, D. & Vogler, A.P. (2008) Towards the phylogeny of chafers (Sericini): analysis of alignment-variable sequences and the evolution of segment numbers in the antennal club. *Molecular Phylogenetics and Evolution*, 47, 783-798.

Ahrens, D., Schwarzer, J. & Vogler, A.P. (2014) The evolution of scarab beetles tracks the sequential rise of angiosperms and mammals. *Proceedings of the Royal Society B*, 281, 20141470.

Ahrens, D., Scott, M., Vogler, A.P. (2011) A phylogeny of monkey beetles based on mitochondrial and ribosomal DNA (Coleoptera: Scarabaeidae: Hopliini). *Molecular Phylogenetics and Evolution*, 60, 408–415.

Ayivi, S.P.G., Tong, Y., Storey, K.B., Yu, D.N., & Zhang, J.Y. (2021) The mitochondrial genomes of 18 new Pleurosticti (Coleoptera: Scarabaeidae) exhibit a novel trnQ-NCR-trnI-trnM gene rearrangement and clarify phylogenetic relationships of subfamilies within Scarabaeidae. *Insects* 14;12(11), 1025.

Balthasar, V. (1963) Monographie der Scarabaeidae und Aphodiidae der palaearktischen und orientalischen Region (Coleoptera: Lamellicornia). Tschechoslowak. Akademie der Wissenschaften, Prague, 1–3, 287 pp.

Berendse, F. & Scheffer, M. (2009) The angiosperm radiation revisited, an ecological explanation for Darwin's 'abominable mystery'. *Ecology Letters*, 12, 865–872.

Beza-Beza, C. F., Jiménez-Ferbans, L., McKenna, D.D. (2020) Phylogeny and systematics of wood-degrading Neotropical bess beetles (Coleoptera: Passalidae: Passalinae). *Arthropod Systematics and Phylogeny*, 78 (2), 287–308.

Bocak, L., Barton, C., Crampton-Platt, A., Chesters, D., Ahrens, D., Vogler, A.P. (2014) Building the Coleoptera tree-of-life for >8000 species: composition of public DNA data and fit with Linnaean classification. *Systematic Entomology*, 39, 97–110

Bouchard, P., Bousquet, Y., Davies, A., Alonso-Zarazaga, M., Lawrence, J., Lyal, C., Newton, A., Reid, C., Schmitt, M., Slipinski, A., Smith, A. (2011) Family-group names in Coleoptera (Insecta). ZooKeys, 88, 1–972.

Browne, D.J. & Scholtz, C.H. (1995) Phylogeny of the families of Scarabaeoidea (Coleoptera) based on the characters of the hindwing articulation, hindwing base and wing venation. *Systematic Entomology*, 20, 145–173.

Browne, J. & Scholtz, C.H. (1998) Evolution of the scarab hindwing articulation and wing base: a contribution toward the phylogeny of the Scarabaeidae (Scarabaeoidea: Coleoptera). *Systematic Entomology*, 23: 307–326.

Browne, J. & Scholtz, C.H. (1999) A phylogeny of the families of Scarabaeoidea (Coleoptera). *Systematic Entomology*, 24: 51–84.

Cai, C., Tihelka, E., Giacomelli, M., Lawrence, J.F., Ślipiński, A., Kundrata, R., Yamamoto, S., Thayer, M.K., Newton, A.F., Leschen, R.A.B., Gimmel, M.L., Lü, L., Engel, M.S., Bouchard, P., Huang, D., Pisani, D., Donoghue, P.C.J. (2022) Integrated phylogenomics and fossil data illuminate the evolution of beetles. *Royal Society Open Science*, 9, 211771.

Caterino, M.S., Hunt, T., & Vogler, A.P. (2005) On the constitution and phylogeny of Staphyliniformia (Insecta: Coleoptera). *Molecular Phylogenetics and Evolution*, 34(3), 655–72.

Cherman, M.A. & Morón, M.A. (2014) Validación de la familia Melolonthidae Leach, 1819 (Coleoptera: Scarabaeoidea). *Acta Zoológica Mexicana (n.s.)*, 30(1), 201–220.

Crowson, R.A. (1955) The natural classification of the families of Coleoptera. Nathaniel Lloyd & Co., Ltd., London, 187 p.

de Boer, H.J., Eppinga, M.B., Wassen, M.J., Dekker, S.C. (2012) A critical transition in leaf evolution facilitated the Cretaceous angiosperm revolution. *Nature Communications*, 3, 1221.

Eberle, J., Sabatinelli, G., Cillo, D., Bazzato, E., Šípek, P., Sehnal, R., Bezděk, A., Král, D., Ahrens, D. (2019) A molecular phylogeny of Melolonthinae chafer

beetles revisits the polyphyly of Tanyproctini (Coleoptera, Scarabaeidae). *Zoologica Scripta*, 48, 349–358.

Eddy, S.R. (2011) Accelerated profile HMM searches. *PLOS Computational Biology*, 7, e1002195.

Endrödi, S. (1966) Monographie der Dynastinae (Coleoptera, Lamellicornia). Teil 1. *Abhandlungen Staatliches Museum für Tierkunde*, 33, 1–460.

Erichson, W.F. (1848) Naturgeschichte der Insecten Deutschlands. Abt. I, Coleoptera 3: 801–968.

Estes, A.M., Hearn, D.J., Snell-Rood, E.C., Feindler, M., Feeser, K., Abebe, T., Duning Hotopp, J.C., Moczek, A.P. (2013) Brood ball-mediated transmission of microbiome members in the dung beetle, *Onthophagus taurus* (Coleoptera: Scarabaeidae). *PLoS ONE*, 8(11), e79061.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 15;29(7), 644–52.

Grebennikov, V.V. & Scholtz, C.H. (2004) The basal phylogeny of Scarabaeoidea (Insecta: Coleoptera) inferred from larval morphology. *Invertebrate Systematics*, 18, 321–348.

Gunter, N.L., Weir, T.A., Slipinksi, A., Bocak, L., Cameron, S.L. (2016) If dung beetles (Scarabaeidae: Scarabaeinae) arose in association with Dinosaurs, did they also suffer a mass co-extinction at the K-Pg boundary? *PLoS ONE*, 11(5), e0153570.

Guo, S., Lin, X., Song, N. (2022) Mitochondrial phylogenomics reveals deep relationships of scarab beetles (Coleoptera, Scarabaeidae). *PLoS ONE*, 17(12): e0278820.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S. (2018) UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522.

Hoyt, C.P., Osborne, G.O., Mulcock, A.P. (1971) Production of an insect sex attractant by symbiotic bacteria. *Nature*, 230, 472–473.

Howden, H. F. (1982) Larval and adult characters of *Frickius* Germain, its relationships to the Geotrupini, and a phylogeny of some major taxa in the Scarabaeoidea (Insecta: Coleoptera). *Canadian Journal of Zoology*, 60, 2713–2724.

Hunt, T., Bergsten, J., Levkanicova, Z., Papadopoulou, A., St. John, O., Wild, R., Hammond, P.M., Ahrens, D., Balke, M., Caterino, M.S., Gómez-Zurita, J., Ribera, I., Barraclough, T.G., Bocakova, M., Bocak, l., Vogler, A.P. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, 318, 1913–1916

Iablokoff-Khnzorian, S. M. (1977) Über die Phylogenie der Lamellicornia. *Entomologische Abhandlungen des Staatlichen Museum für Tierkunde*, 41, 135–199.

Janssens, A. (1949) Table synoptique et essai de classification pratique des Coléoptères Scarabaeidae. *Bulletin de l'Institut royal des sciences naturelles de Belgique*, 25, 1–30.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S. (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14, 587–589.

Katoh, K., Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780.

Kishino, H., Hasegawa, M. (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *Journal of Molecular Evolution*, 29, 170–179.

Kishino, H., Miyata, T., Hasegawa, M. (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *Journal of Molecular Evolution*, 31, 151–160.

Kohlmann, B. & Morón, M. A. (2003) Análisis histórico de la clasificación de los Coleoptera Scarabaeoidea o Lamellicornia. *Acta Zoológica Mexicana (n.s.)*, 90, 175–280.

Krell, F.-T. (2007) Catalogue of fossil Scarabaeoidea (Coleoptera: Polyphaga) of the Mesozoic and Tertiary – Version 2007. *DMNS Technical* Report 2007-8: 1–79.

Krueger, F., James, F., Ewels, P., Afyounian, E. & Schuster-Boeckler, B. (2021) FelixKrueger/TrimGalore: v0.6.6 – DOI via Zenodo (TrimGalore v.0.6.7). Zenodo. https://doi.org/10.5281/zenodo.5127899

Kück, P., Meusemann, K., Dambach, J., Thormann, B., von Reumont, B.M., Wägele, J.W., Misof, B. (2010) Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Frontiers in Zoology*, 7, 10.

Lacroix, M. (2007) Pachydeminae du Monde (Scarabaeoidea, Melolonthidae). Genera et catalogue commenté. M. Lacroix/collection Hannetons, 2007, 450pp.

Lacroix, M. (2014) Melolonthidae Malgaches (Coleoptera, Scarabaeoidea). Catalogue commenté. M. Lacroix/collection Hannetons, 2014, 260pp.

Lawrence, J.F. & Newton, A.F. Jr. (1982) Evolution and classification of beetles. *Annual Review of Ecology, Evolution, and Systematics*, 13, 261–290.

Lawrence, J.F. & Newton, A.F. (1995) Families and subfamilies of Coleoptera (with selected genera, notes and references, and data on family-group names), pp. 779-1006. In: Pakaluk, J. & Slipinski, S. A. (Eds*.) Biology, Phylogeny, and Classification of Coleoptera: Papers Celebrating the 80th Birthday of Roy A. Crowson*. Warsaw: Muzeum I Instytut Zoologii PAN. 277 p.

Lawrence, J.F. Ślipiński, A., Seago, A.E., Thayer, M.K., Newton, A.F., Marvaldi, A.E. (2011) Phylogeny of the Coleoptera based on morphological characters of adults and larvae. *Annales Zoologici*, 61, 1–217.

Li, B., Lopes, J.S., Foster, P.G., Embley, T.M., and Cox, C.J. (2014) Compositional biases among synonymous substitutions cause conflict between gene and protein trees for plastid origins. *Mol. Biol. Evol.* 31, 1697–1709.

Lu, Y., Ballerio, A., Wang, S., Zou, Z., Gorb, S.N., Wang, T., Li, L.,Ji, S., Zhao, Z., Li, S., Tong, Y., Chen, Y., Zhuo, D., Luo, C., Zhang, W., Liu, N., Gu, Q., Bai, M. (2022) The evolution of conglobation in Ceratocanthinae. *Communications Biology*, 5, 777

Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., Zdobnov, E.M. (2021) BUSCO Update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, 38(10), 4647–4654.

Martin, M.M. (1983) Cellulose digestion in insects. *Comp. Biochem. Physiol.* 75A(3), 313–324.

Martin, M.M., Jones, C.G., Bernays, E.A. (1991) The evolution of cellulose digestion in insects [and Discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 333(1267), 281–288.

McKenna, D.D., Farrell, B.D., Caterino, M.S., Farnum, C.W., Hawks, D.C., Maddison, D.R., Seago, A.E., Short, A.E.Z., Newton, A.F., Thayer, M.K. (2014) Phylogeny and evolution of Staphyliniformia and Scarabaeiformia: Forest litter as a stepping stone for diversification of nonphytophagous beetles. *Systematic Entomology*, 40, 35–60

McKenna, D.D., Wild, A.L., Kanda, K., Bellamy, C.L., Beutel, R.G., Caterino, M.S., Farnum, C.W., Hawks, D.C., Ivie, M.A., Jameson, M.L., Leschen, R.A.B., Marvaldi, A.E., McHugh J.V., Newton, A.F., Robertson, J.A., Thayer, M.K., Whiting, M.F., Lawrence J.F., Ślipiński, A., Maddison, D.R., Farrell, B.D. (2015) The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution. *Systematic Entomology*, 40, 835–880.

Medvedev, S.I. (1976) Sistematike i filogenii plastinchatouyx zhukov Palearktiki. *Entomologischeskoe Obozrenje*, 55, 400–409.

Misof, B. & Misof, K. (2009) A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. *Systematic Biology*, 58, 21–34.

Misof, B., et al. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346(6210), 763–767.

Morón, M.A. (1984) *Escarabajos, 200 millones de años de evolución*. Publicación 14. Instituto de Ecología, México. 131 p.

Morón, M.A. (1997) Antecedentes, pp.1-8. In: Morón, M. A., B. C. Ratcliffe, & C. Deloya (Eds.). Atlas de los escarabajos de México. Coleoptera Lamellicornia. Vol. I. Familia Melolonthidae. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad y Sociedad Mexicana de Entomología, México. 280 p.

Morón, M.A. (2003) *Atlas de los escarabajos de México. Coleoptera: Lamellicornia. Vol. II Familias Scarabaeidae, Trogidae, Passalidae y Lucanidae*. Argania editio. Barcelona. 227 p.

Neita-Moreno, J.C., Agrain, F.A., Eberle, J., Ahrens, D., Pereyra, V. (2019) On the phylogenetic position and systematics of extant and fossil Aclopinae (Coleoptera: Scarabaeidae). *Systematic Entomology*, 44, 709–727.

Pacheco, T.L., Bohacz, C., Ballerio, A., Schoolmeesters, P., Ahrens, D. (2022) Revisiting trends of morphology of antennal sensilla in Scarabaeoid beetles. *Zoomorphology* 141, 315-326.

Paulian, R. & Baraud, J. (1982) *Faune des Coléoptères de France, II. Lucanoidea et Scarabaeoidea.* Éditions Lechevalier, Paris. 477 pp.

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., Klopfstein, 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. *Current Biology*, 27(7), 1013–1018.

Petersen, M., Meusemann, K., Donath, A., Dowling, D., Liu, S., Peters, R.S., Podsiadlowski, L., Vasilikopoulos, A., Zhou, X., Misof, B., Niehuis, O. (2017) Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics*, 18, 111.

Ratcliffe, B.C. & Jameson, M.L. (2004) The revised classification for Scarabaeoidea: What the hell is going on? *Scarabs*, 15, 3–10.

Scholtz, C.H. (1990) Phylogenetic trends in the Scarabaeoidea (Coleoptera). *Journal of Natural History*, 24, 1027–1066.

Scholtz, C.H. & Grebennikov, V.V. (2005) Scarabaeoidea Latreille, 1802, pp. 368–425. In: R.G. Beutel & R.A.B. Leschen (Eds.) *Handbook of Zoology, Coleoptera, Beetles. Volume 1. Morphology and Systematics (Archostemata, Adephaga, Myxophaga, Polyphaga partim*. Berlin, New York: Walter de Gruyter. 567 p.

Sharma, P.P., Kaluziak, S.T., Pérez-Porro, A.R., González, V.L., Hormiga, G., Wheeler, W.C., & Giribet, G. (2014) Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal. *Molecular Biology and Evolution*, 31, 2963–2984.

Shimodaira, H. (2002) An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, 51, 492–508.

Shimodaira, H. & Hasegawa, M (1999) Multiple comparisons of log-likelihoods with applications to phylogene.tic inference. *Molecular Biology and Evolution*, 16, 1114.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31, 3210–2.

Šípek, P., Fabrizi, S., Eberle, J., Ahrens, D. (2016) A molecular phylogeny of rose chafers (Coleoptera: Scarabaeidae: Cetoniinae) reveals a complex and concerted morphological evolution related to their flight mode. *Molecular Phylogenentics and Evolution*, 101, 163–175.

Smith, A.B.T. (2006) A review of the family-group names for the superfamily Scarabaeoidea (Coleoptera) with corrections to nomenclature and a current classification. *Coleopterists Society Monograph*, 5,144–204.

Smith, A.B.T., Hawks, D.C. & Heraty, J.M. (2006) An overview of the classification and evolution of the major scarab beetle clades (Coleoptera: Scarabaeoidea) based on preliminary molecular analyses. *Coleopterists Society Monograph*, 5, 35–46.

Song, N. & Zhang, H. (2018) The mitochondrial genomes of phytophagous scarab beetles and systematic implications. *Journal of Insect Science*, 18, 11.

Strimmer, K. & Rambaut, A. (2002) Inferring confidence sets of possibly misspecified gene trees. *Proceedings of the Royal Society B*, 269, 137–142.

Strimmer, K. & von Haeseler, A. (1997) Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences*, 94, 6815–6819.

Suyama, M., Torrents, D. & Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, 34, W609–W612.

Taylor, L., Banwart, S., Leake, J., Beerling, D.J. (2011) Modeling the evolutionary rise of ectomycorrhiza on sub-surface weathering environments and the geochemical carbon cycle. *American Journal of Science*, 311, 369–403.

Taylor, L.L., Banwart, S.A., Valdes, P.J., Leake, J.R., Beerling, D.J. (2012) Evaluating the effects of terrestrial ecosystems, climate and carbon dioxide on weathering over geological time: a global-scale process-based approach. *Philosophical Transactions of the Royal Society B*, 367, 565–582.

Timmermans, M.J.T.N., Barton, C., Haran, J., Ahrens, D., Culverwell, C.L., Ollikainen, A., Dodsworth, S., Foster, P.G., Bocak, L., Vogler, A.P. (2016) Family-level sampling of mitochondrial genomes in Coleoptera: compositional heterogeneity and phylogenetics. *Genome Biology and Evolution*, 8, 161–175.

Toussaint, E.F.A., Seidel, M., Arriaga-Varela, E., Hájek, J., Král, D., Sekerka, L., Short, A.E.Z., Fikáček, M. (2017) The peril of dating beetles. *Systematic Entomology*, 42, 1–10.

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S. (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19, 153.

Table 1. Overview of results on monophyly of Scarabaeidae retrieved in different previous phylogenetic analyses.

Reference	Scarab dung beetles + Pleurosticti	Data/ tree building method
Grebennikov & Scholtz (2004)	no	Larval morphology/ Parsimony
Caterino et al. (2005)	no	18S/ ML, Parsimony
Smith et al. (2006)	no	18S, 28S/ Parsimony
Hunt et al. (2007)	no	28S, 16S, COI/ BI
Lawrence et al. (2011)	no	Morphology/ Parsimony
Ahrens et al. (2014)	no	18S, 28S, 16S, COI/ BI
Bocak et al. (2014)	no	18S, 28S, 16S, COI/ ML
Timmermanns et al. (2015)	yes	mt genomes/ BI
McKenna et al. (2014)	yes	CAD+28S/ BI
McKenna et al. (2015)	no	8 nuclear genes/ BI
Gunter et al. (2016)	yes	28S, 16S, COI/ BI
Toussaint et al. (2017)	no	8 nuclear genes/ BI
Song & Zhang (2018)	Yes	mt genomes/ ML
Zhang et al. (2018)	yes	95 PCG; AA/ ML (RaxML)
	yes	95 PCG; Nucleotides/ ML (RaxML)
	yes	AA/ ML (IQ-TREE)
	no	95 PCG; Nucleotides/ ML (IQ-TREE)
	no	95 PCG; AA/ BI
	yes	Nucleotides/ BI
McKenna et al. (2019)	yes	AA of Transcriptomes (4,818 nuclear genes)/ ML
Ayivi et al. (2021)	yes	mt genomes/ BI, ML
Cai et al. (2022)	no	68 single-copy nuclear protein-coding (NPC) genes/ BI (Phylobayes), ML (IQ-TREE)
Guo et al. (2022)	yes	mt genomes/ BI, ML

Table 2. Likelihood tests regarding the monophyly of Scarabaeidae using the coleopteran-specific dataset, with constraint trees using a variety of resampling tests in IQ-TREE using the RELL approximation including bootstrap proportion (BP), Kishino-Hasegawa test (KH), Shimodaira-Hasegawa test (SH), expected likelihood weights (ELW), and the approximately unbiased (AU) test. The confirmed most likely topology is highlighted in bold.

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	c-ELW	p-AU
hmmalign NT12							
1: Dung scarabs + Pleurosticti	-48463502	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-48465916	2414	0	0	0	0	2.45E-05
3: Hybosoridae + Dung scarabs	-48465122	1620.3	0	0	0	0	1.68E-59
hmmalign NT123							
1: Dung scarabs + Pleurosticti	-124696860	1522.3	0	0	0	0	1.08E-05
2: Hybosoridae + Pleurosticti	-124695338	0	1	1	1	1	1
3: Hybosoridae + Dung scarabs	-124697335	1997.4	0	0	0	0	1E-07
hmmalign AA							
1: Dung scarabs + Pleurosticti	-44349511	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-44353242	3730.9	0	0	0	0	1.48E-08
3: Hybosoridae + Dung scarabs	-44351488	1976.8	0	0	0	0	7.45E-09
MAFFT NT12							
1: Dung scarabs + Pleurosticti	-31506595	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-31508450	1855	0	0	0	0	0.000172
3: Hybosoridae + Dung scarabs	-31507967	1372	0	0	0	0	3.29E-69
MAFFT NT123							
1: Dung scarabs + Pleurosticti	-87783308	2052.7	0	0	0	0	4.4E-82
2: Hybosoridae + Pleurosticti	-87781255	0	1	1	1	1	1
3: Hybosoridae + Dung scarabs	-87783596	2341.4	0	0	0	0	1.48E-48
MAFFT AA							
1: Dung scarabs + Pleurosticti	-26130488	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-26133387	2898.7	0	0	0	0	6.66E-06
3: Hybosoridae + Dung scarabs	-26132268	1780	0	0	0	0	4.07E-06

Table 3. Likelihood tests regarding the monophyly of Rutelinae using the coleopteran-specific dataset, with constraint trees using a variety of resampling tests in IQ-TREE using the RELL approximation including bootstrap proportion (BP), Kishino-Hasegawa test (KH), Shimodaira-Hasegawa test (SH), expected likelihood weights (ELW), and the approximately unbiased (AU) test. The confirmed most likely topology is highlighted in bold.

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	c-ELW	p-AU
hmmalign NT12							
1: Adoretus + Anomala	-48463550	0	0.986	0.983	1	0.986	0.991
2: Adoretus + Dynastinae	-48463962	412.1	0.0006	0.0012	0.0012	0.000607	0.0013
3: Anomala + Dynastinae	-48463842	292.12	0.0134	0.0168	0.0277	0.0135	0.0129
hmmalign NT123							
1: Adoretus + Anomala	-124695262	0	0.97	0.976	1	0.97	0.991
2: Adoretus + Dynastinae	-124695650	387.47	0.0053	0.0063	0.0129	0.00534	0.00767
3: Anomala + Dynastinae	-124695590	327.26	0.0244	0.0236	0.0407	0.0245	0.0198
hmmalign AA							
1: Adoretus + Anomala	-44349512	0	1	1	1	1	1
2: Adoretus + Dynastinae	-44349994	482.06	0.0002	0.0001	0.0002	0.000195	0.000127
3: Anomala + Dynastinae	-44350021	508.52	0.0002	0	0.0002	0.00019	0.000164
MAFFT NT12							
1: Adoretus + Anomala	-31506594	0	1	1	1	1	1
2: Adoretus + Dynastinae	-31507065	470.61	0.0001	0.0002	0.0003	0.0001	7.03E-05
3: Anomala + Dynastinae	-31507098	503.3	0	0	0	5.19E-05	0.000247
MAFFT NT123							
1: Adoretus + Anomala	-87781254	0	0.998	0.997	1	0.998	0.998
2: Adoretus + Dynastinae	-87781673	419.51	0.0021	0.0031	0.0048	0.00214	0.0029
3: Anomala + Dynastinae	-87781779	524.79	0.0001	0.0002	0.0005	0.00013	0.000376
MAFFT AA							
1: Adoretus + Anomala	-26130489	0	1	1	1	1	1
2: Adoretus + Dynastinae	-26130951	462.27	0.0003	0.0001	0.0005	0.000317	0.000697
3: Anomala + Dynastinae	-26130995	506.31	0.0002	0.0001	0.0002	0.000164	0.000818

Table 4. Likelihood tests regarding the monophyly sister group relationship of Hopliini and Melolonthinae s. str. (including Melolonthini and *Sparrmannia*, excluding *Pachypus* and Sericini) using the coleopteran-specific dataset, with constraint trees using a variety of resampling tests in IQ-TREE using the RELL approximation including bootstrap proportion (BP), Kishino-Hasegawa test (KH), Shimodaira-Hasegawa test (SH), expected likelihood weights (ELW), and the approximately unbiased (AU) test. The confirmed most likely topology is highlighted in bold.

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	c-ELW	p-AU
hmmalign NT12							
1: **Hopliini + Dynastinae etc.**	-48463522	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-48466907	3384.2	0	0	0	0	9.98E-06
3: Melolonthinae + Hopliini	-48466288	2765.9	0	0	0	0	9.98E-06
hmmalign NT123							
1: **Hopliini + Dynastinae etc.**	-124695344	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-124703111	7766.7	0	0	0	0	1.01E-08
3: Melolonthinae + Hopliini	-124701830	6486	0	0	0	0	2.33E-43
hmmalign AA							
1: **Hopliini + Dynastinae etc.**	-44349512	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-44352720	3208.2	0	0	0	0	1.65E-52
3: Melolonthinae + Hopliini	-44352216	2704.6	0	0	0	0	1.22E-57
MAFFT NT12							
1: **Hopliini + Dynastinae etc**.	-31506595	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-31509134	2539.6	0	0	0	0	2.63E-44
3: Melolonthinae + Hopliini	-31508736	2141.3	0	0	0	0	1.64E-67
MAFFT NT123							
1: **Hopliini + Dynastinae etc**.	-87781254	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-87787291	6037.7	0	0	0	0	4.16E-50
3: Melolonthinae + Hopliini	-87786383	5129.7	0	0	0	0	4.65E-07
MAFFT AA							
1: **Hopliini + Dynastinae etc.**	-26130487	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-26132818	2331	0	0	0	0	1.09E-45
3: Melolonthinae + Hopliini	-26132612	2124.8	0	0	0	0	6.7E-09

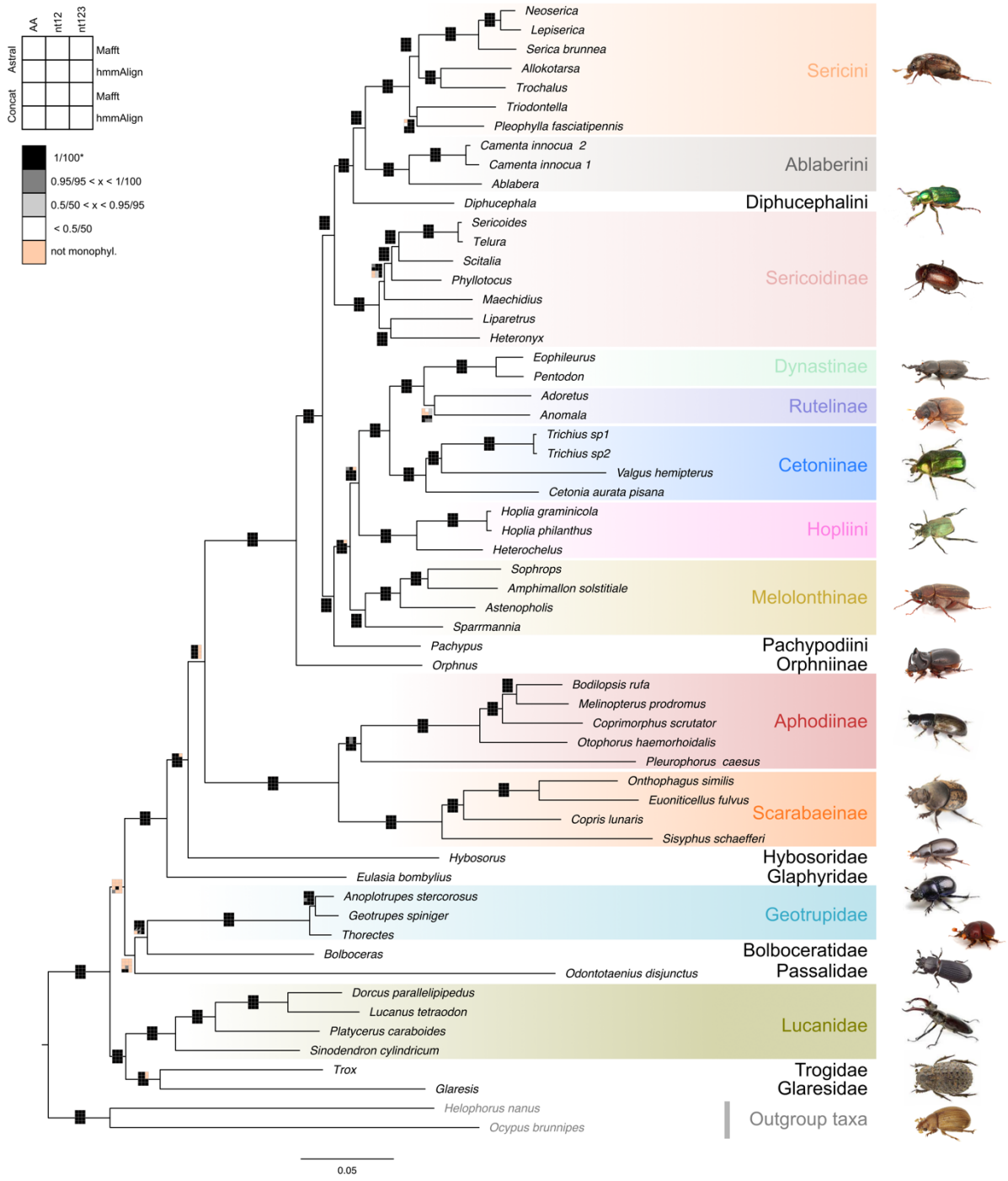


Fig. 1. Phylogenetic tree from the concatenated data of the Mafft alignment of the nucleotide data containing (Coleoptera single copy orthologs, full) only the first and second base pair (nt12). Branch support (nn/ xx; see legend) and results of retrieval with other alignments and tree reconstruction approaches as well as the current systematic assignment of lineages are mapped onto the branches.

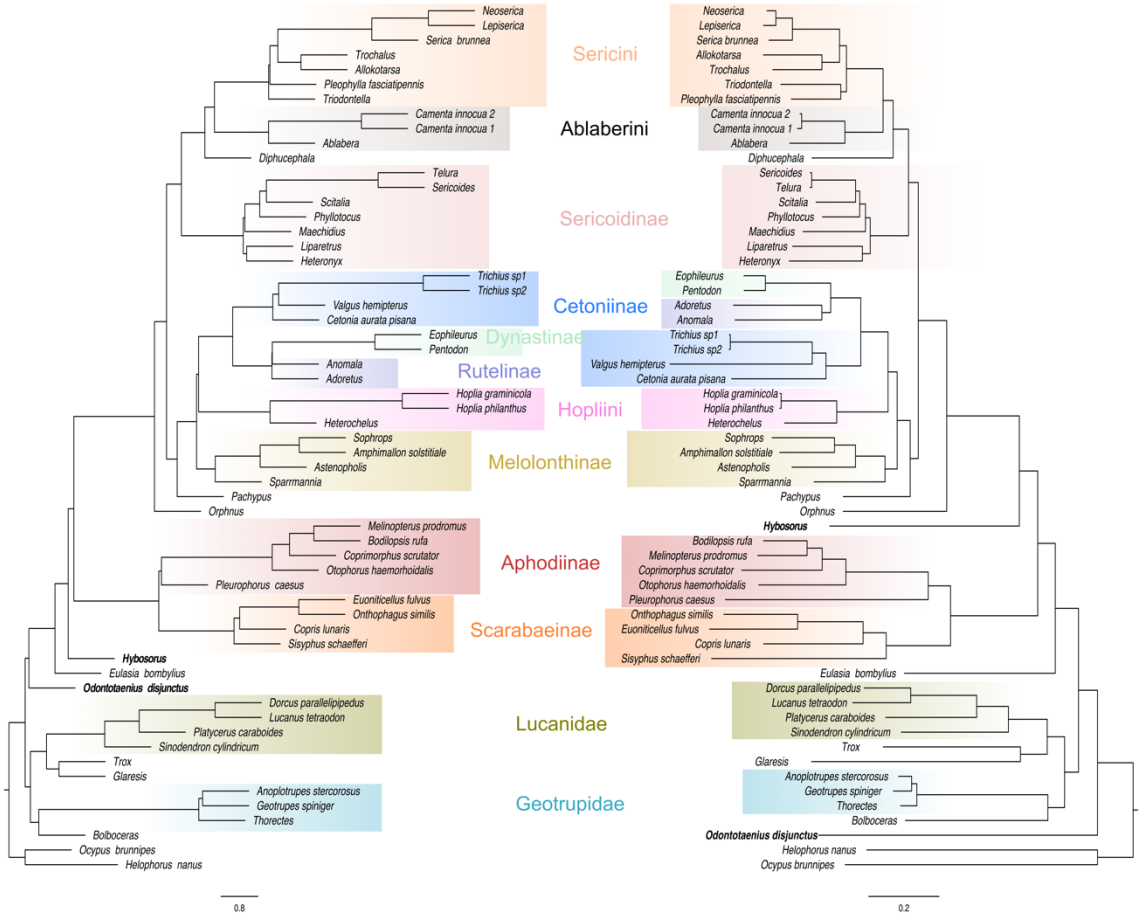


Fig. 2. Contrasting tree topologies based on the Coleoptera single copy orthologs obtained with coalescent tree search with Astral (amino acid sequences; left side) and with concatenated data using all nucleotides (right side); major lineages are highlighted; single taxa changing topology are marked in bold.

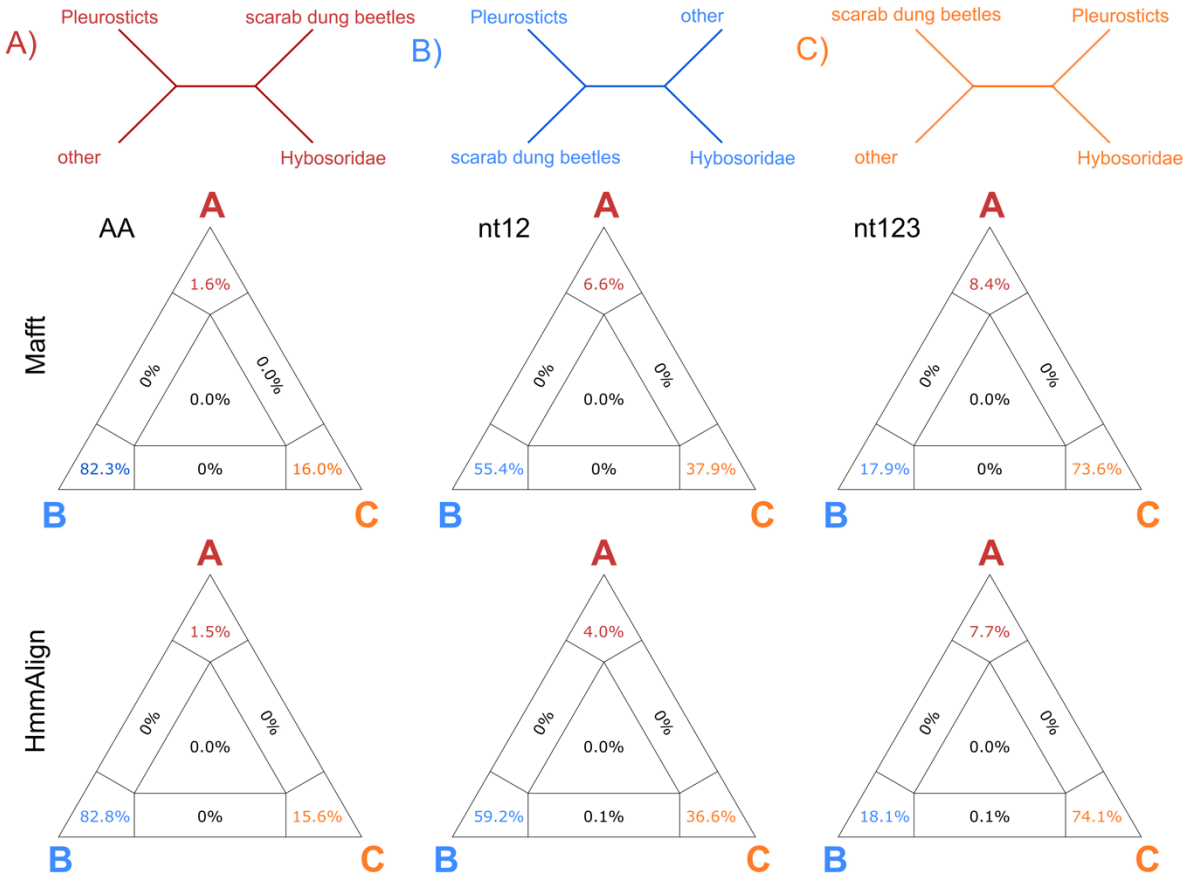


Figure 3. Quartet comparison of Coleoptera - single copy orthologs based on alignments with hmmalign and Mafft, both based on the AA, nt12, and nt123 data set.

Figure S1. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (Mafft alignment; full dataset; Coleoptera orthologs).

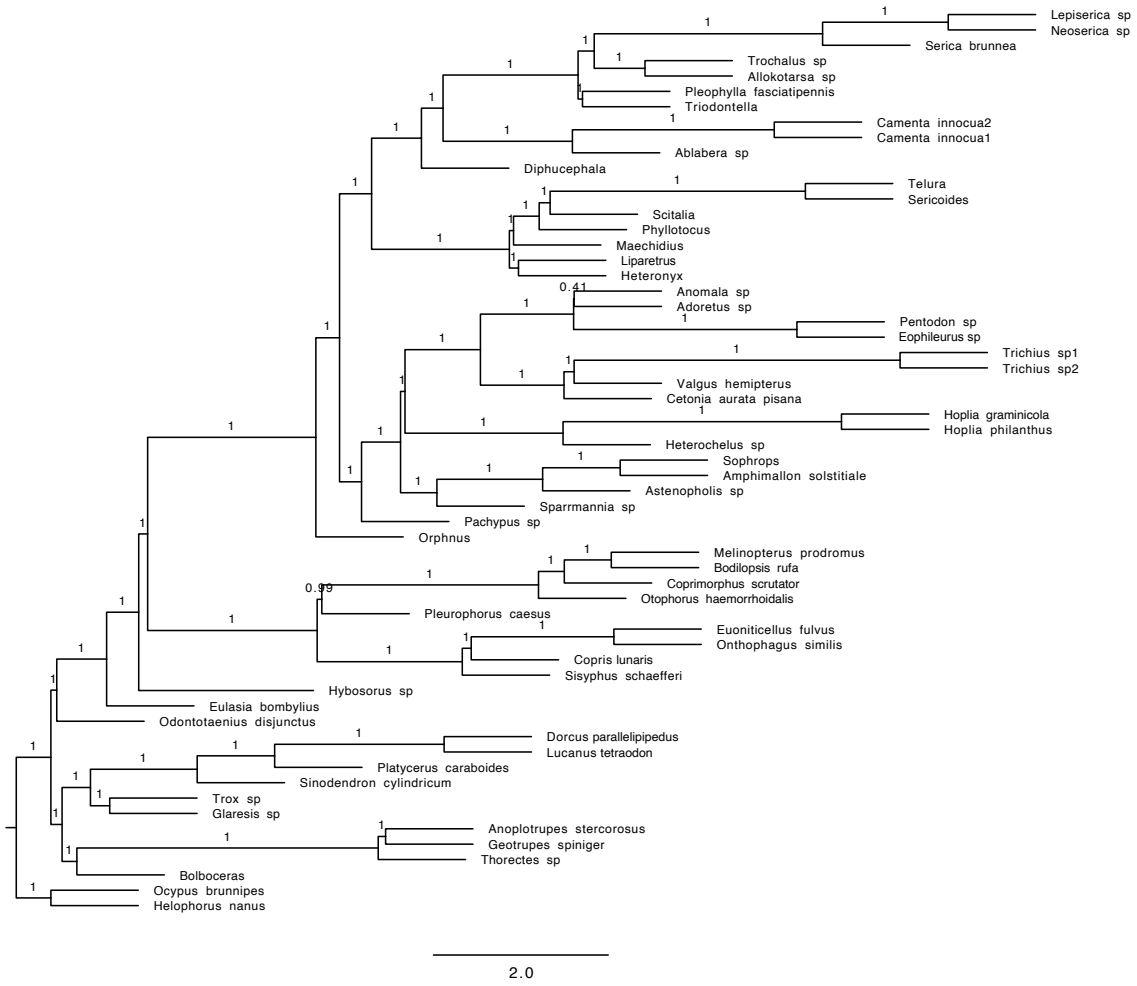


Figure S2. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; full dataset; Coleoptera orthologs).

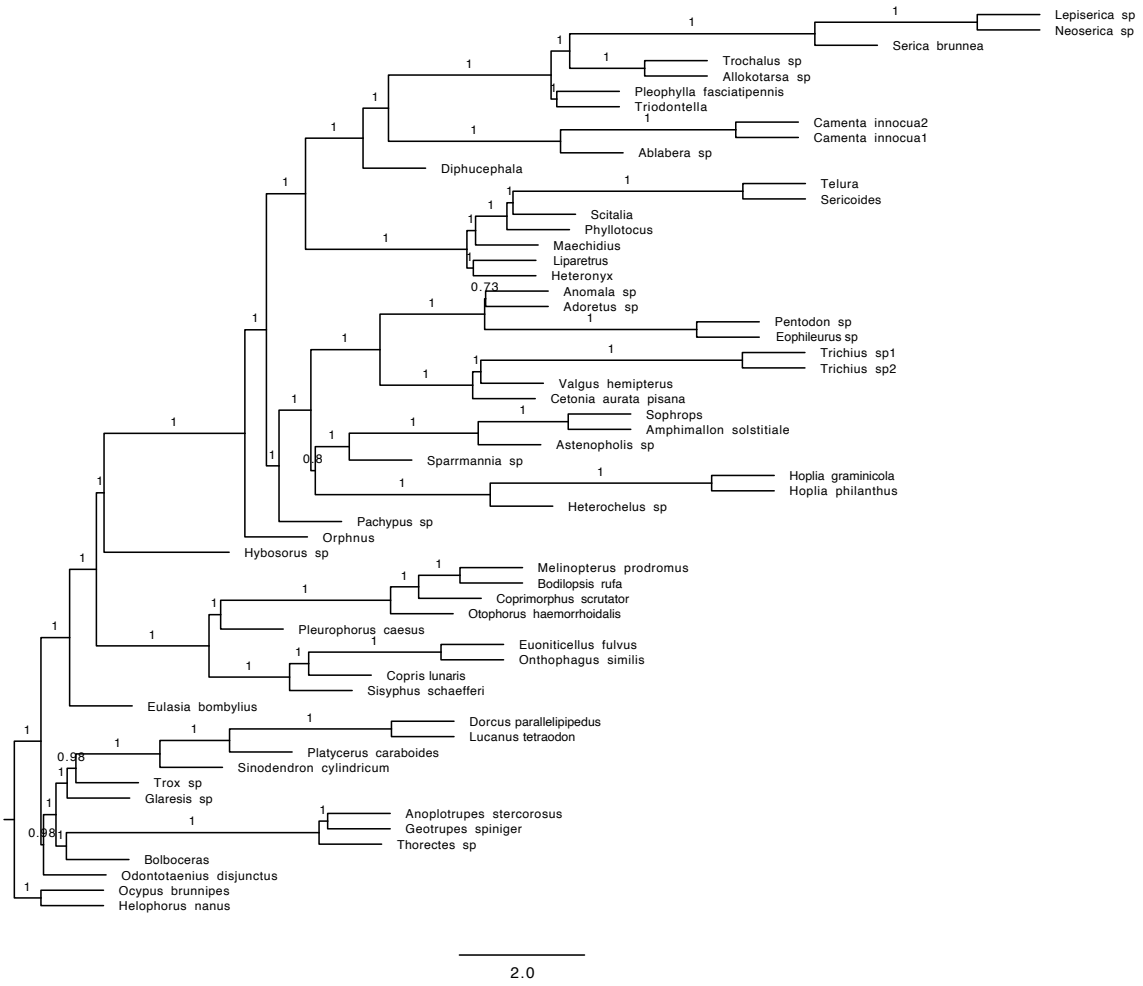


Figure S3. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (Mafft alignment; full dataset; Coleoptera orthologs).

Figure S4. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (HmmAlign alignment; full dataset; Coleoptera orthologs).

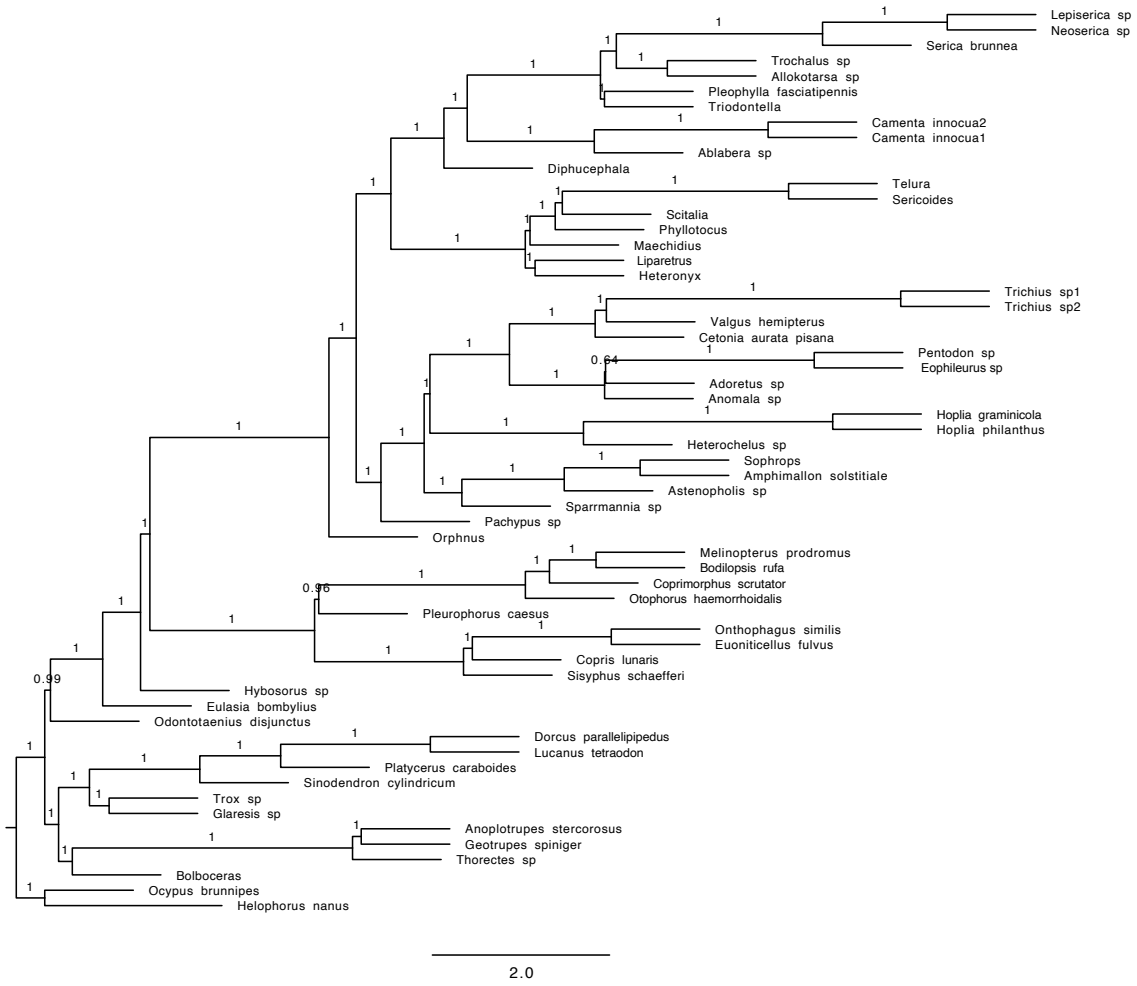


Figure S5. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; full dataset; Coleoptera orthologs).

Figure S6. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; full dataset; Coleoptera orthologs).

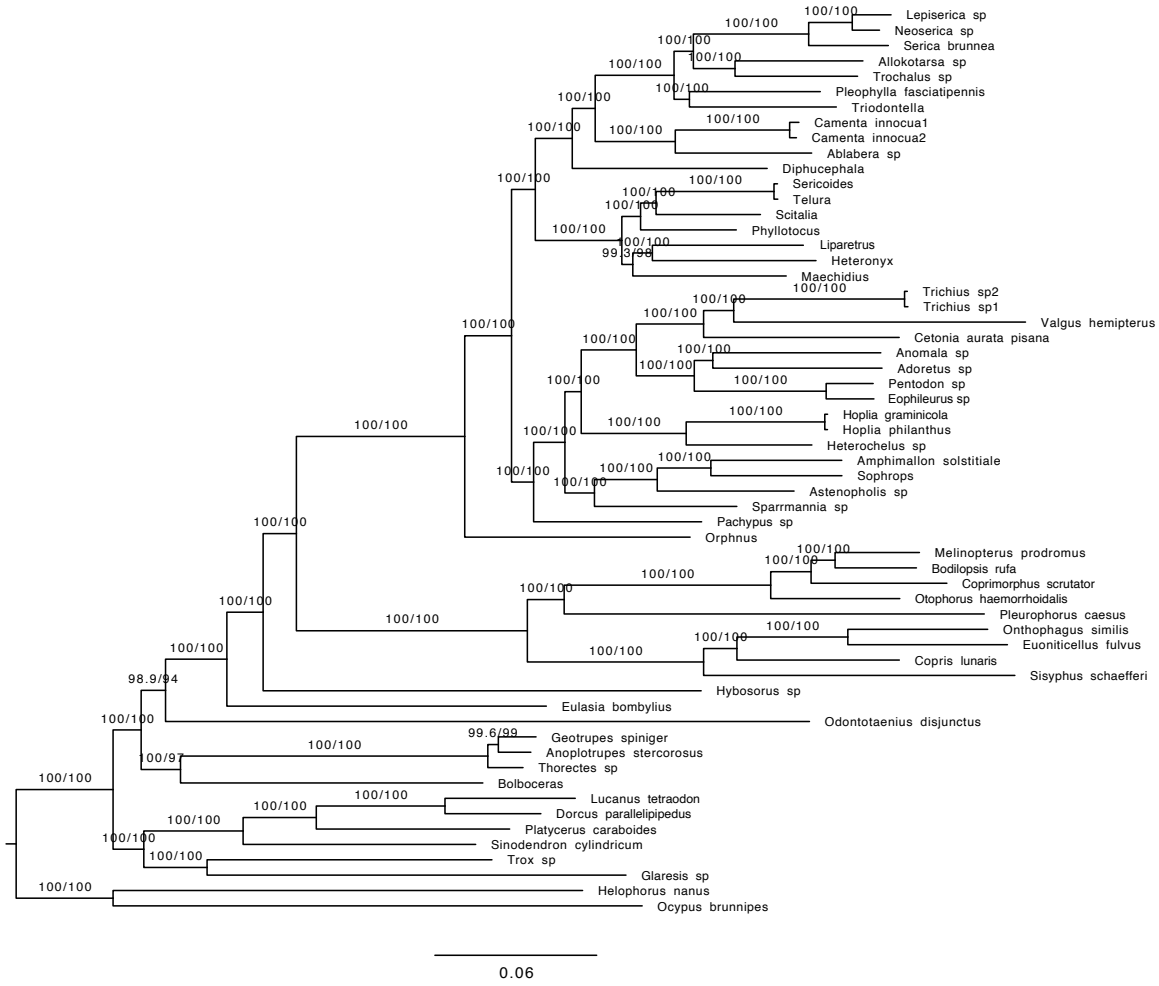


Figure S7. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (Mafft alignment; full dataset; Coleoptera orthologs).

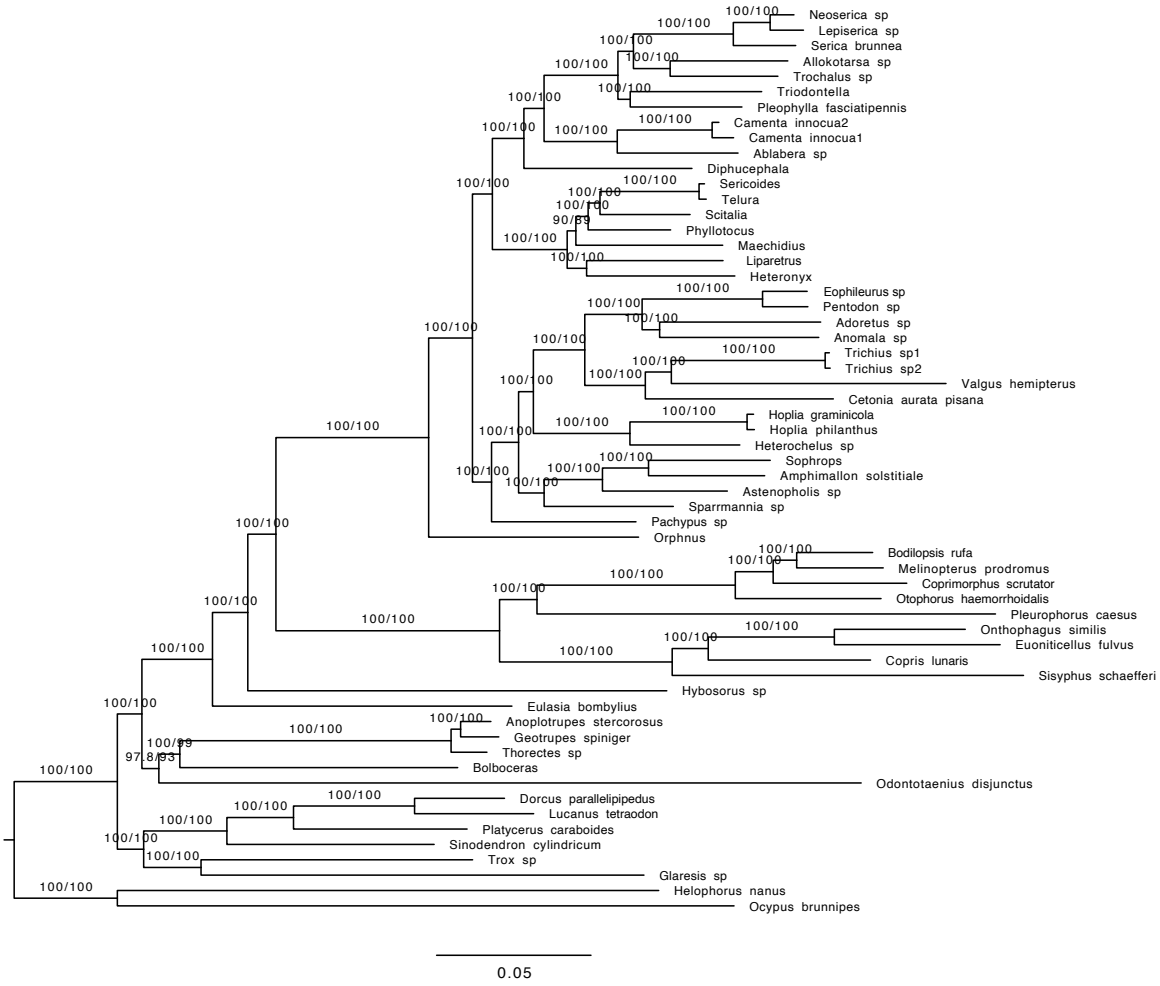


Figure S8. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; full dataset; Coleoptera orthologs).

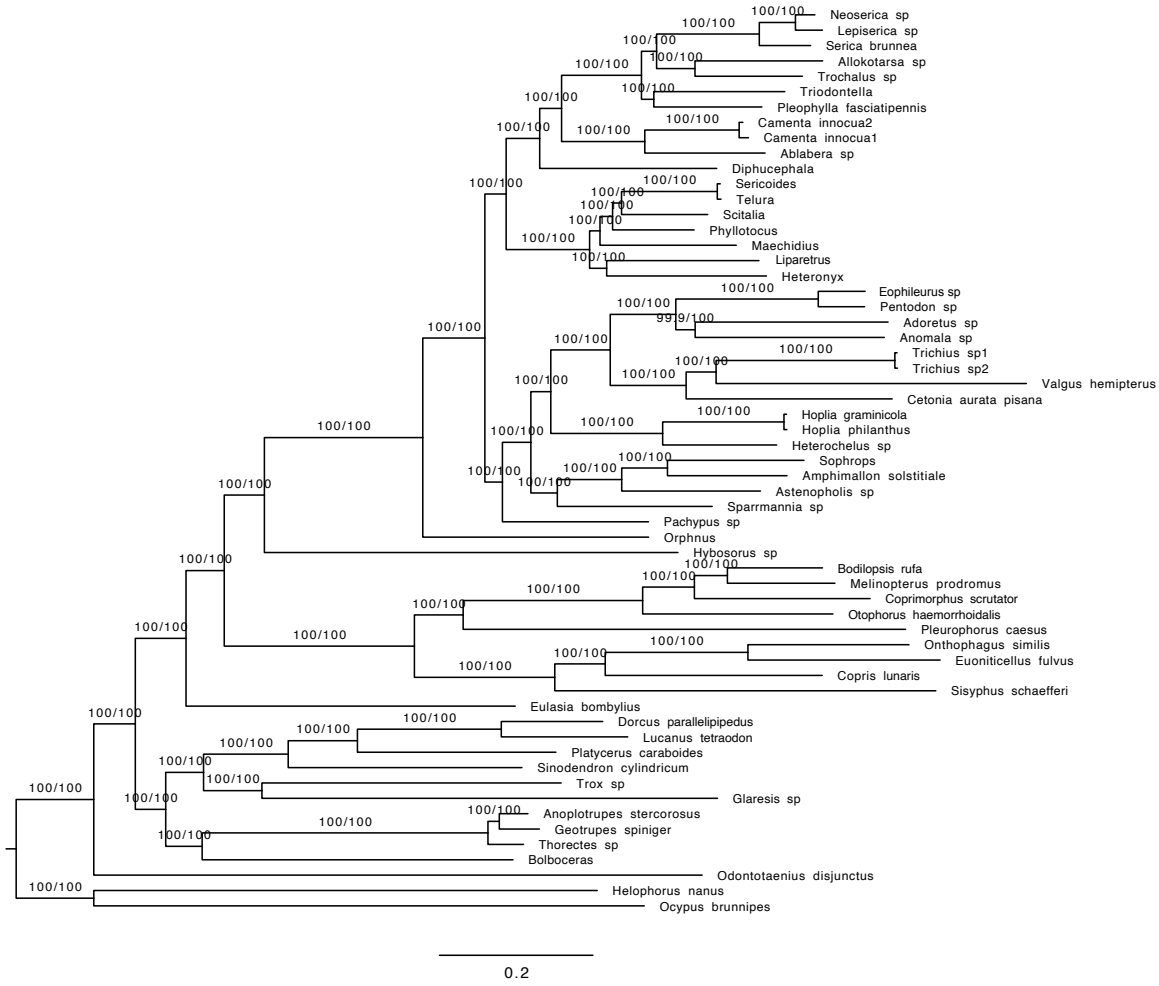


Figure S9. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (Mafft alignment; full dataset; Coleoptera orthologs).

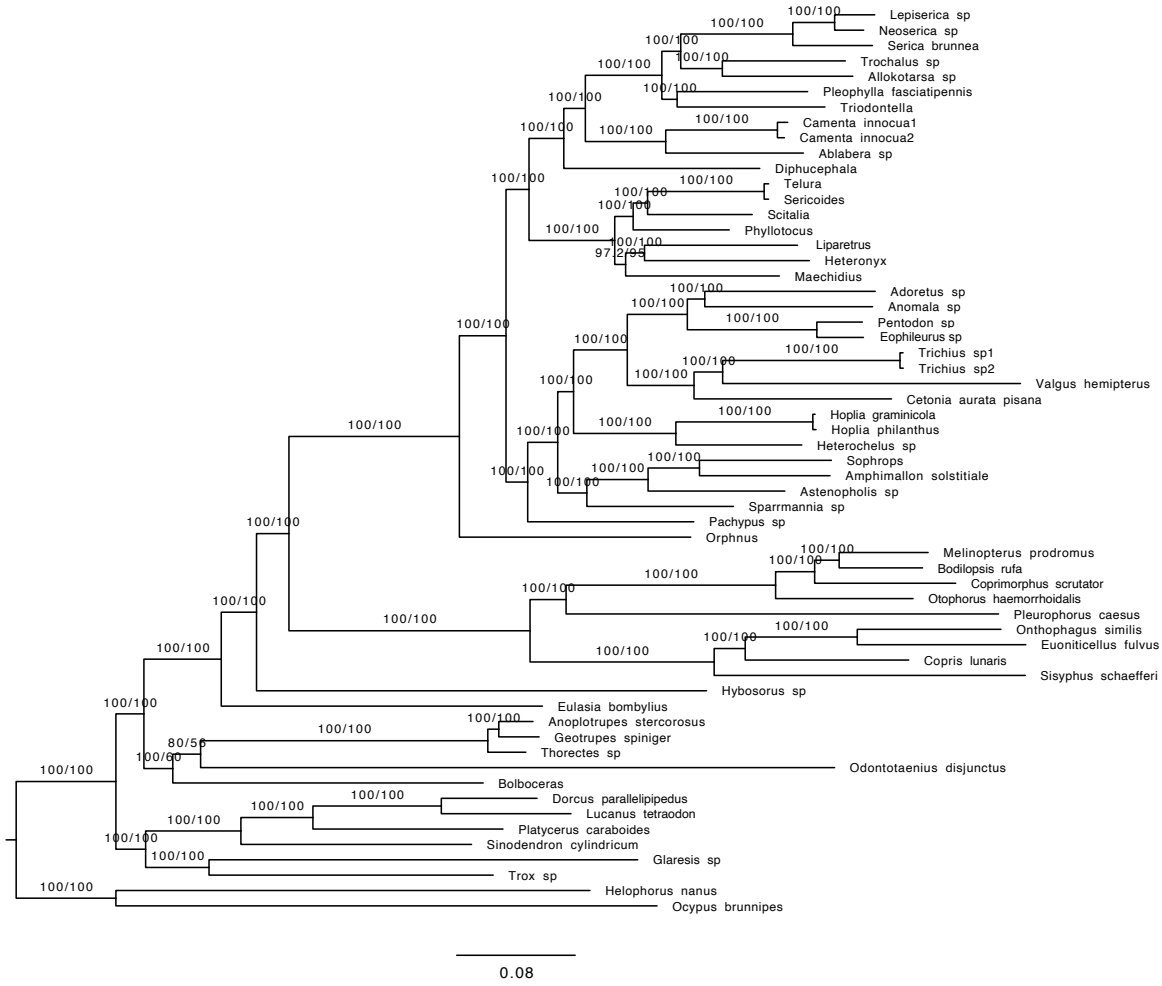


Figure S10. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (HmmAlign alignment; full dataset; Coleoptera orthologs).

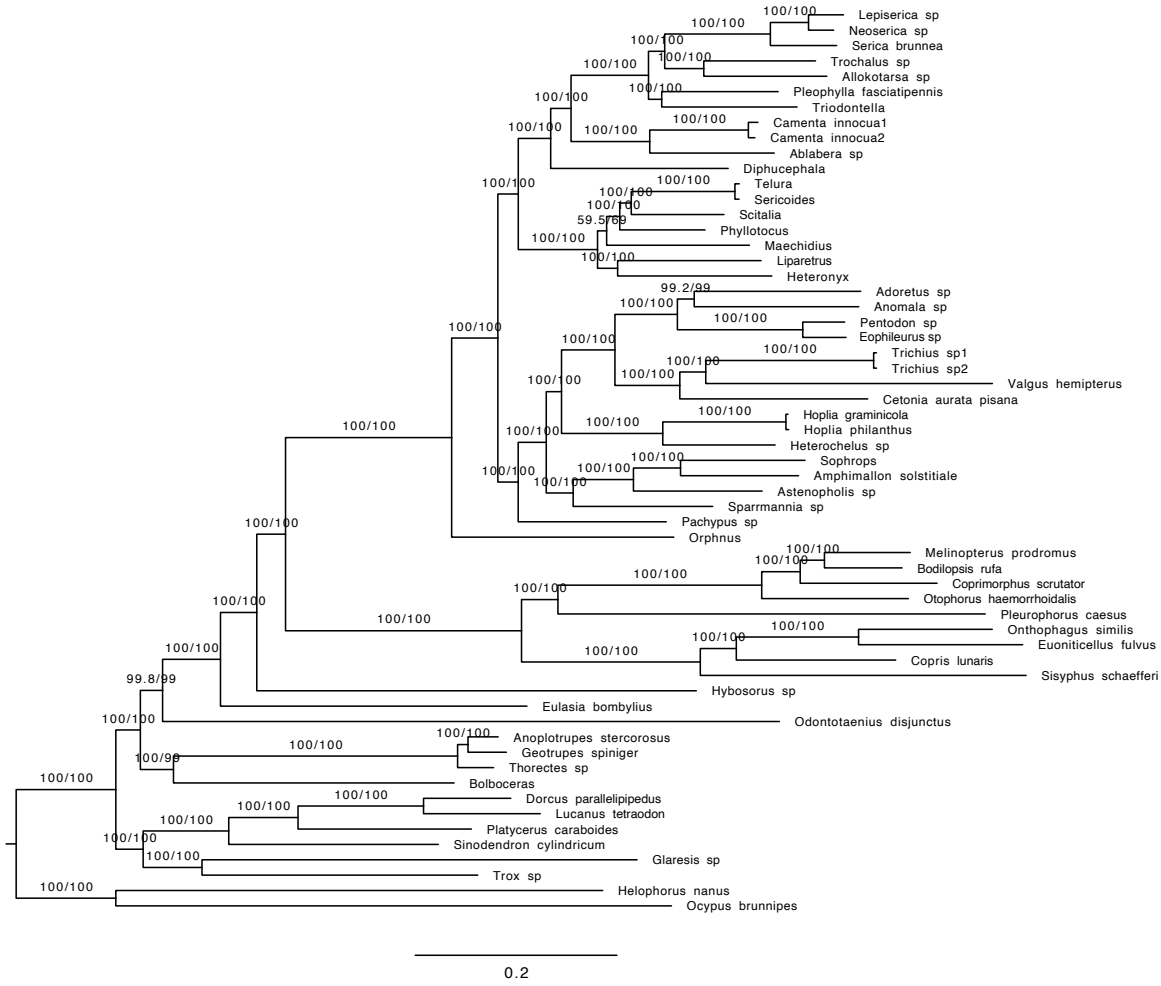


Figure S11. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; full dataset; Coleoptera orthologs).

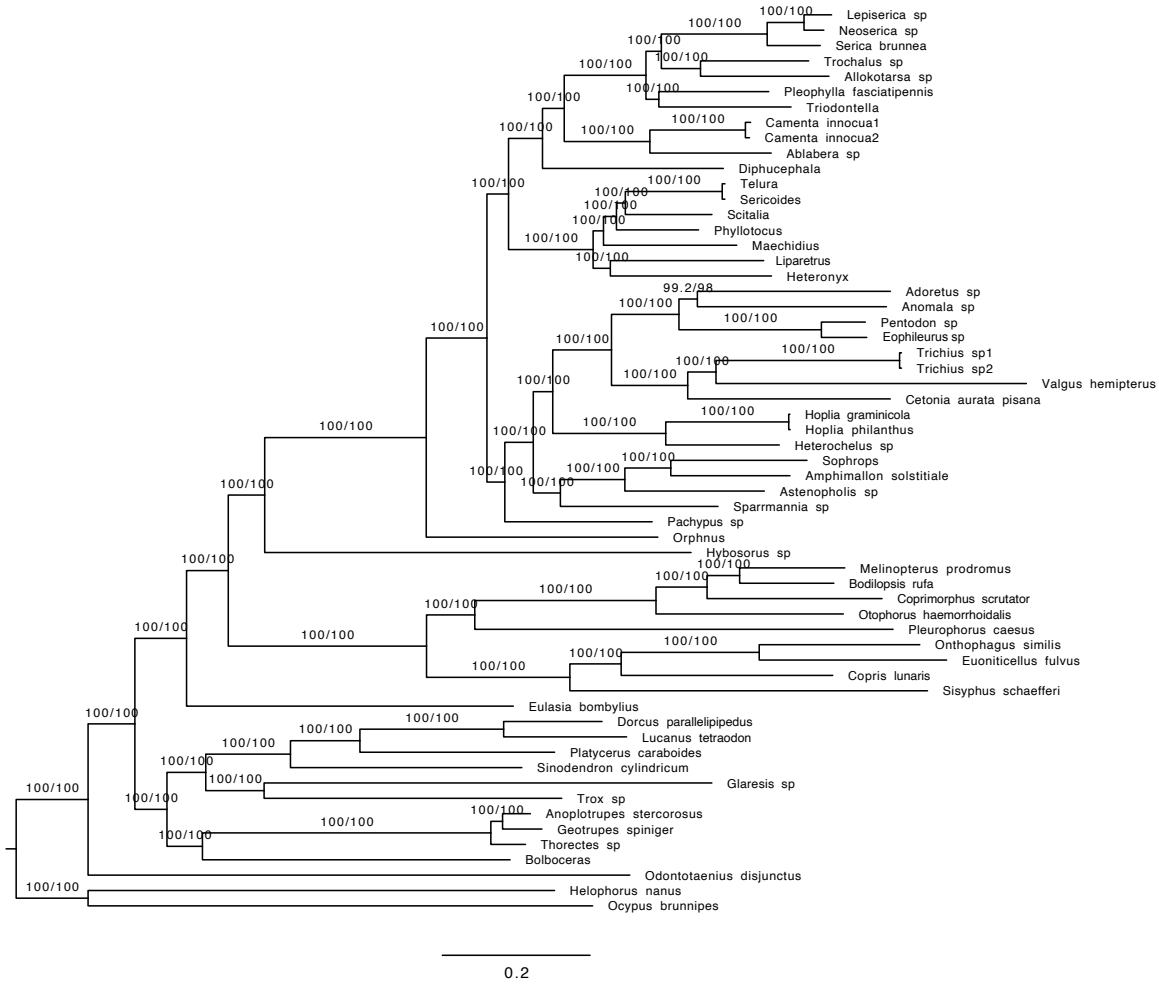


Figure S12. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; full dataset; Coleoptera orthologs).

Figure S13. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

Figure S14. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

Figure S15. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

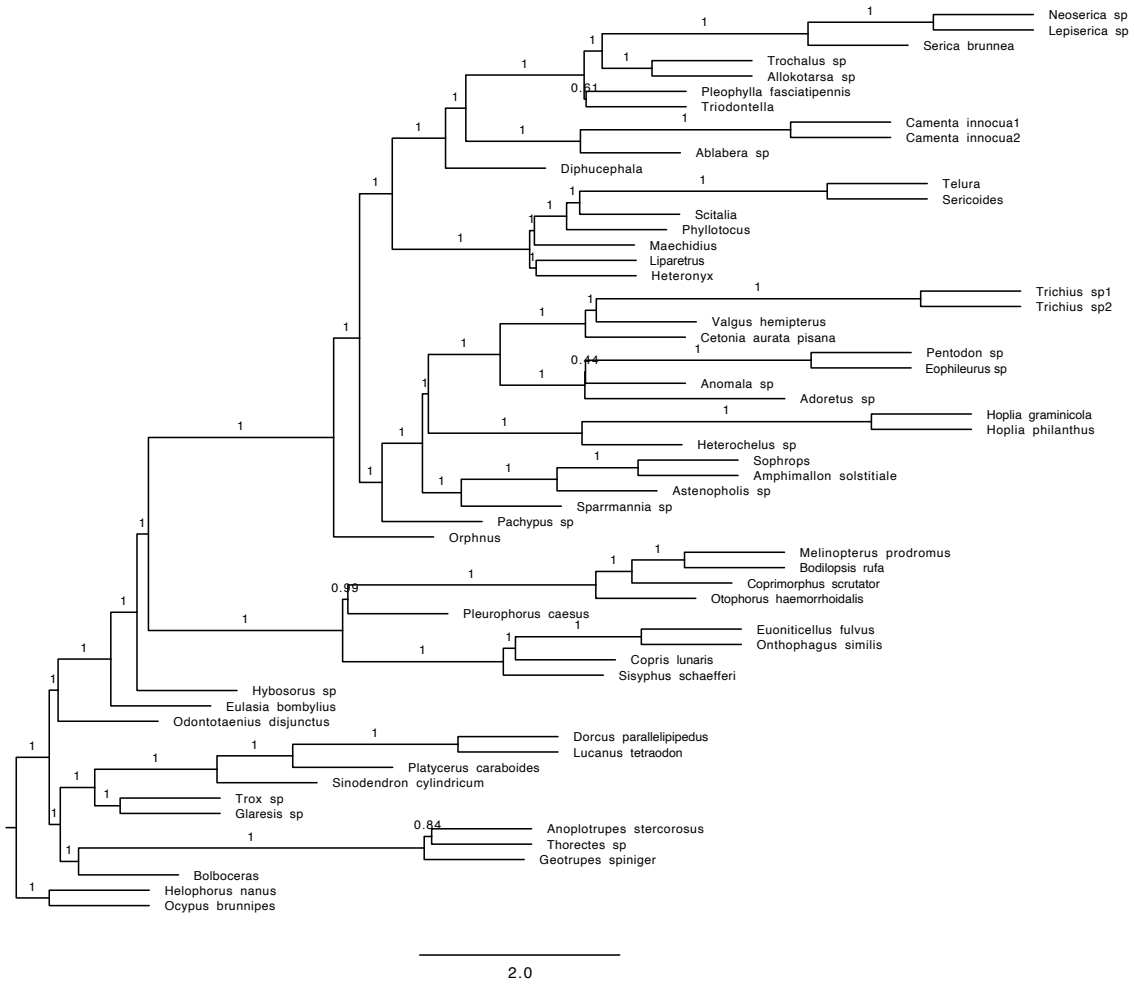


Figure S16. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (HmmAlign alignment; dataset with 70% complete data; Coleoptera orthologs).

Figure S17. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

Figure S18. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; dataset with 70% complete data; Coleoptera orthologs).

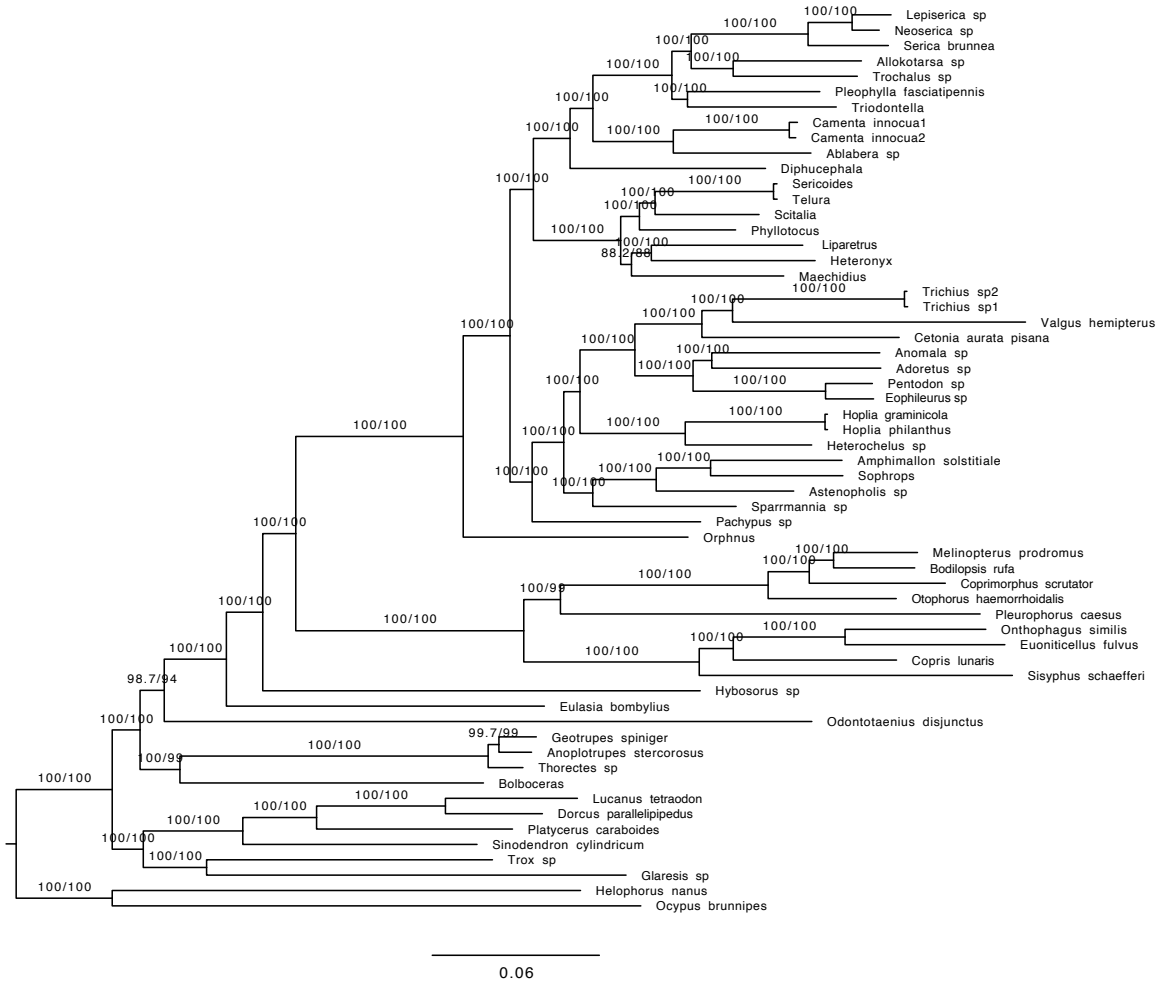


Figure S19. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

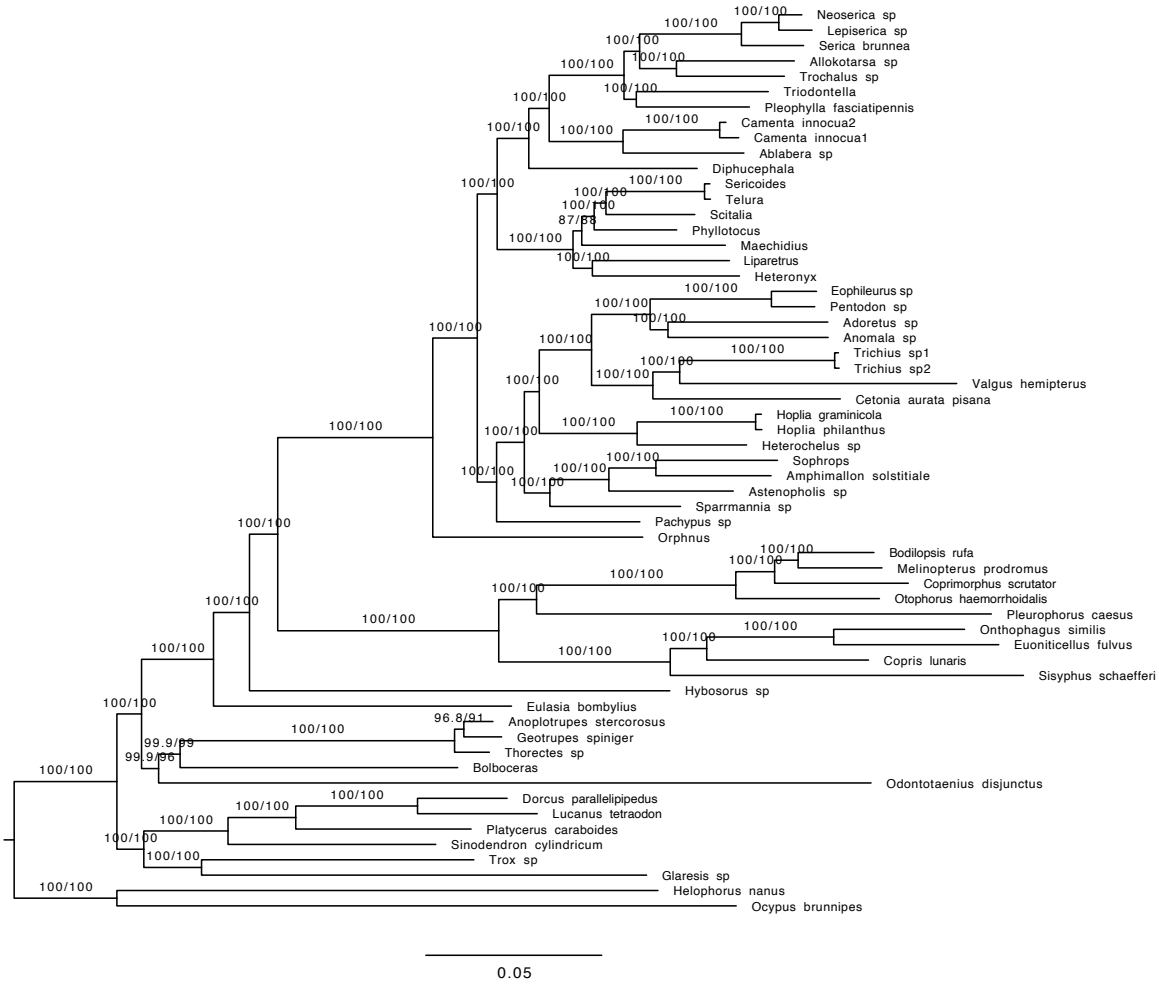


Figure S20. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

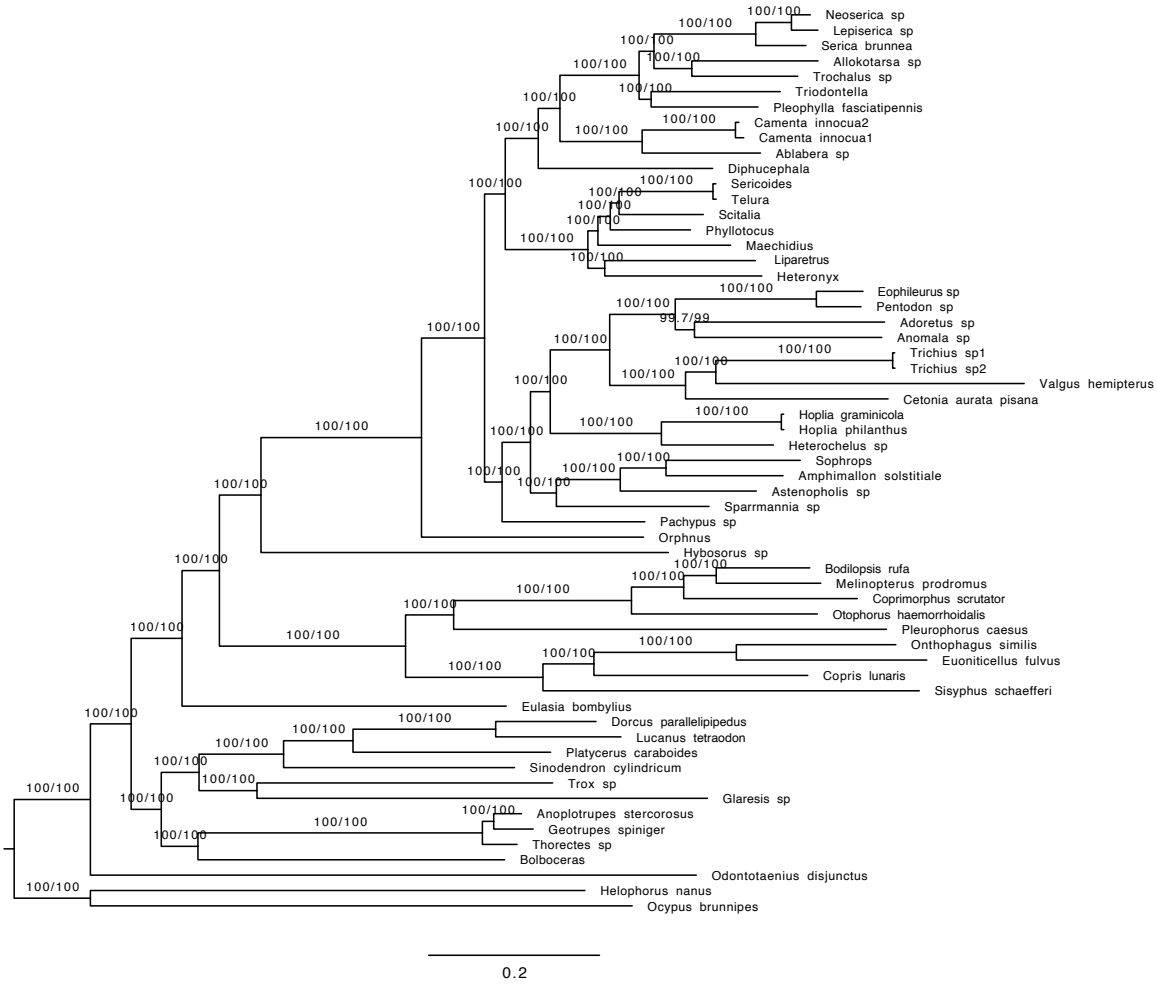


Figure S21. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (Mafft alignment; dataset with 70% complete data; Coleoptera orthologs).

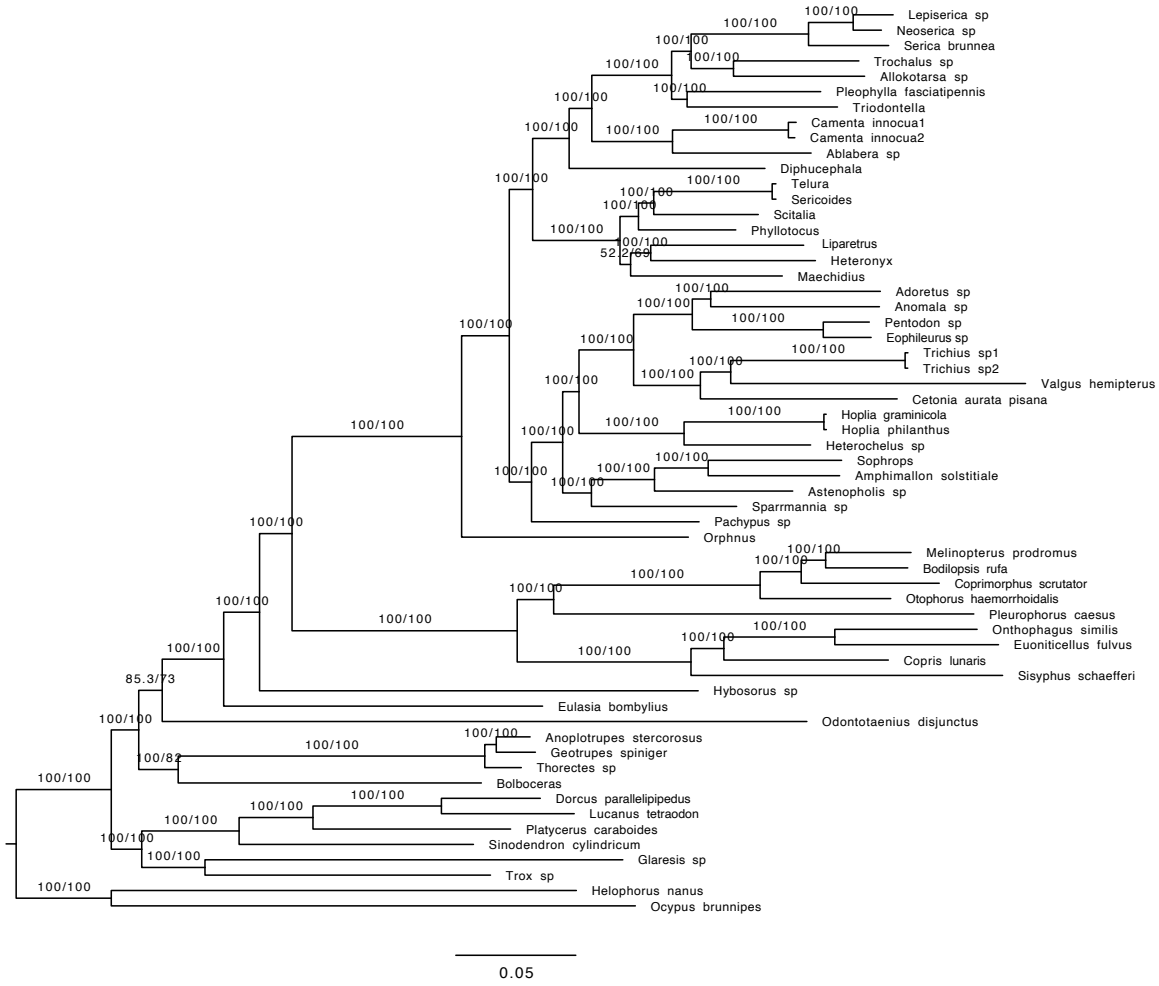


Figure S22. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (HmmAlign alignment; dataset with 70% complete data; Coleoptera orthologs).

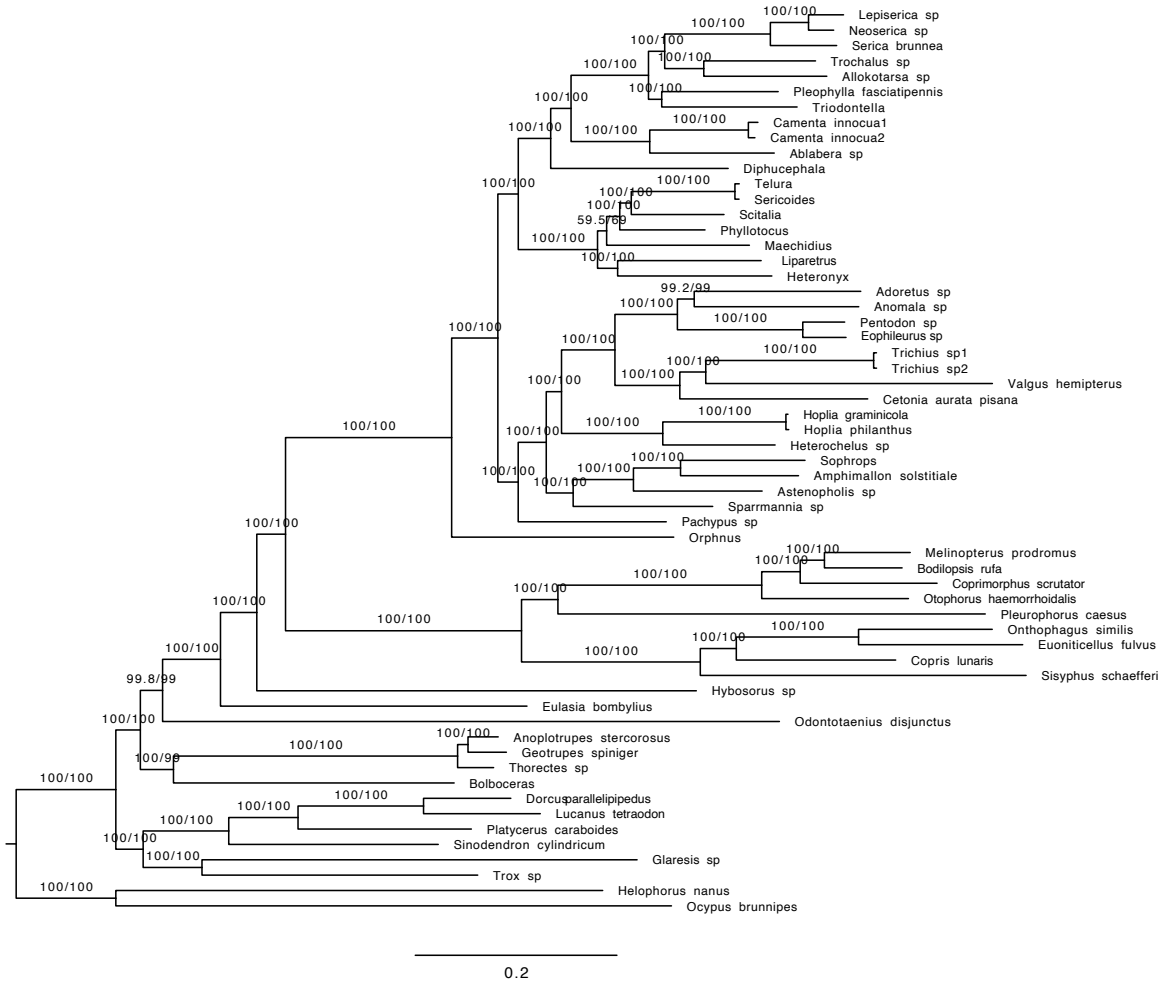


Figure S23. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; dataset with 70% complete data; Coleoptera orthologs).

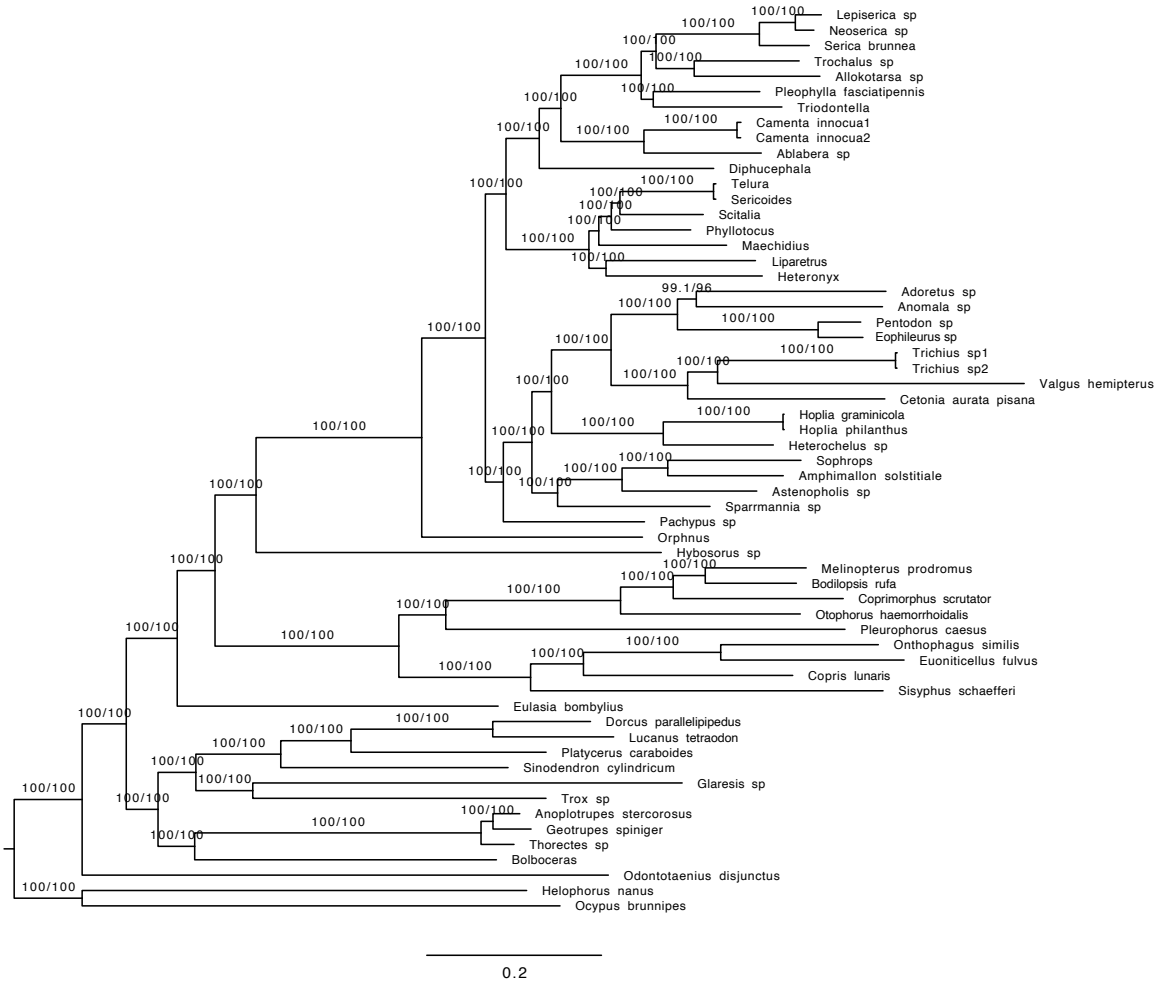


Figure S24. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; dataset with 70% complete data; Coleoptera orthologs).

Figure S25. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (Mafft alignment; fast evolving genes only; Coleoptera orthologs).

Figure S26. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; fast evolving genes only; Coleoptera orthologs).

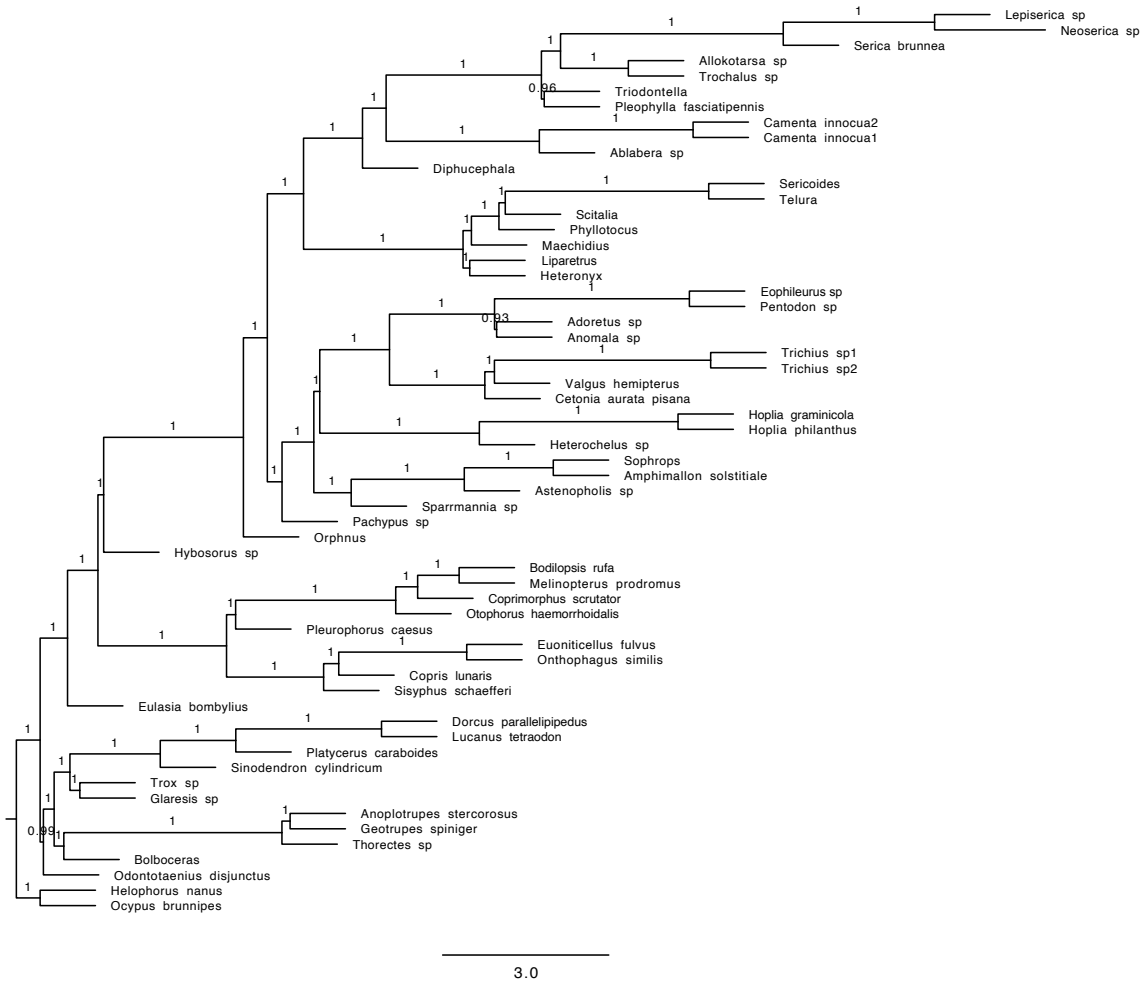


Figure S27. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (Mafft alignment; fast evolving genes only; Coleoptera orthologs).

Figure S28. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (HmmAlign alignment; fast evolving genes only; Coleoptera orthologs).

Figure S29. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; fast evolving genes only; Coleoptera orthologs).

Figure S30. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; fast evolving genes only; Coleoptera orthologs).

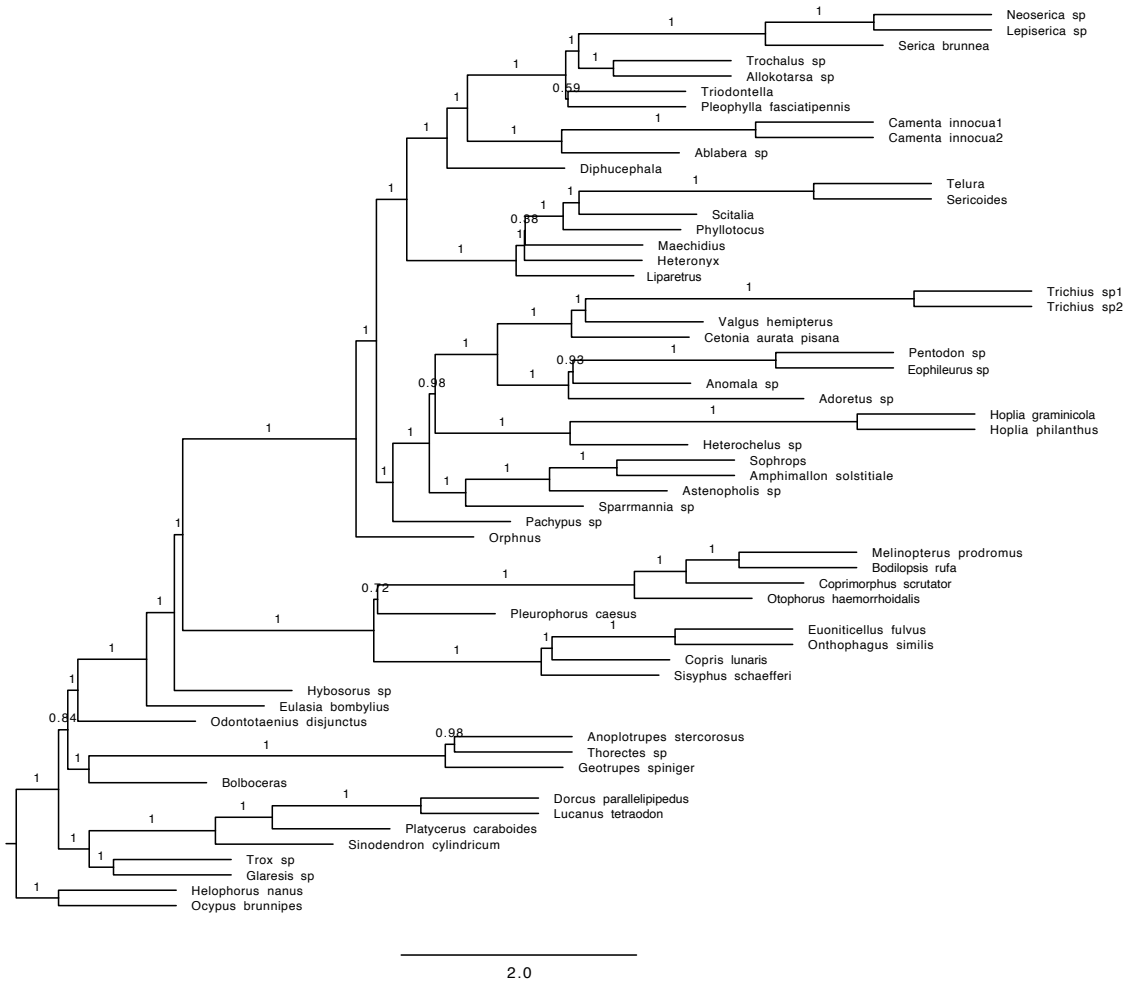


Figure S31. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (Mafft alignment; slowly evolving genes only; Coleoptera orthologs).

Figure S32. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; slowly evolving genes only; Coleoptera orthologs).

Figure S33. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (Mafft alignment; slowly evolving genes only; Coleoptera orthologs).

Figure S34. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (HmmAlign alignment; slowly evolving genes only; Coleoptera orthologs).

Figure S35. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; slowly evolving genes only; Coleoptera orthologs).

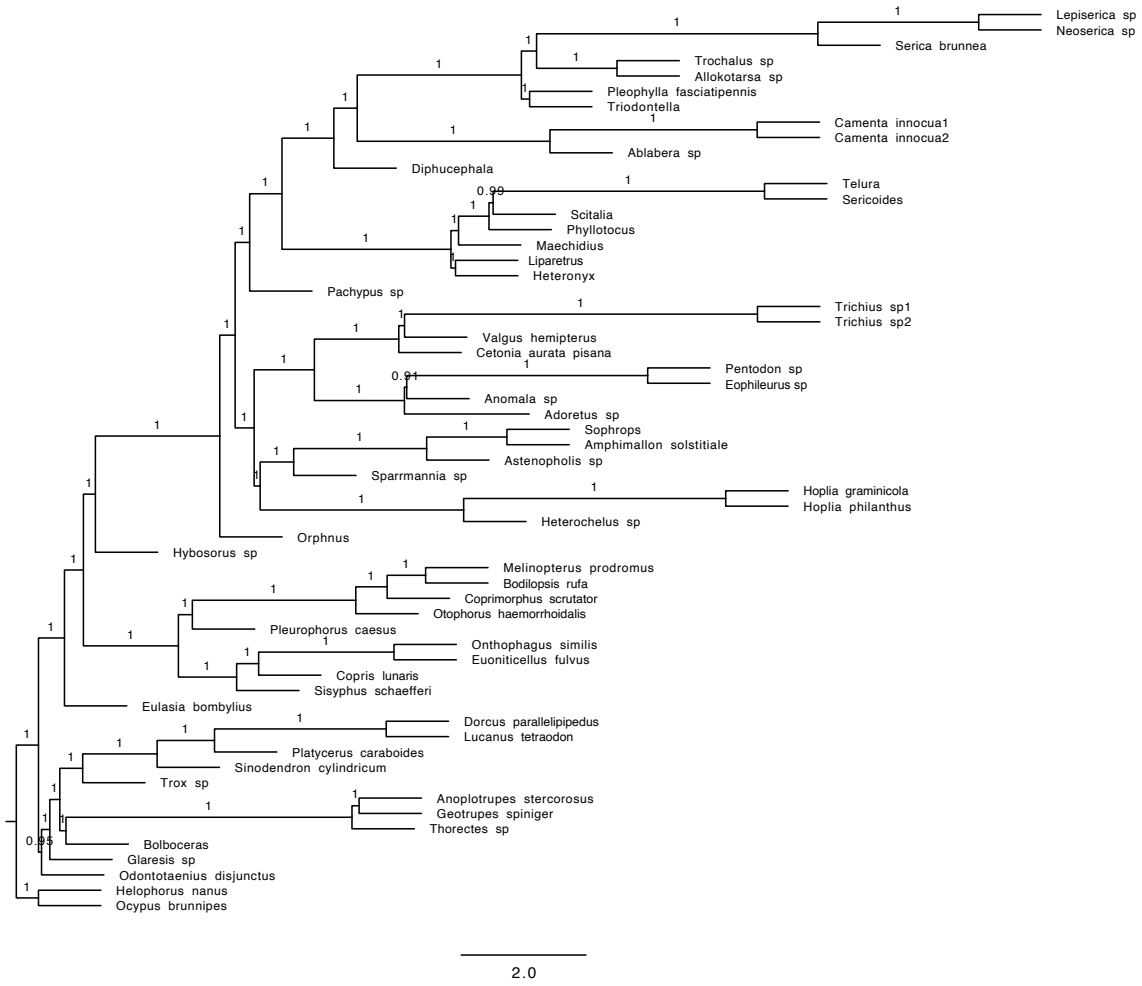


Figure S36. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; slowly evolving genes only; Coleoptera orthologs).

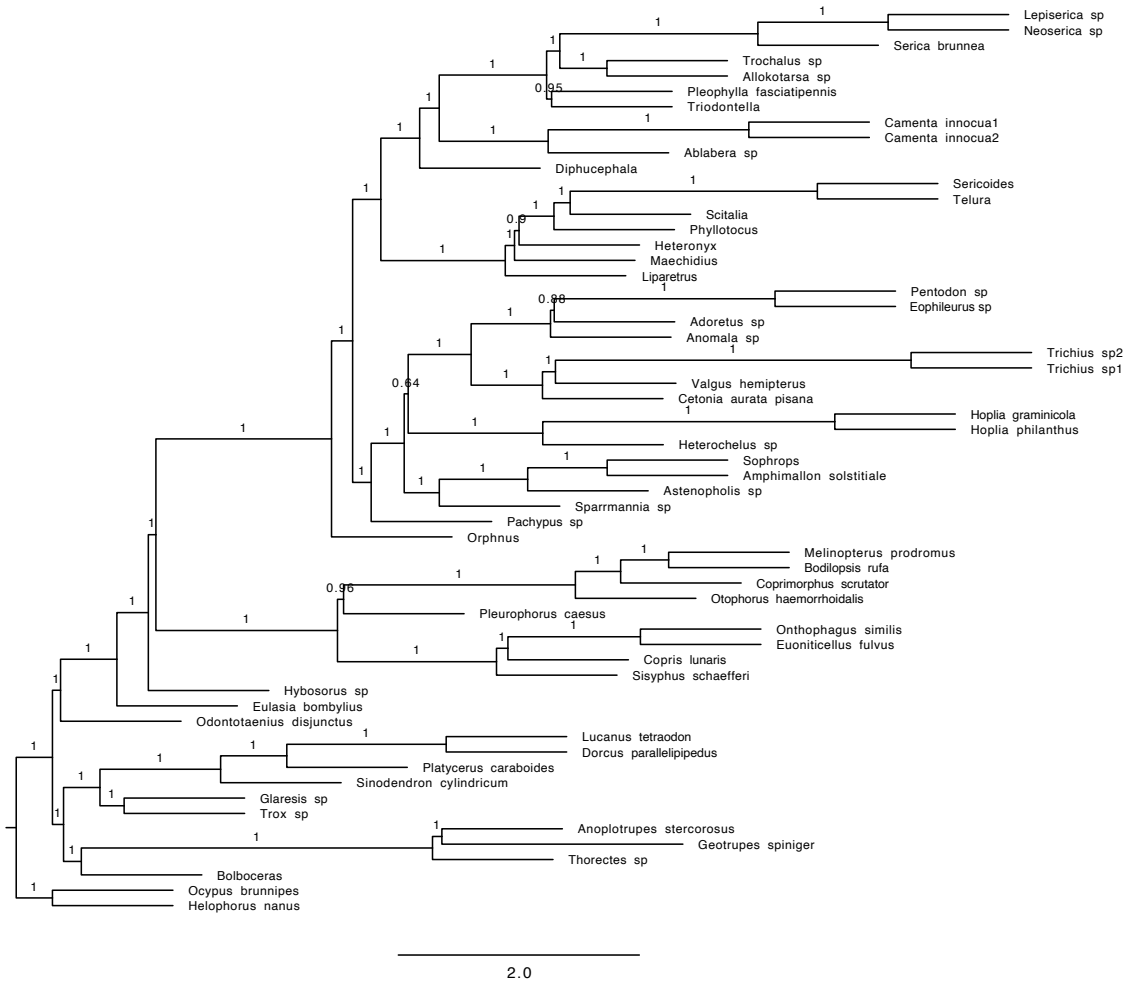


Figure S37. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (Mafft alignment; full dataset; Endopterygota orthologs).

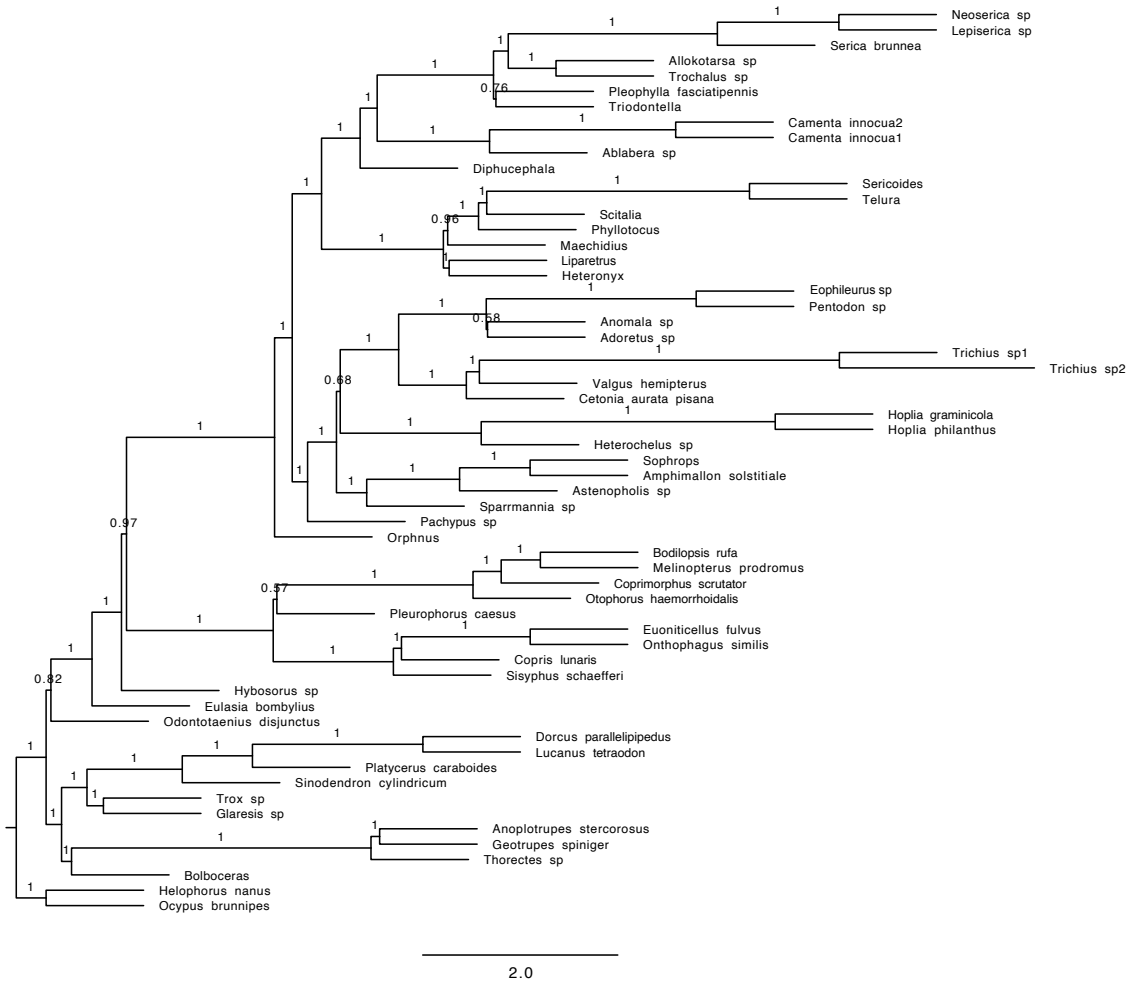


Figure S38. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; full dataset; Endopterygota orthologs).

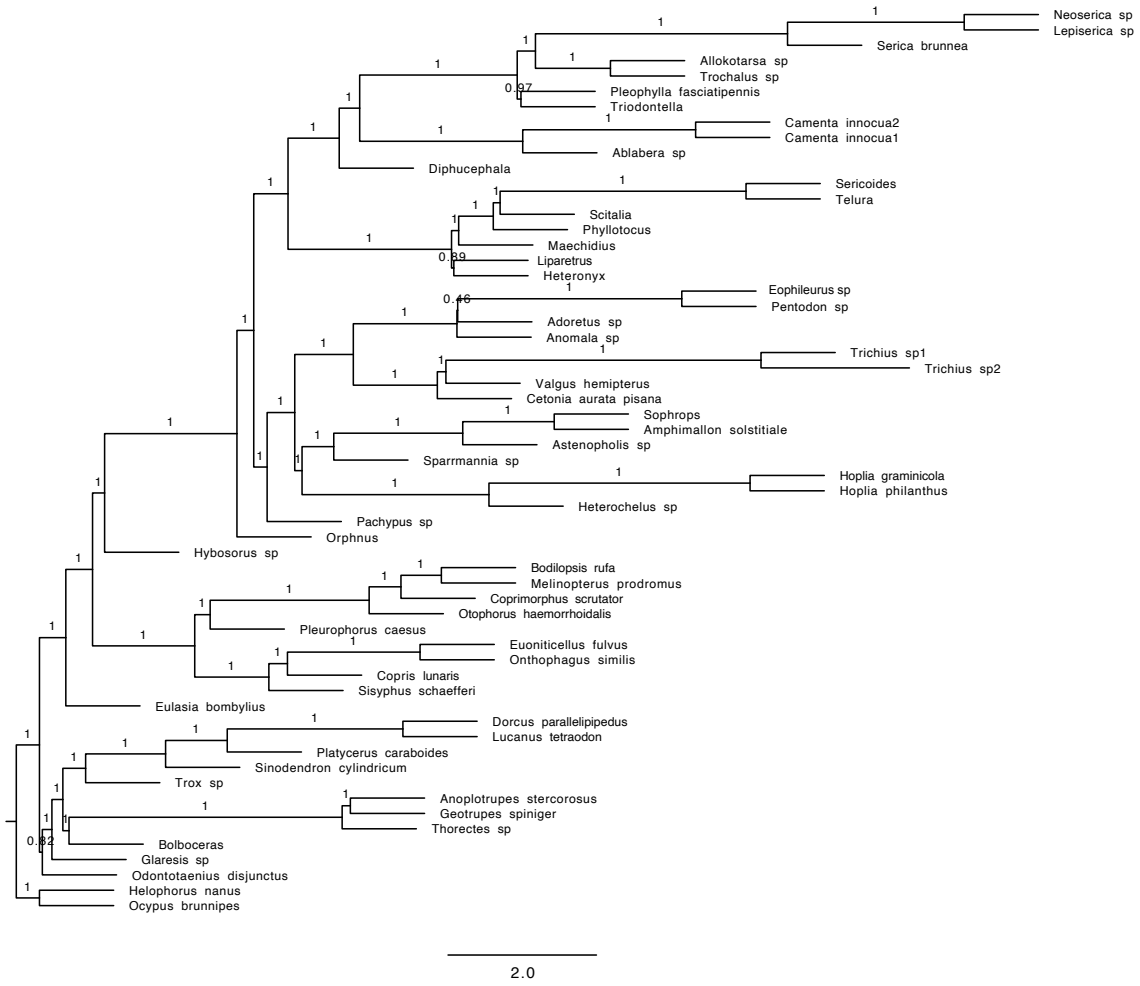


Figure S39. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (Mafft alignment; full dataset; Endopterygota orthologs).

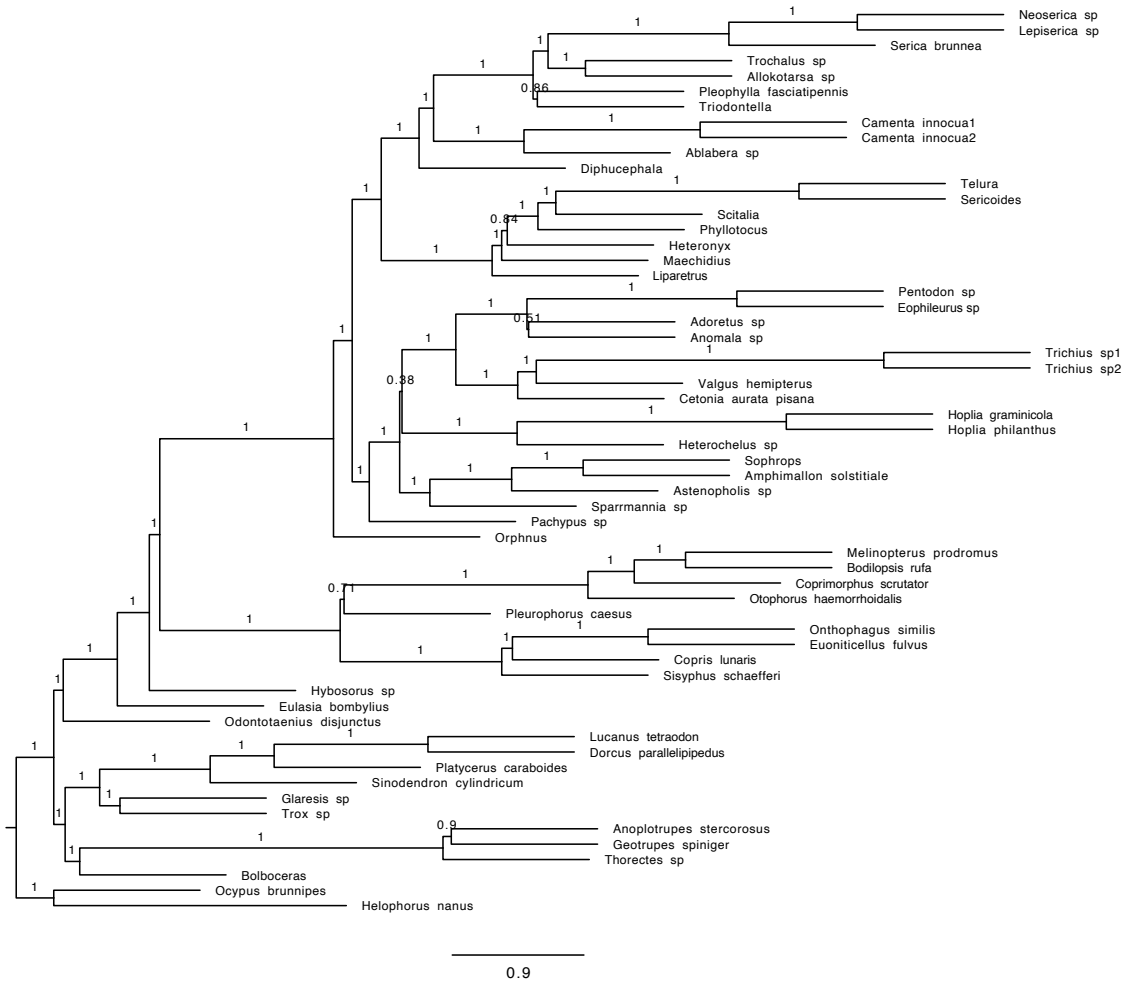


Figure S40. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (HmmAlign alignment; full dataset; Endopterygota orthologs).

Figure S41. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; full dataset; Endopterygota orthologs).

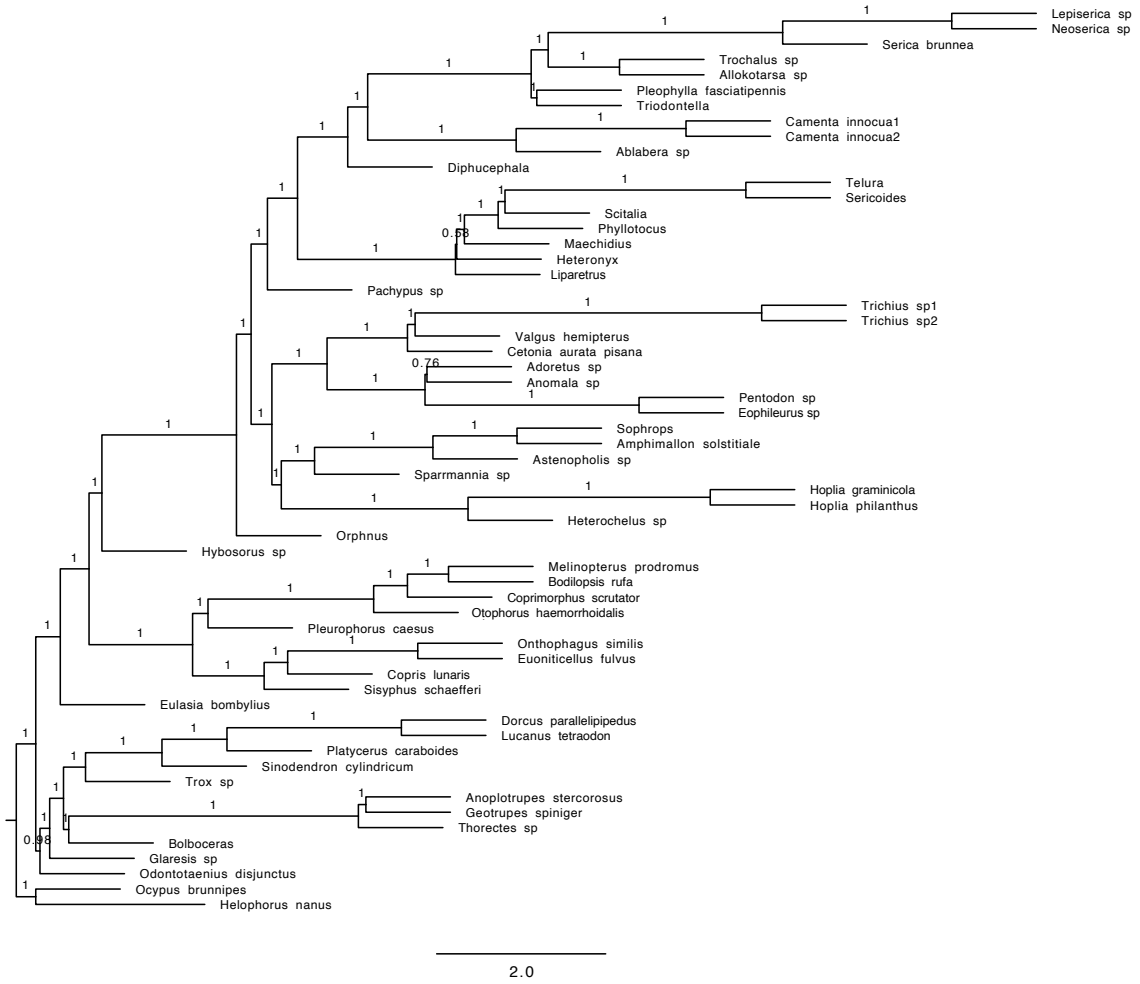


Figure S42. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; full dataset; Endopterygota orthologs).

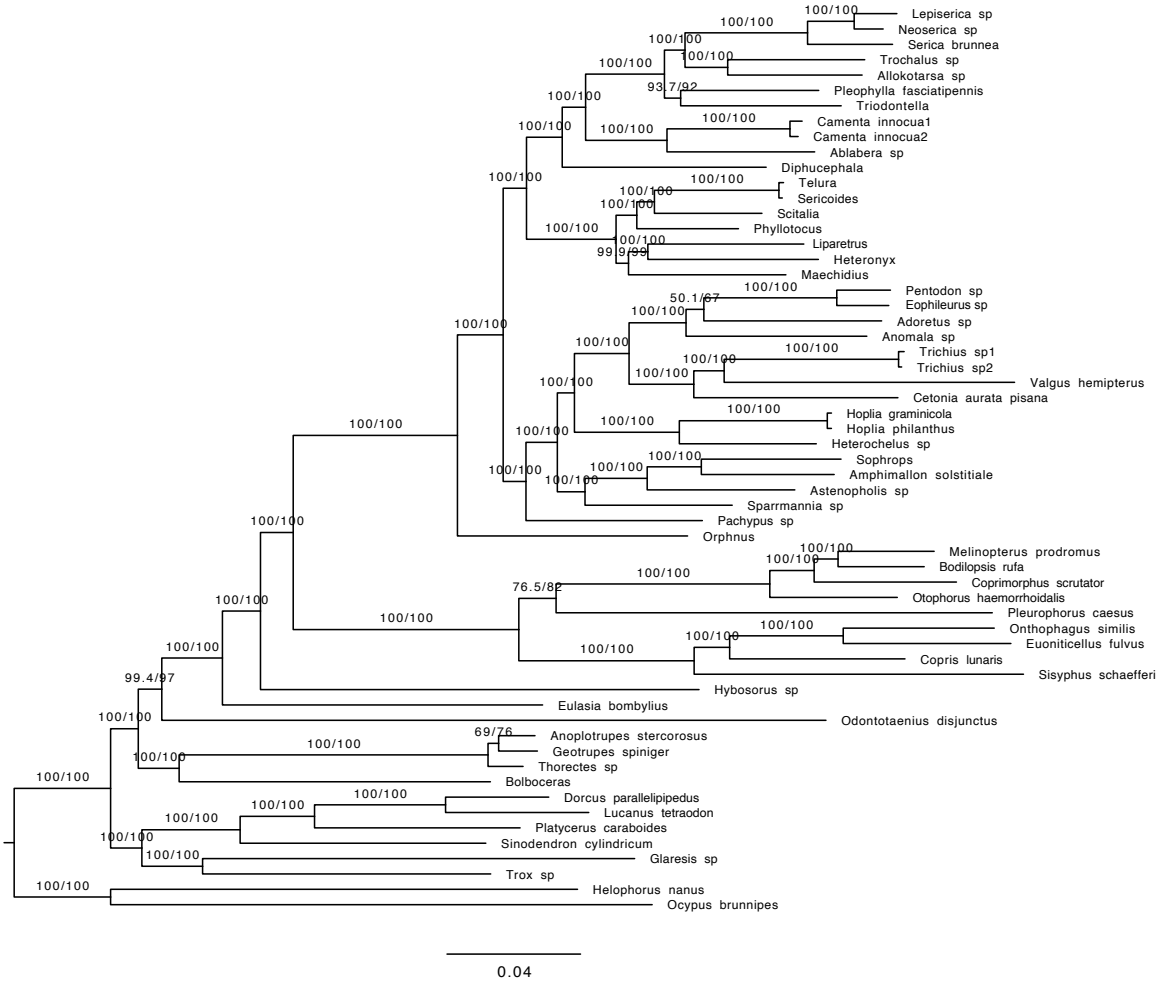


Figure S43. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (Mafft alignment; full dataset; Endopterygota orthologs).

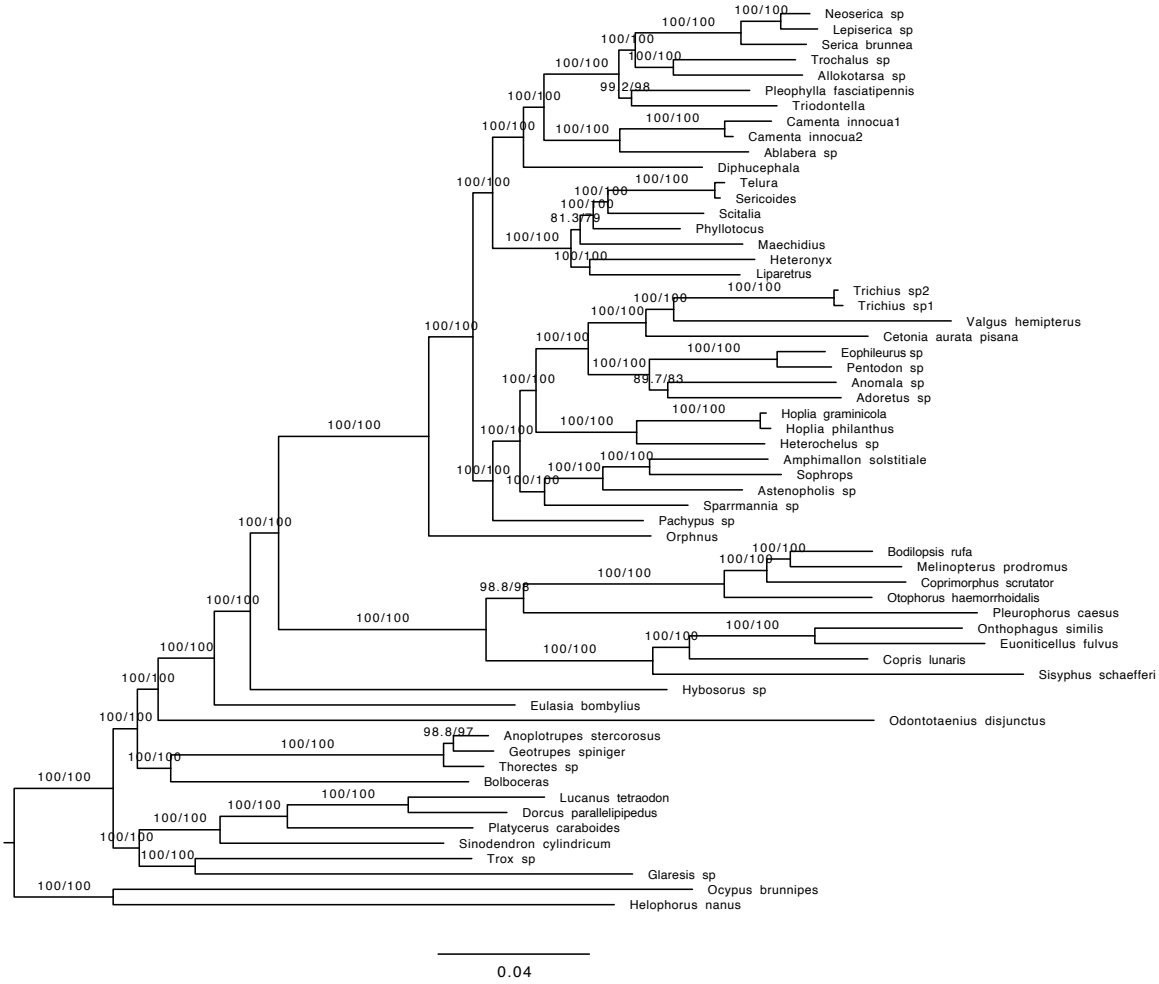


Figure S44. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; full dataset; Endopterygota orthologs).

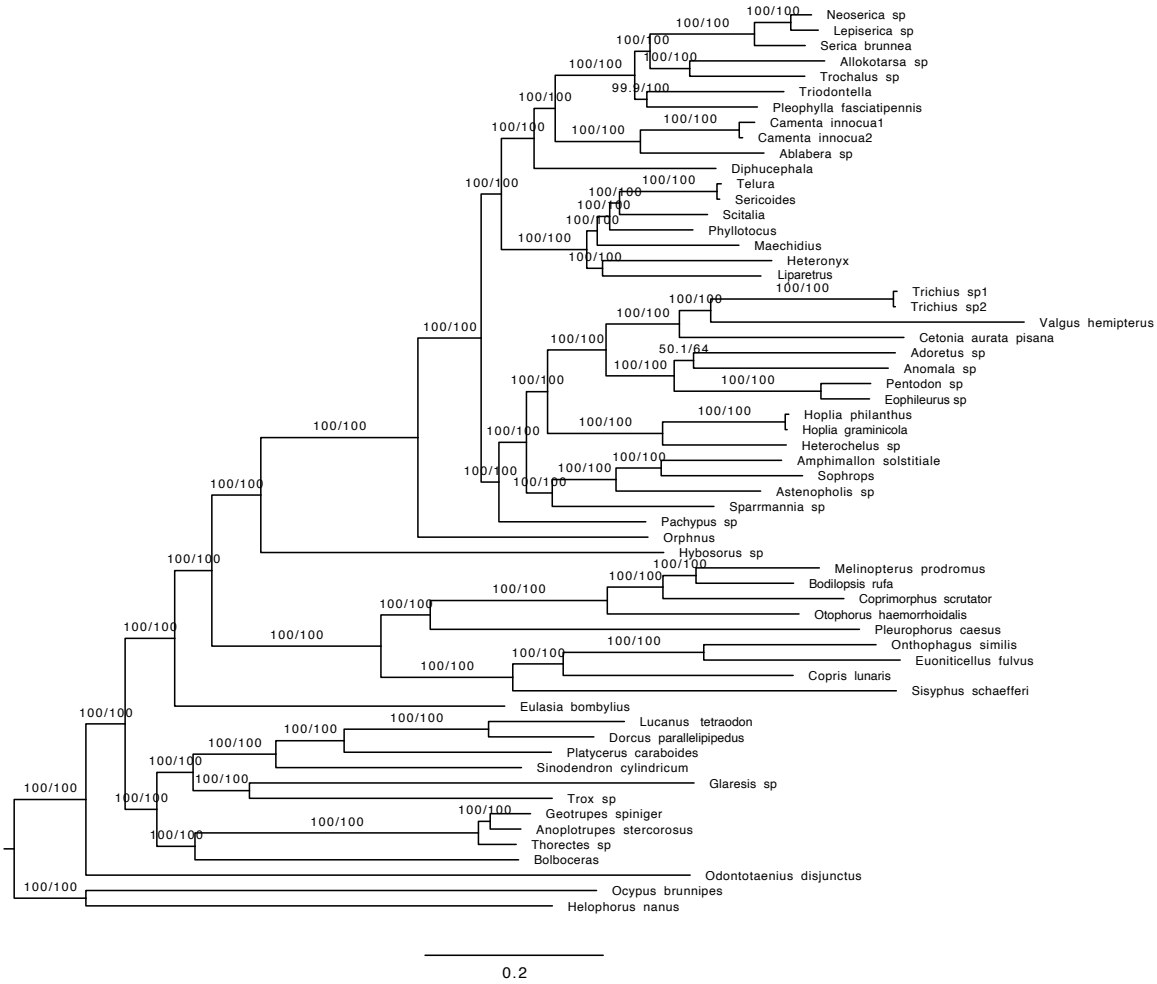


Figure S45. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (Mafft alignment; full dataset; Endopterygota orthologs).

Figure S46. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (HmmAlign alignment; full dataset; Endopterygota orthologs).

Figure S47. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; full dataset; Endopterygota orthologs).

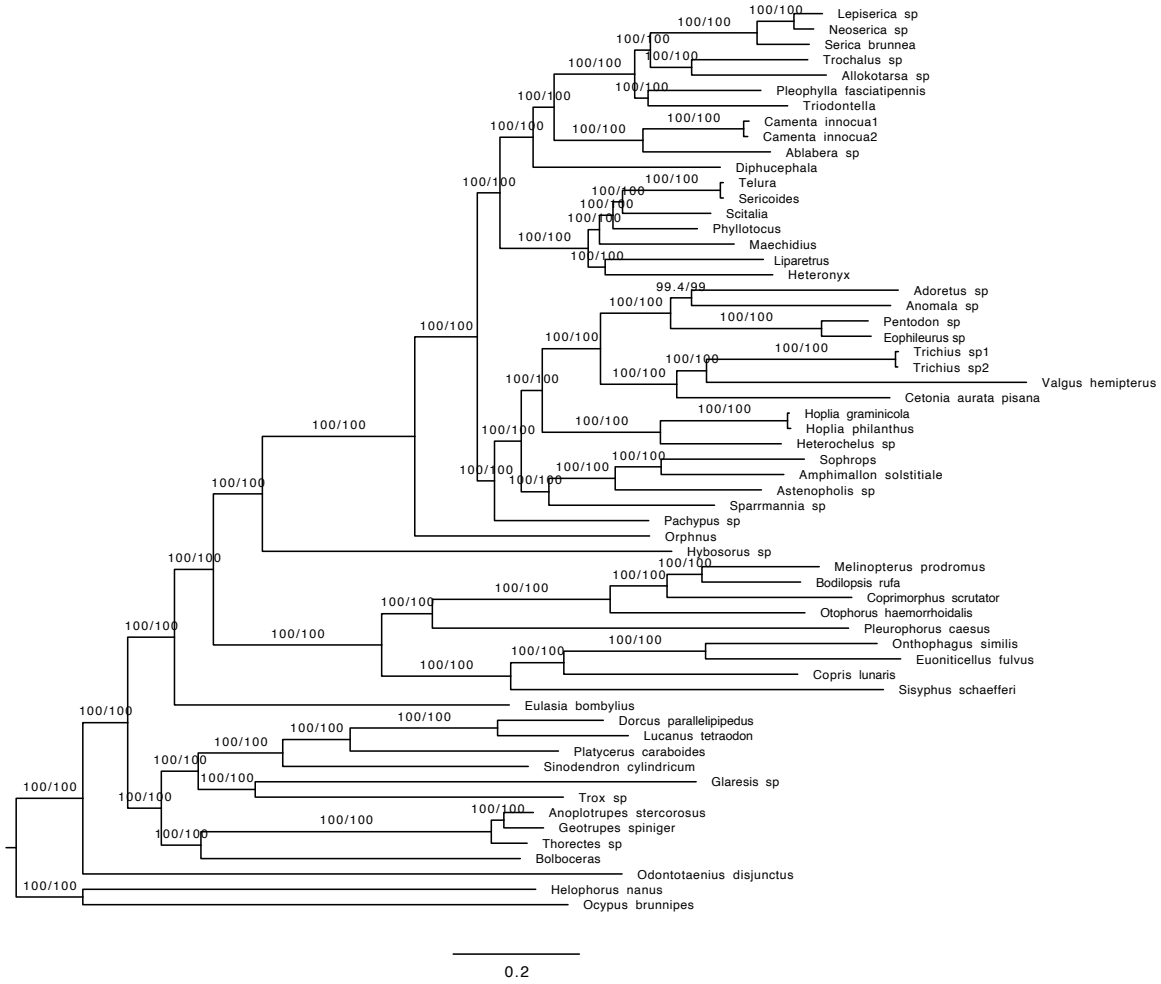


Figure S48. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; full dataset; Endopterygota orthologs).

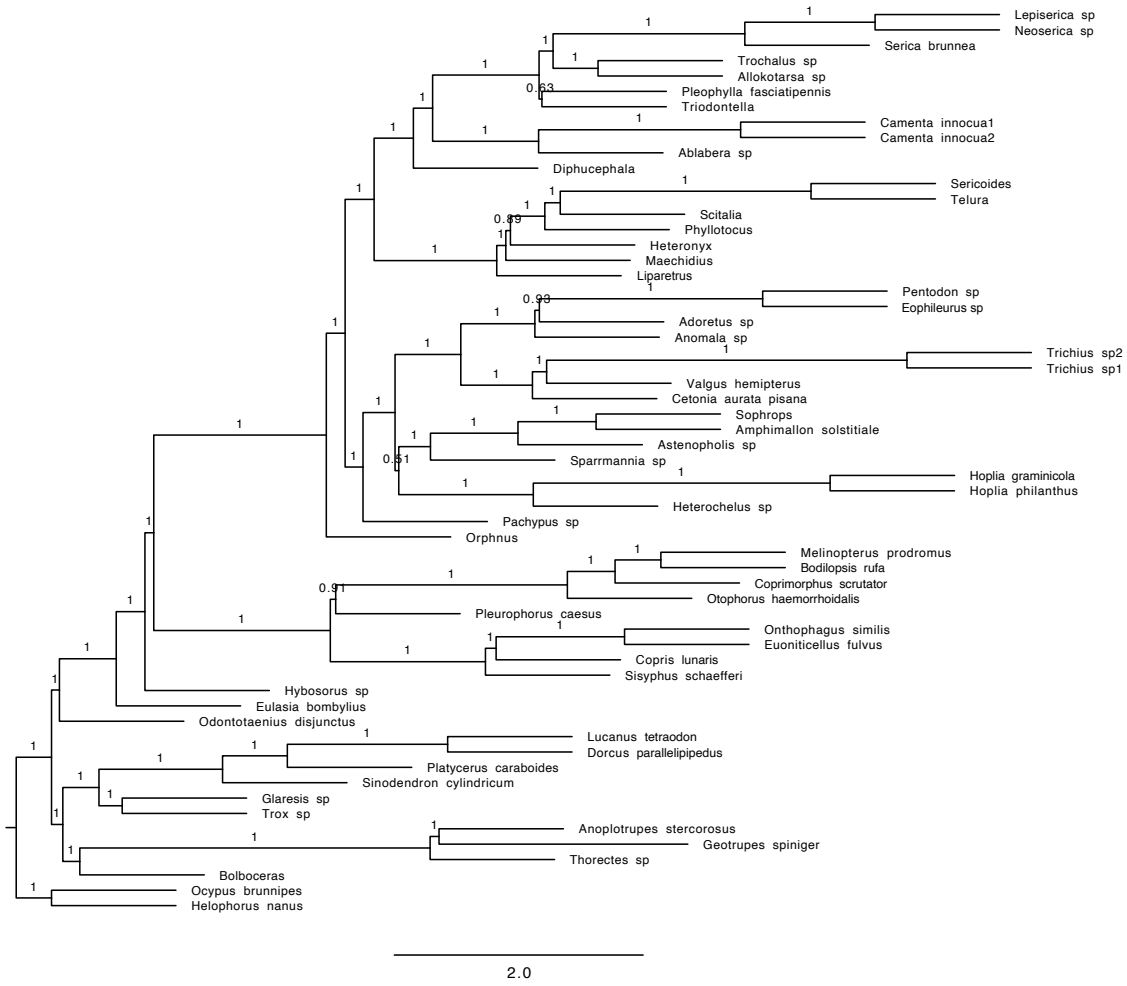


Figure S49. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (Mafft alignment; dataset with 70% complete data; Endopterygota orthologs).

Figure S50. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; dataset with 70% complete data; Endopterygota orthologs).

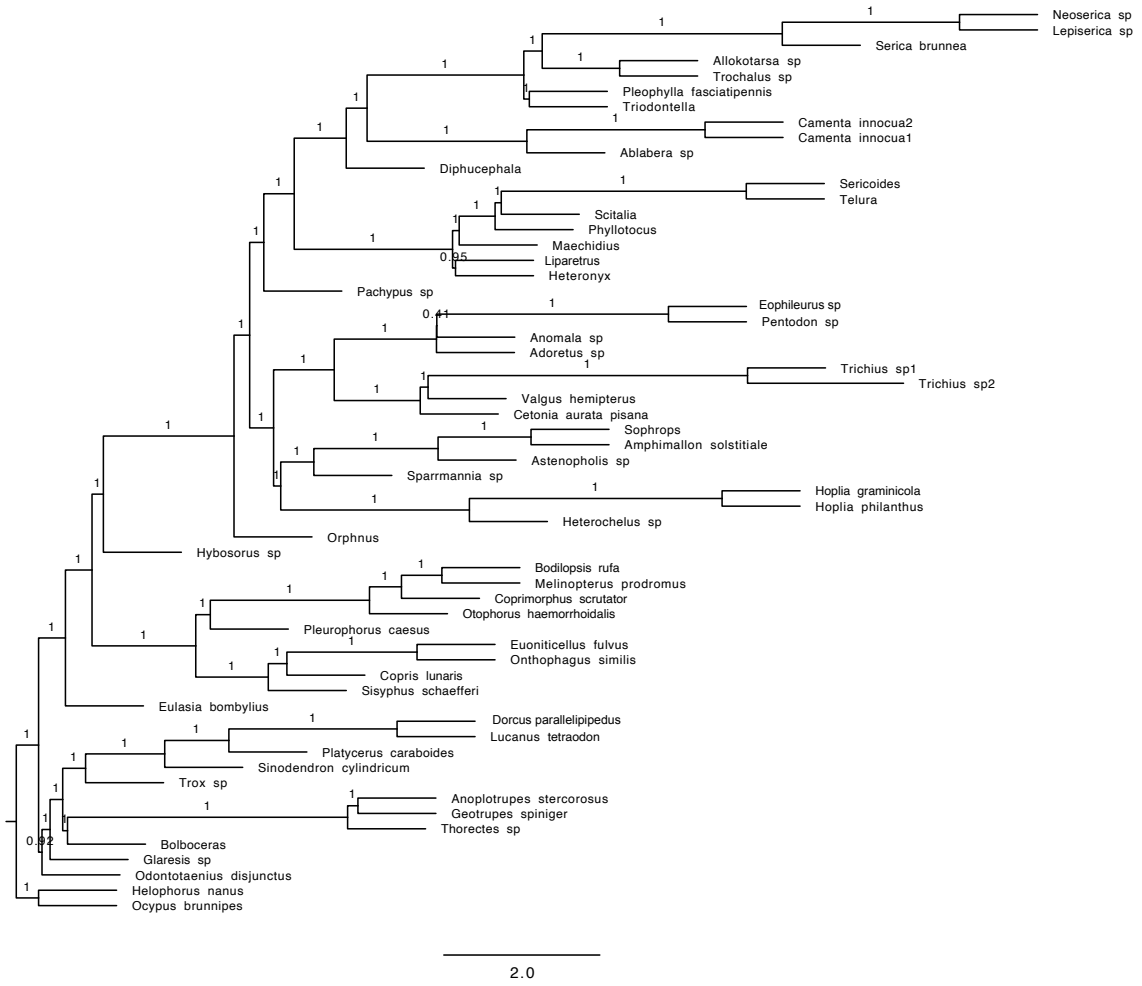


Figure S51. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (Mafft alignment; dataset with 70% complete data; Endopterygota orthologs).

Figure S52. Phylogenetic tree obtained with the coalescent tree search with Astral and amino acid sequences (HmmAlign alignment; dataset with 70% complete data; Endopterygota orthologs).

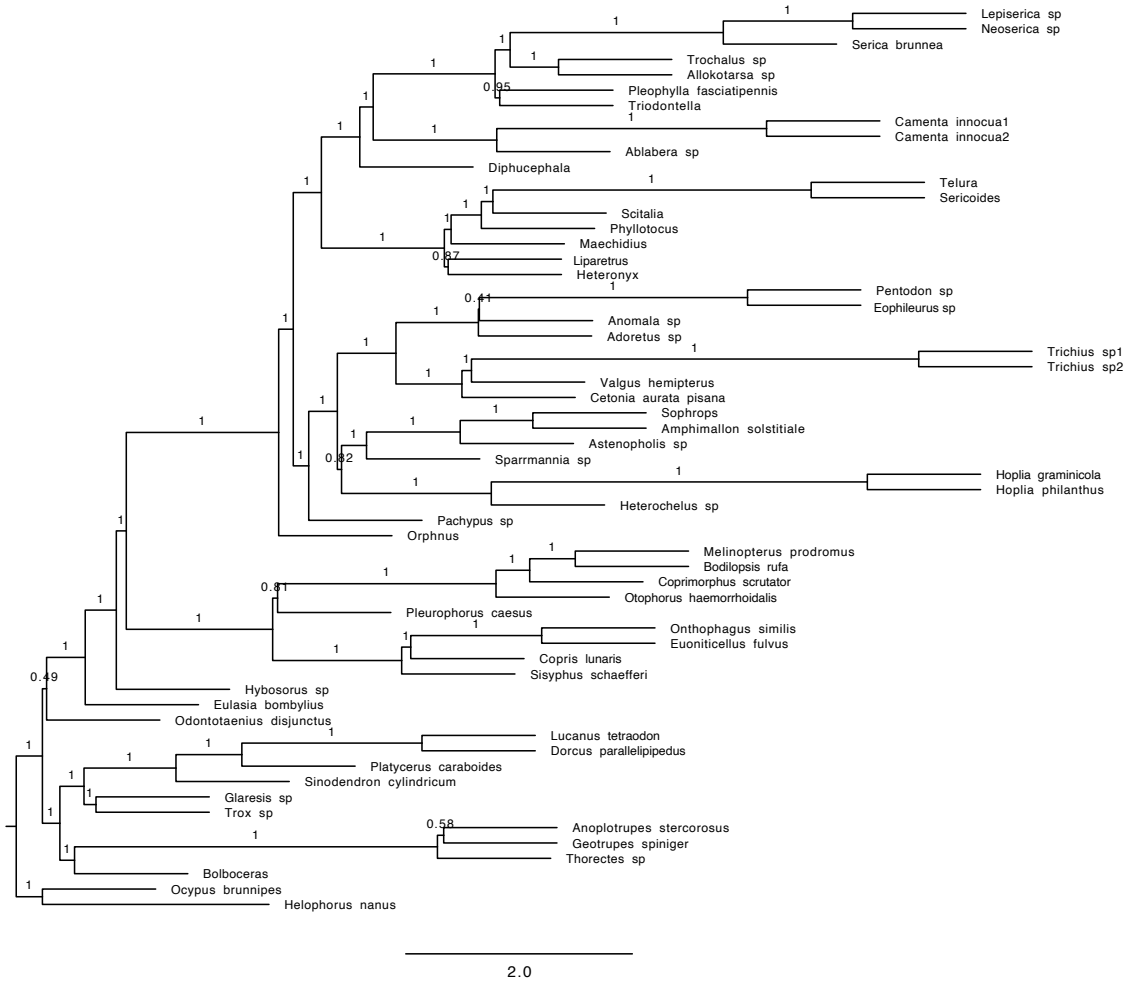


Figure S53. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; dataset with 70% complete data; Endopterygota orthologs).

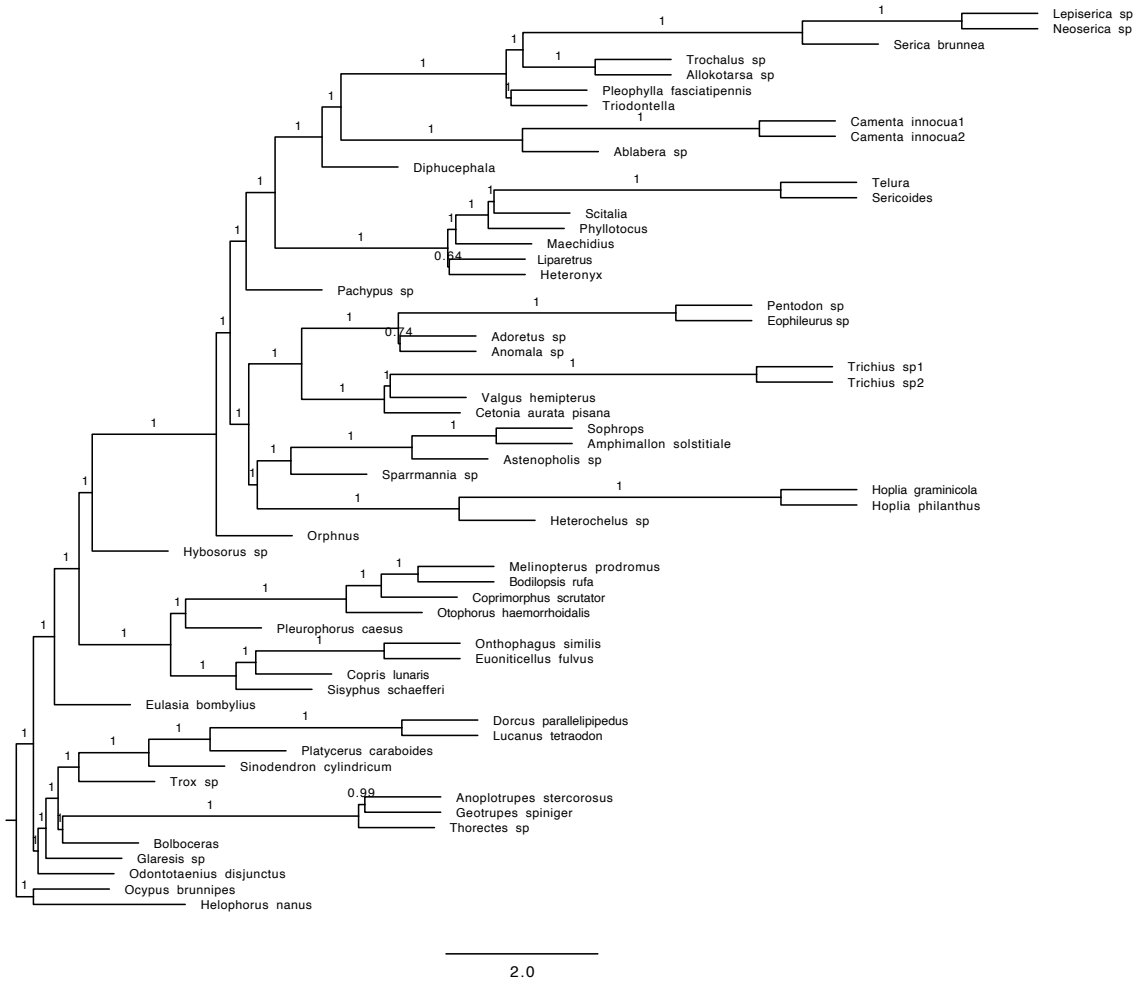


Figure S54. Phylogenetic tree obtained with the coalescent tree search with Astral and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; dataset with 70% complete data; Endopterygota orthologs).

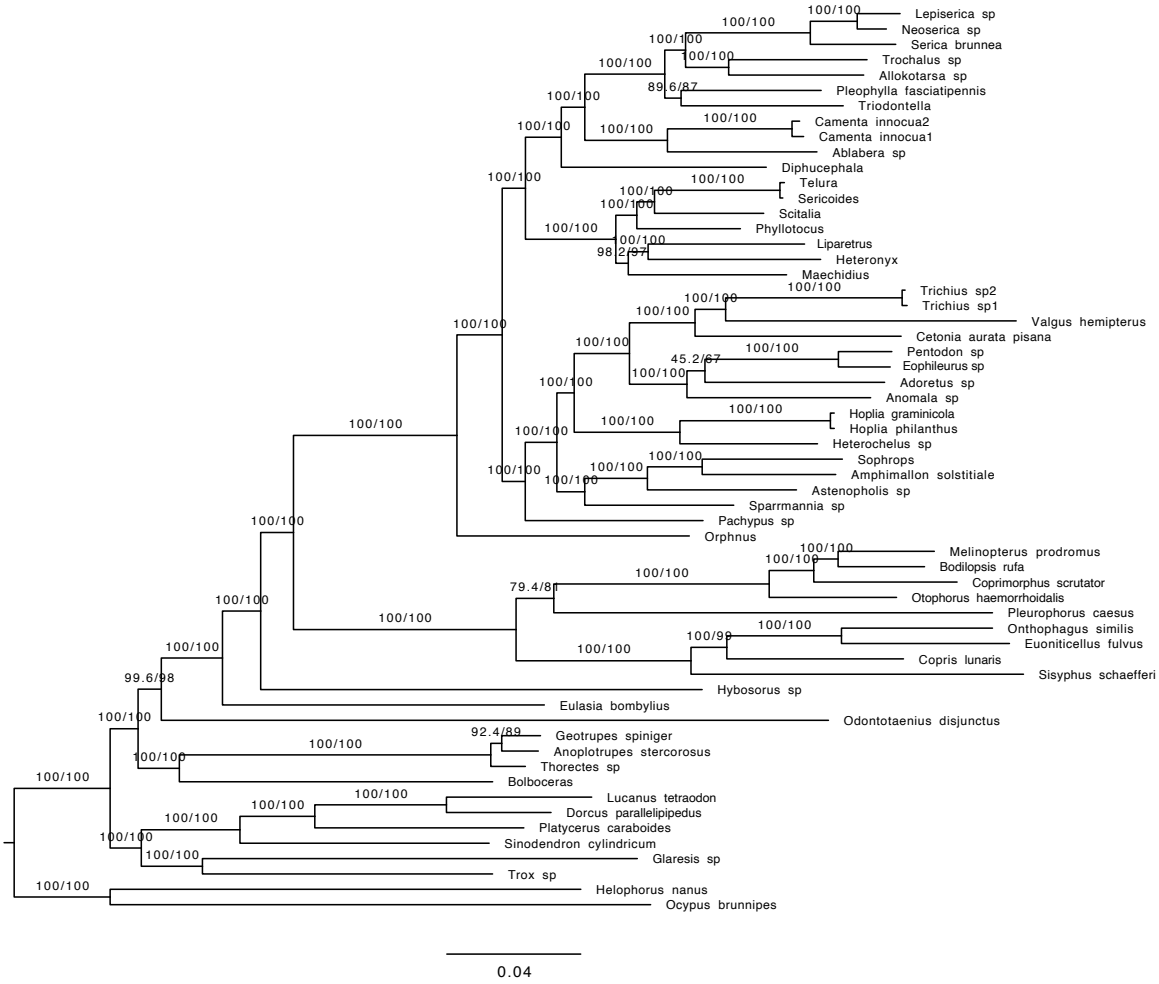


Figure S55. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) (Mafft alignment; dataset with 70% complete data; Endopterygota orthologs).

Figure S56. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (Mafft alignment; dataset with 70% complete data; Endopterygota orthologs).

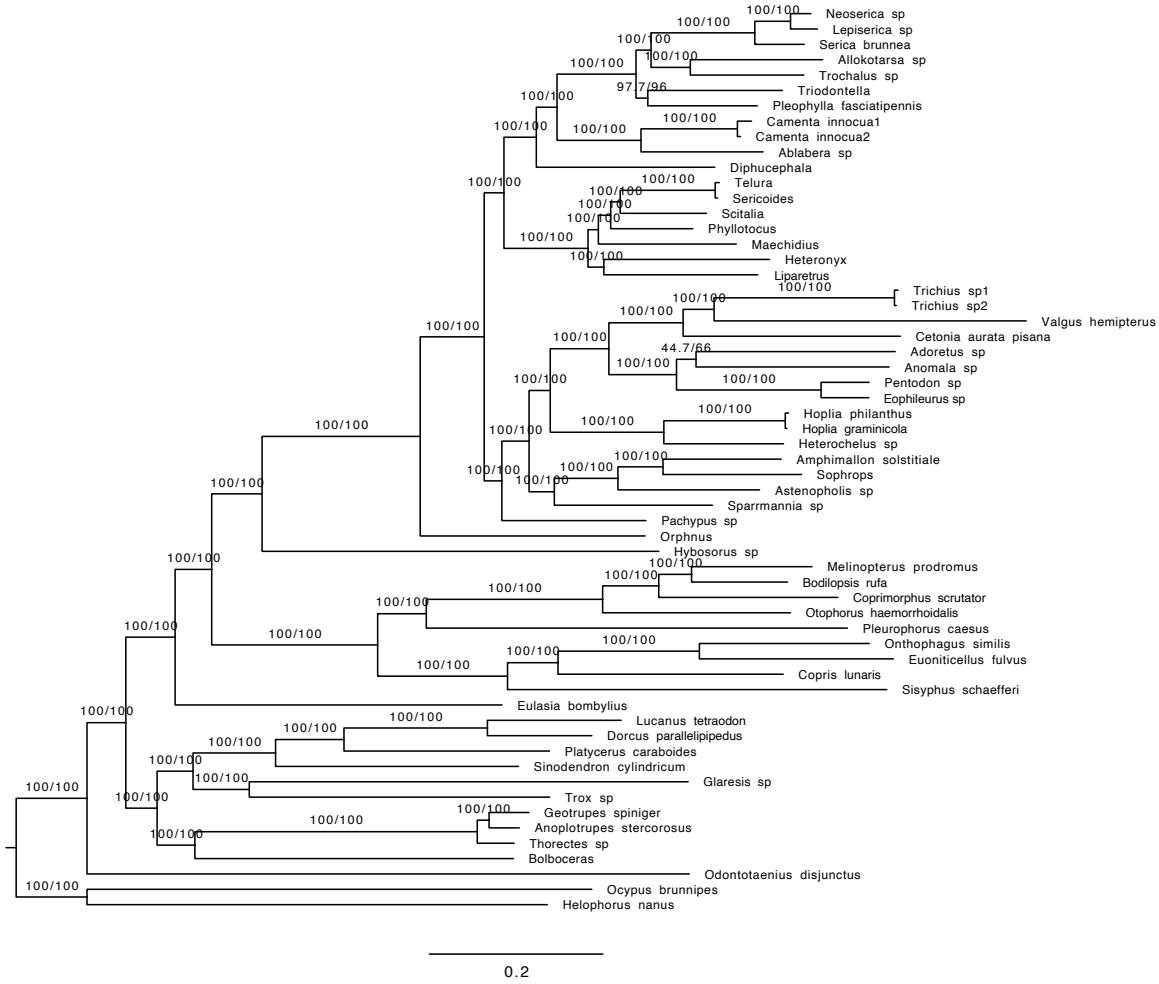


Figure S57. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (Mafft alignment; dataset with 70% complete data; Endopterygota orthologs).

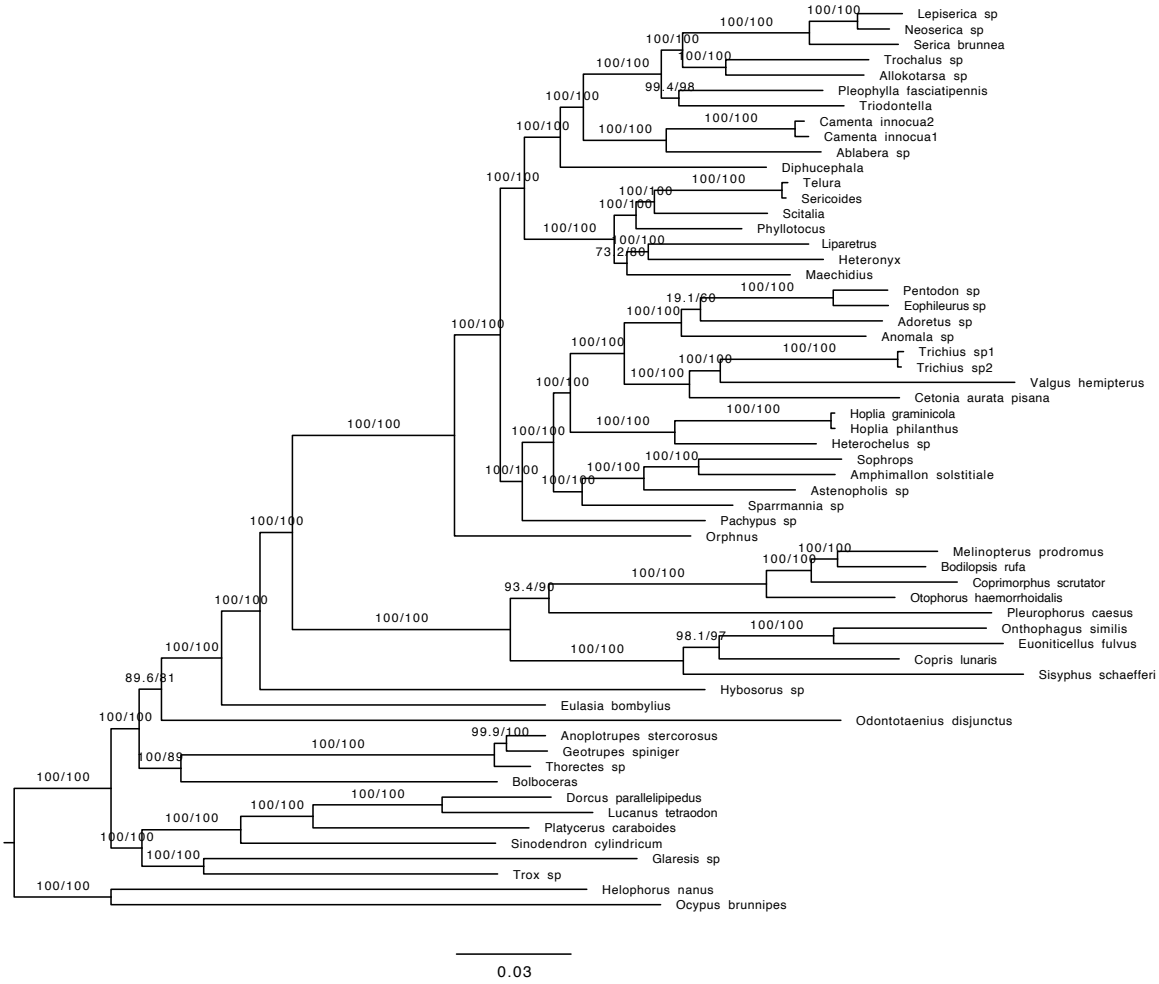


Figure S58. Phylogenetic tree obtained with the concatenated data and amino acid sequences (AA) HmmAlign alignment; dataset with 70% complete data; Endopterygota orthologs).

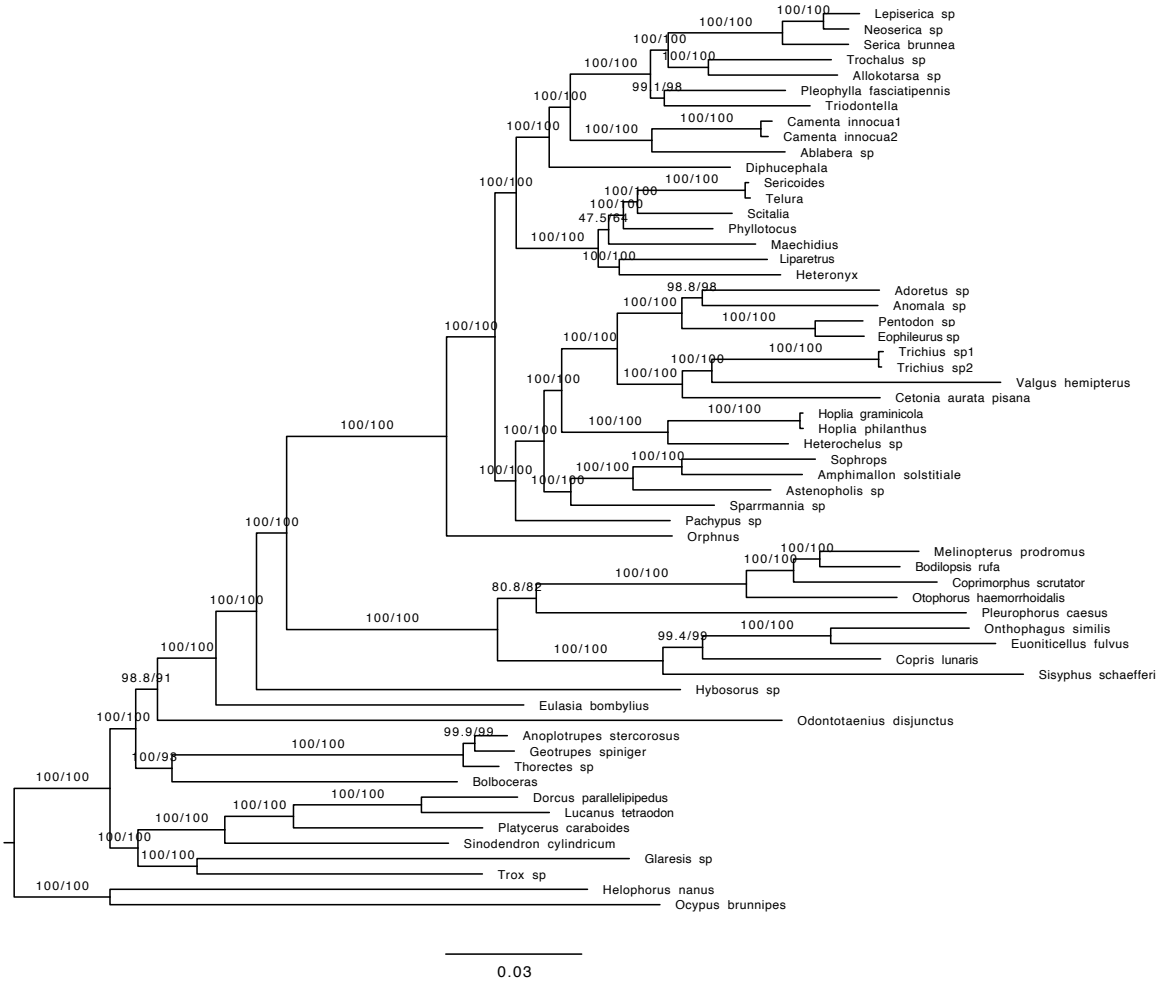


Figure S59. Phylogenetic tree obtained with the concatenated data and nucleotide data containing only the first and second base pair (nt12) (HmmAlign alignment; dataset with 70% complete data; Endopterygota orthologs).

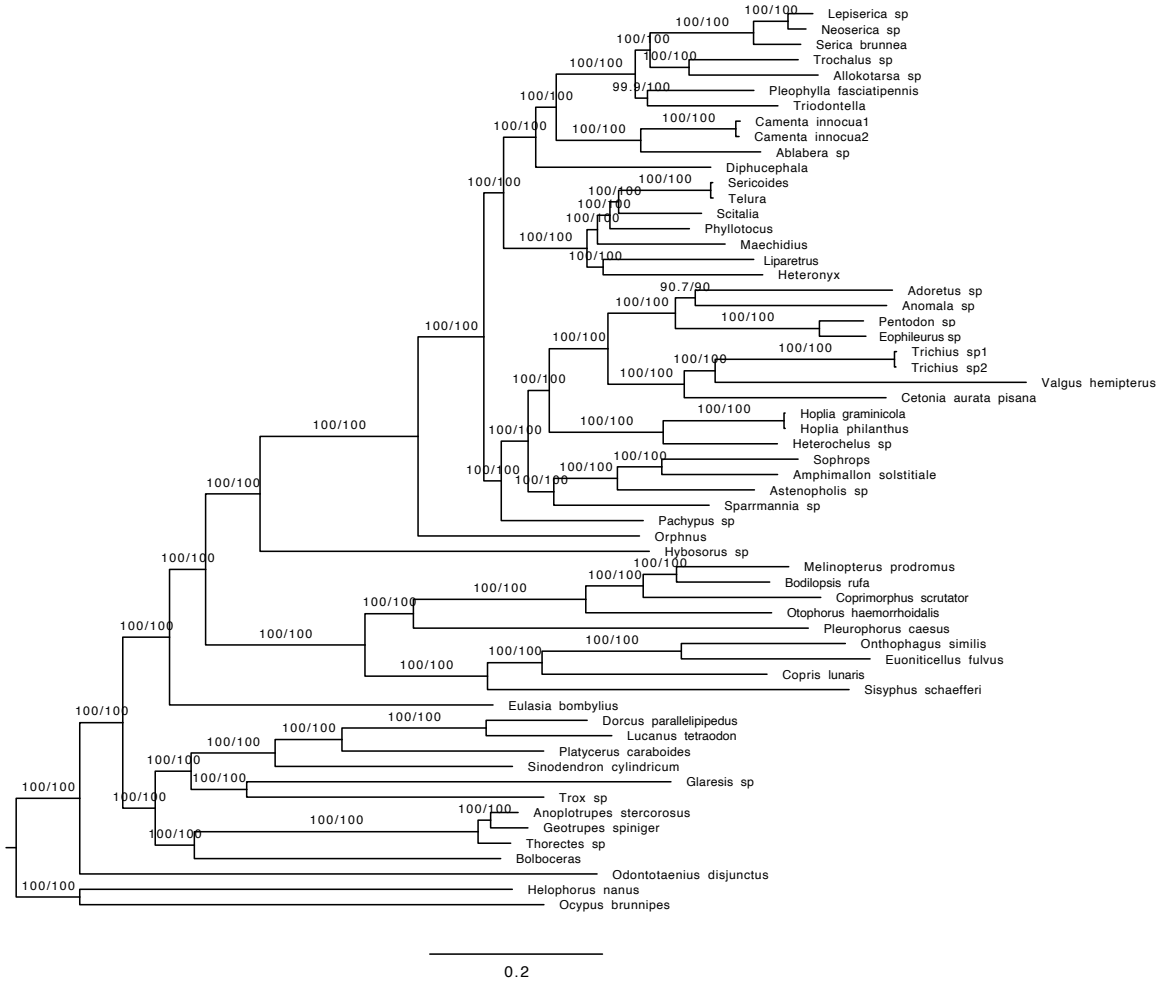
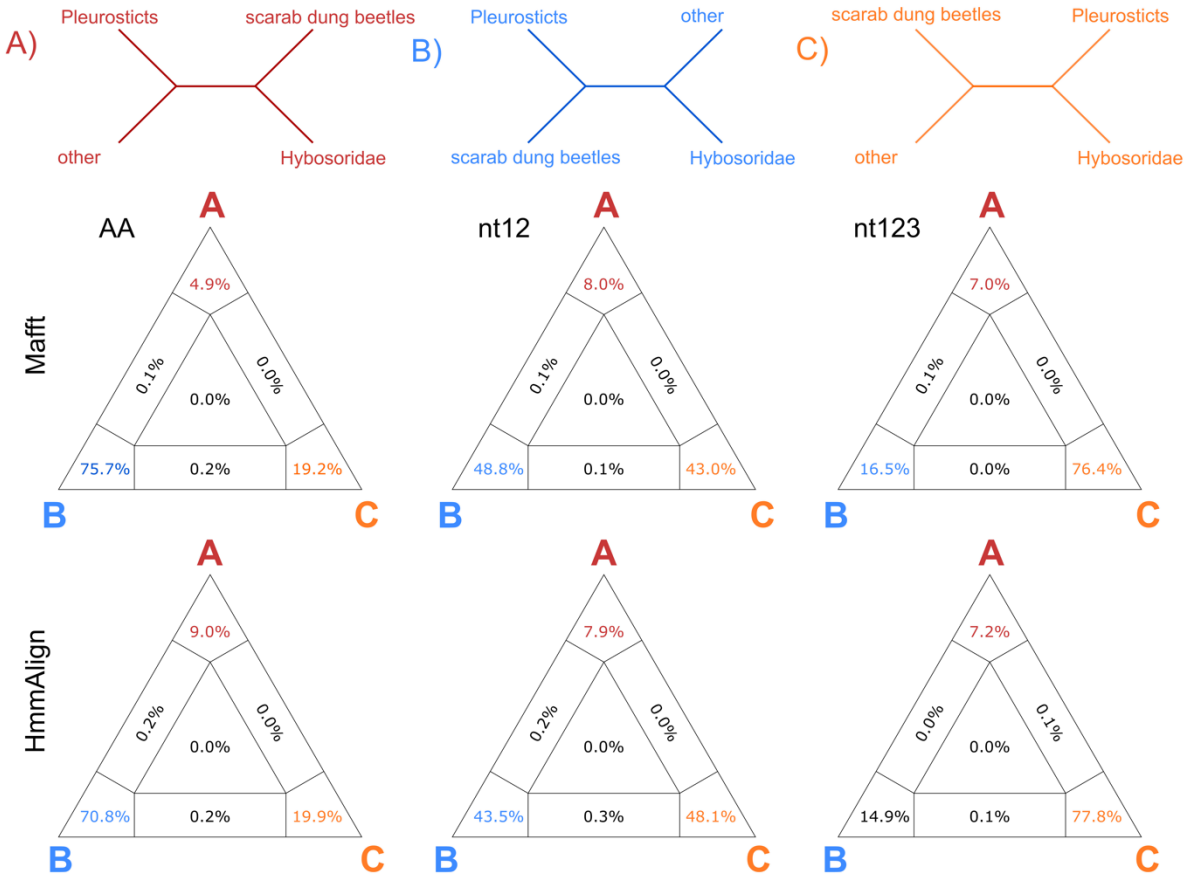
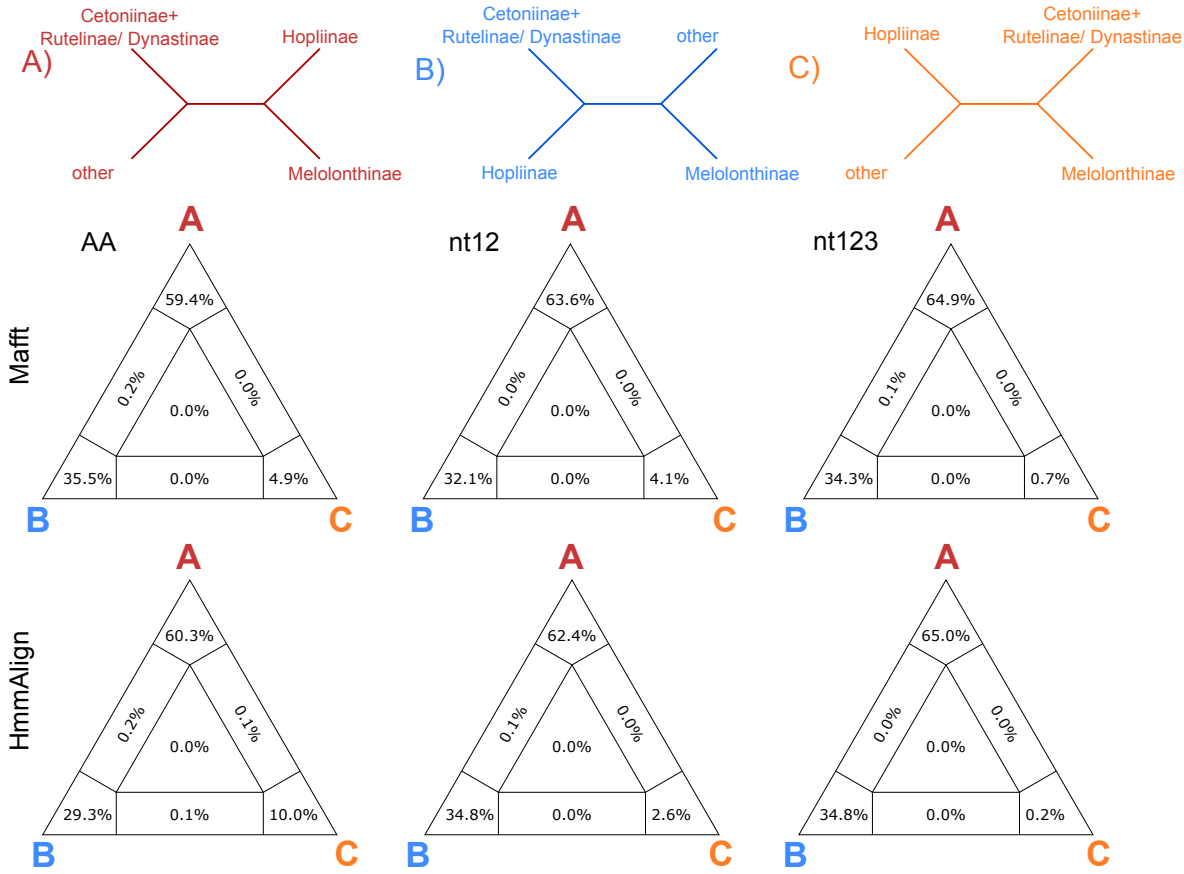


Figure S60. Phylogenetic tree obtained with the concatenated data and nucleotide data containing all three base pairs (nt123) (HmmAlign alignment; dataset with 70% complete data; Endopterygota orthologs).



Suppl. Figure 61. Quartet comparison of Endopterygota - single copy orthologs based on alignments with Hmmalign and Mafft, both based on the AA, nt12, and nt123 data set.



Suppl. Figure 62. Quartet comparison of Endopterygota - single copy orthologs based on alignments with Hmmalign and Mafft, both based on the AA, nt12, and nt123 data set, regarding the monophyly of Melolonthinae + Hopliinae.

Table S1. Overview on samples included in this study, referring to taxon name, current systematic placement, collection site, voucher number, accession numbers of sequence data repositories. The table also includes information on the number of recovered genes (Endopterygota USCOs vs. Coleoptera-specific genes), alignment completeness [in %] (Endopterygota USCOs, hmmalign vs. MAFFT vs. Coleoptera-specific genes, hmmalign vs. MAFFT).

Species	Family	Subfamily	Tribe	Locality	Voucher no	Accession number (Raw reads)	Accessions Transcriptome assemblies	Reference	# recovered genes (Endo USCOs)	# recovered genes (Coleo USCOs)	alignment completeness [%] (Endo USCOs, hmmalign)	alignment completeness [%] (Endo USCOs, MAFFT)	alignment completeness [%] (Coleo USCOs, hmmalign)	alignment completeness [%] (Coleo USCOs, MAFFT)
Neoserica sp	Scarabaeidae	Sericinae	Sericini	RSA: SA2013-16; South Africa/ Northern Cape: River Destiny Lodge (near Colesberg), 30°35'16,7"S, 25°25'11,7"E, 1207m, 22.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat07	SRR23446704	GKKO00000000	-	1978	3898	72.63	87.72	70.05	85.2
Lepiserica sp	Scarabaeidae	Sericinae	Sericini	RSA: SA2013-36; South Africa/ Mpumalanga: Graskop env., 10 km before Pilgrim's Rest; "Zur Alten Mine" Lodge, 24°55'46,6"S, 30°48'29,9"E, 1525m, 13.xii.2013, at light, open meadow, lg. Ahrens, Eberle, Fabrizi	RNA Lat18	SRR23446703	GKKQ00000000	-	1954	3943	69.06	85.43	67.83	84.55
Serica brunnea	Scarabaeidae	Sericinae	Sericini	Germany, Bonn, 2015	RNA Lat33	SRR23446692	GKJD00000000	-	1962	3836	65.58	86.61	63.09	84.39
Allokotarsa sp	Scarabaeidae	Sericinae	Sericini	RSA: SA2013-16; South Africa/ Northern Cape: River Destiny Lodge (near Colesberg), 30°35'16,7"S, 25°25'11,7"E, 1207m, 22.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat05	SRR23446681	GKKN00000000	-	1963	3926	64.73	84.78	62.71	82.8
Trochalus sp	Scarabaeidae	Sericinae	Sericini	RSA: SA2013-36; South Africa/ Mpumalanga: Graskop env., 10 km before Pilgrim's Rest; "Zur Alten Mine" Lodge, 24°55'46,6"S, 30°48'29,9"E, 1525m, 13.xii.2013, at light, open meadow, lg. Ahrens, Eberle, Fabrizi	RNA Lat17	SRR2083734	GKJW00000000	McKenna et al. (2019)	1969	3901	60.51	80	57.73	78.9
Triodontella sp	Scarabaeidae	Sericinae	Sericini	Italy: Sardinia, Cagliari 01.XI.2020 lg. D. Cillo	ZFMK-DA-RNAL62	SRR23446670	GKKL00000000	-	1974	3912	66.26	85.4	64.71	84.1
Pleophylla fasciatipennis	Scarabaeidae	Sericinae	Sericini	RSA: SA2013-19; South Africa/Western Cape: Goukamma Nature Reserve; Dunes near Blacktail Rondavell, 34°04'02,9"S, 22°56'50,8"E, 10m, 24.-26.xi.2013, daytime, lg. Ahrens, Eberle, Fabrizi	RNA Lat12	SRR23446669	GKJA00000000	-	1965	3845	67.05	85.89	64.39	83.4
Camenta innocua 1	Scarabaeidae	Sericinae	Ablaberini	RSA: SA2013-19; South Africa/Western Cape: Goukamma Nature Reserve; Dunes near Blacktail Rondavell, 34°04'02,9"S, 22°56'50,8"E, 10m, 24.-26.xi.2013, daytime, lg. Ahrens, Eberle, Fabrizi	RNA Lat11a	SRR2083641	GKJK00000000	McKenna et al. (2019)*	1791	3519	35.72	51.41	32.35	49.4
Camenta innocua 2	Scarabaeidae	Sericinae	Ablaberini	RSA: SA2013-19; South Africa/Western Cape: Goukamma Nature Reserve; Dunes near Blacktail Rondavell, 34°04'02,9"S, 22°56'50,8"E, 10m, 24.-26.xi.2013, daytime, lg. Ahrens, Eberle, Fabrizi	RNA Lat11b	SRR23446668	GKJB00000000	-	1953	3919	68.28	86	66.41	84.24
Ablabera sp	Scarabaeidae	Sericinae	Ablaberini	RSA: SA2013-19; South Africa/Western Cape: Goukamma Nature Reserve; Dunes near Blacktail Rondavell, 34°04'02,9"S, 22°56'50,8"E, 10m, 24.-26.xi.2013, daytime, lg. Ahrens, Eberle, Fabrizi	RNA Lat10	SRR23446667	GKKP00000000	-	1935	3875	64.46	84.4	62.91	83.65
Diphucephala sp	Scarabaeidae	Sericinae	Diphucephalini	AUSTRALIA: New South Wales, Jilliby Rd. 5 km NNW of Alison, -33.235689 151.378419; 20m, 18.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-16	ZFMK-DA-RNAL50	SRR23446666	GKKE00000000	-	1954	3839	63.18	85.32	60.91	82.71
Sericoides sp	Scarabaeidae	Liparetrinae	Sericoidiini	AUSTRALIA: New South Wales, picknick area @ Glendon Rd. 2 km E of Singleton, -32.560333, 151.197467; 40m 19.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-23	ZFMK-DA-RNAL60	SRR23446665	GKKK00000000	-	1897	3693	57.96	81.05	54.31	77.71
Telura sp	Scarabaeidae	Liparetrinae	Scitaliini	AUSTRALIA: New South Wales, picknick area @ Glendon Rd. 2 km E of Singleton, -32.560333, 151.197467; 40m 19.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-23	ZFMK-DA-RNAL59	SRR23446702	GKKJ00000000	-	1911	3659	56.76	79.15	52.66	75.86
Scitalia sp	Scarabaeidae	Liparetrinae	Scitaliini	AUSTRALIA: New South Wales, Wollombi valley N of Sydney, Cedar Creek Rd. and Wollombi Rd. junction; -32.874714, 151.203703; 110m; 18.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-18	ZFMK-DA-RNAL53	SRR23446701	GKKH00000000	-	1915	3794	59.48	82.48	57.71	81.1
Phyllotocus sp	Scarabaeidae	Liparetrinae	Phyllotocini	AUSTRALIA: New South Wales, Wollombi valley N of Sydney, Cedar Creek Rd. and Wollombi Rd. junction; -32.874714, 151.203703; 110m; 18.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-18	ZFMK-DA-RNAL51	SRR23446700	GKKF00000000	-	1952	3940	64.69	85.86	63.99	84.82
Maechidius sp	Scarabaeidae	Liparetrinae	Maechidiini	AUSTRALIA: New South Wales, picknick area @ Glendon Rd. 2 km E of Singleton, -32.560333, 151.197467; 40m 19.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-23	ZFMK-DA-RNAL57	SRR23446699	GKKI00000000	-	1831	3547	49.41	70.16	45.27	66.81
Liparetrus sp	Scarabaeidae	Liparetrinae	Liparetrini	AUSTRALIA: Victoria, Cann River @ Can River caravan park, -37.567642, 149.146703; 100m; 28.xi.2019, Fikáček, Seidel & Sýkora lgt. AU-2019-53	ZFMK-DA-RNAL47	SRR23446698	GKKD00000000	-	1949	3842	60.62	83.6	57.96	81.25

Heteronyx sp	Scarabaeidae	Liparetrinae	Heteronycini	AUSTRALIA: New South Wales, Wollombi valley N of Sydney, Cedar Creek Rd. and Wollombi Rd. junction; -32.874714, 151.203703; 110m; 18.xi.2019; Fikáček, Seidel & Sýkora lgt. AU-2019-18	ZFMK-DA-RNAL52	SRR23446697	GKKG00000000	-	1948	3836	60.81	82.36	58.02	80.89
Eophileurus sp	Scarabaeidae	Dynastinae	Phileurini	Sri Lanka: 22.X.19, Hiyare	ZFMK-DA-RNAL45	SRR23446696	GKKC00000000	-	1929	3897	65.13	84.93	64.46	82.55
Pentodon sp	Scarabaeidae	Dynastinae	Pentodini	RSA: SA2013-18; South Africa/ Western Cape: Koka Tsara Bush Camp, near Karoo Nat. Park; 32°15'20,3"S, 22°34'25,1"E, 980m, 23.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat08	SRR2083712	GKJU00000000	McKenna et al. (2019)	1831	3585	52.26	73.86	49.7	71.45
Adoretus sp	Scarabaeidae	Rutelinae	Adoretini	RSA: SA2013-16; South Africa/ Northern Cape: River Destiny Lodge (near Colesberg), 30°35'16,7"S, 25°25'11,7"E, 1207m, 22.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat03	SRR2083620	GKIO00000000	McKenna et al. (2019)*	1950	3924	62.8	83.76	62	82.51
Anomala sp	Scarabaeidae	Rutelinae	Anomalini	RSA: SA2013-16; South Africa/ Northern Cape: River Destiny Lodge (near Colesberg), 30°35'16,7"S, 25°25'11,7"E, 1207m, 22.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat04	SRR2083625	GKJH00000000	McKenna et al. (2019)	1966	3902	66.19	85.64	65.7	83.94
Trichius sp1	Scarabaeidae	Cetoniinae	Trichiini	Germany: Bonn, Duisdorf, Messdorfer Feld, Helnholtzstrasse, 25.5.2018, leg. Ahrens & Fabrizi	RNA Lat31	SRR23446695	GKKR00000000	-	1976	3930	71.38	87.2	69.74	84.04
Trichius sp2	Scarabaeidae	Cetoniinae	Trichiini	Germany, Rheinland Pfalz, Bausenberg, 12.6.2015	RNA Lat36	SRR23446694	GKKT00000000	-	1977	3957	72.47	88.01	71.08	85.51
Valgus hemipterus	Scarabaeidae	Cetoniinae	Valgini	Germany: Bonn, Duisdorf, Messdorfer Feld, Helnholtzstrasse, 25.5.2019, leg. Ahrens & Fabrizi	ZFMK-DA-RNAL37	SRR23446693	GKIZ00000000	-	1966	3935	69.1	85.74	66.53	83.28
Cetonia aurata pisana	Scarabaeidae	Cetoniinae	Cetoniini	Italy: Lazio, Lariano, 7.11.2013, lg. Ahrens	RNA Lat01	SRR2083643	GKJL00000000	McKenna et al. (2019)	1886	3652	48.85	69.14	45.33	66.07
Hoplia graminicola	Scarabaeidae	Melolonthinae	Hopliini	Germany: Bonn, Duisdorf, Messdorfer Feld, Helnholtzstrasse, 25.5.2018, leg. Ahrens & Fabrizi	RNA Lat30	SRR23446691	GKJC00000000	-	1943	3855	68.23	84.67	67.63	83.3
Hoplia philanthus	Scarabaeidae	Melolonthinae	Hopliini	Germany, Bonn, 2015	RNA Lat34	SRR23446690	GKJE00000000	-	1918	3735	62.26	82.45	59.69	80.6
Heterochelus sp	Scarabaeidae	Melolonthinae	Hopliini	RSA: SA2013-18; South Africa/ Western Cape: Koka Tsara Bush Camp, near Karoo Nat. Park; 32°15'20,3"S, 22°34'25,1"E, 980m, 23.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat09	SRR2083671	GKJP00000000	McKenna et al. (2019)	1923	3819	55.95	76.91	53.86	74.92
Sophrops sp	Scarabaeidae	Melolonthinae	Rhizotrogini	Sri Lanka: Kandy District, Deenston, Knuckles South, 7,35771944N, 80,8500611E, 16-X-2019, Light sheet	ZFMK-DA-RNAL43	SRR23446689	GKKB00000000	-	1943	3869	64.37	85.25	62.08	82.78
Amphimallon solstitiale	Scarabaeidae	Melolonthinae	Rhizotrogini	Germany: Bonn, Duisdorf, Messdorfer Feld, Helnholtzstrasse, 2.6.2021, leg. Ahrens	ZFMK-DA-RNAL66	SRR23446688	GKIQ00000000	-	1926	3730	63.94	83.32	59.56	79.27
Astenopholis sp	Scarabaeidae	Melolonthinae	Leucophilini	RSA: SA2013-36; South Africa/ Mpumalanga: Graskop env., 10 km before Pilgrim's Rest; "Zur Alten Mine" Lodge, 24°55'46,6"S, 30°48'29,9"E, 1525m, 13.xii.2013, at light, open meadow, lg. Ahrens, Eberle, Fabrizi	RNA Lat19	SRR2083634	GKJJ00000000	McKenna et al. (2019)	1911	3753	57.43	77.13	55.35	75.99
Sparrmannia sp	Scarabaeidae	Melolonthinae	Sparrmanniini	RSA: SA2013-16; South Africa/ Northern Cape: River Destiny Lodge (near Colesberg), 30°35'16,7"S, 25°25'11,7"E, 1207m, 22.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat02	SRR23446687	GKKM00000000	-	1950	3935	67.68	85.46	67.79	84.85
Pachypus sp	Scarabaeidae	Pachypodinae	Pachypodini	France, Corsica: Camping Villata (10 km N Porto Vecchio) (daytime), 41°39.565'N, 009°22.366'E; 21.-27.vi.2014; Ahrens & Fabrizi leg.	RNA Lat32, DA4284	SRR23446686	GKKS00000000	-	1932	3828	66.46	84.77	65.04	82.66
Orphnus sp	Scarabaeidae	Orphninae	-	Sri Lanka: Matale District, Dambulla, NIFS Arboretum, 7,86011766N, 80,67441844E, 11-X-2019, Light sheet	ZFMK-DA-RNAL40	SRR23446685	GKJZ00000000	-	1915	3845	61.58	83.68	59.74	81.76
Bodilopsis rufa	Scarabaeidae	Aphodiinae	Aphodiini	Lazio: Monte Lupone, near Rocca Massima, 7.ix.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL73	SRR23446684	GKIX00000000	-	1965	3901	69.77	85.78	64.61	82.78
Melinopterus prodromus	Scarabaeidae	Aphodiinae	Aphodiini	Germany: Kottenforst, 28.3.2014, leg. Ahrens	RNA Lat28	SRR2083628	GKJI00000000	McKenna et al. (2019)	1964	3854	60.79	81.18	56.53	79.66
Coprimorphus scrutator	Scarabaeidae	Aphodiinae	Aphodiini	Lazio: Monte Lupone, near Rocca Massima, 22.viii.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL71	SRR23446683	GKIV00000000	-	1967	3890	65.86	85.42	60.88	81.19
Otophorus haemorhoidalis	Scarabaeidae	Aphodiinae	Aphodiini	Lazio: Monte Lupone, near Rocca Massima, 22.viii.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL72	SRR23446682	GKIW00000000	-	1983	3943	66.6	84.84	62.16	82.23

Pleurophorus caesus	Scarabaeidae	Aphodiinae	Psammodiini	Italy: Velletri, 1.6.2021 leg. Ahrens	ZFMK-DA-RNAL63	SRR23446680	GKIN00000000	-	1955	3871	63.36	82.18	56.05	78
Onthophagus similis	Scarabaeidae	Scarabaeinae	Onthophagini	Germany: Kottenforst, 28.3.2014, leg. Ahrens	RNA Lat27	SRR2083707	GKJT00000000	-	1969	3903	61.28	80.02	60.32	79.04
Euoniticellus fulvus	Scarabaeidae	Scarabaeinae	Euoniticellini	Lazio: Monte Lupone, near Rocca Massima, 22.viii.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL69	SRR23446679	GKIT00000000	-	1956	3895	66.5	84.52	62.47	82.35
Copris lunaris	Scarabaeidae	Scarabaeinae	Coprini	Lazio: Monte Lupone, near Rocca Massima, 22.viii.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL70	SRR23446678	GKIU00000000	-	1928	3684	60.81	80.64	53.54	75.42
Sisyphus schaefferi	Scarabaeidae	Scarabaeinae	Sisiphynini	Lazio: Monte Lupone, near Rocca Massima, 22.viii.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL68	SRR23446677	GKIS00000000	-	1931	3773	61.25	79.81	52.71	74.64
Hybosorus sp	Hybosoridae	Hybosorinae	-	RSA: Pretoria; Rietfontein, 15/12/2013	RNA Lat26	SRR2083672	GKJQ00000000	McKenna et al. (2019)	1942	3770	64.42	84.46	59.8	81.33
Eulasia bombylius	Glaphyridae	Glaphyrinae		TUNISIA, Tunis, Carthage, 1.IV.2014. Leg.G.Sabatinelli	-	SRR2083662	GKJM00000000	McKenna et al. (2019)	1947	3843	64.44	83.14	60.4	80.6
Anoplotrupes stercorosus	Geotrupidae	Geotrupinae	Geotrupini	Germany: Kottenforst, 28.3.2014, leg. Ahrens	RNA Lat29	SRR2083624	GKJG00000000	McKenna et al. (2019)	1994	3985	71.26	87.57	68.41	85.84
Geotrupes spiniger	Geotrupidae	Geotrupinae	Geotrupini	Lazio: Monte Lupone, near Rocca Massima, 7.ix.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL74	SRR23446676	GKIY00000000	-	1953	3832	66.75	84.53	62.82	82.07
Thorectes sp	Geotrupidae	Geotrupinae	Geotrupini	Lazio: Monte Lupone, near Rocca Massima, 7.ix.2021 leg. Ahrens & Pacheco	ZFMK-DA-RNAL75	SRR23446675	GKJY00000000	-	1967	3870	68.33	85.28	62.74	81.42
Bolboceras sp	Bolboceratidae			Sri Lanka: Matale District, Dambulla, NIFS Arboretum, 7,86011766N, 80,67441844E, 11-X-2019, Light sheet	ZFMK-DA-RNAL42	SRR23446674	GKKA00000000	-	1920	3827	56.74	78.76	53.53	76.79
Odontotaenius disjunctus	Passalidae	Passalinae	Proculini	-	-	SRR2083704	GKJS00000000	McKenna et al. (2019)	1983	3931	55.76	77.31	50.06	74.29
Dorcus parallelipedus	Lucanidae	Lucaninae	Dorcini	Germany: Bonn, Duisdorf, Messdorfer Feld, Helnholtzstrasse, 2.6.2021, leg. Ahrens	ZFMK-DA-RNAL65	SRR23446673	GKIP00000000	-	1929	3780	62.14	80.35	56.92	76.09
Lucanus tetraodon	Lucanidae	Lucaninae	Lucanini	Italy: Lariano, 9.8.2021 leg. Ahrens	ZFMK-DA-RNAL67	SRR23446672	GKIR00000000	-	1890	3607	56.15	74.78	49.22	69.98
Platycerus caraboides	Lucanidae	Platycerinae		Germany: Röttgen, Kottenforst, 1.6.2015, leg. Ahrens	RNA Lat35	SRR23446671	GKJF00000000	-	1990	3942	71.73	87.63	68.39	83.37
Sinodendron cylindricum	Lucanidae	Sindesinae	Sinodendrini	-	-	SRR2083727	GKJV00000000	McKenna et al. (2019)	1923	3727	62.07	81.41	58.28	78.24
Trox sp	Trogidae	Troginae		RSA: SA2013-16; South Africa/ Northern Cape: River Destiny Lodge (near Colesberg), 30°35'16,7"S, 25°25'11,7"E, 1207m, 22.xi.2013, at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat06	SRR2083735	GKJX00000000	McKenna et al. (2019)	1969	3901	63.91	85.22	60.47	81.67
Glaresis sp	Glaresidae	-		RSA: SA2013-31; South Africa/ Western Cape: Vrolijkheid Nature Reserve, 33°55'06,5'S, 19°52'31,9"E, 182m, 8-9.xii.2013 at light, lg. Ahrens, Eberle, Fabrizi	RNA Lat14	SRR2083664	GKJN00000000	McKenna et al. (2019)	1958	3841	63.14	83.34	57.36	79.7
Helophorus nanus	Hydrophilidae	Helophorinae		-	-	SRR2083668	GKJO00000000	McKenna et al. (2019)	1995	3976	60.26	81.66	52.19	79.18
Ocypus brunnipes	Staphylinidae	Staplhylininae		-	-	SRR2083703	GKJR00000000	McKenna et al. (2019)	1936	3770	57.57	77.87	49.36	74.88

Table S2. Likelihood tests regarding the monophyly of different groupings using the Endopterygota USCO dataset, with constraint trees using a variety of resampling tests in IQ-TREE using the RELL approximation including bootstrap proportion (BP), Kishino-Hasegawa test (KH), Shimodaira-Hasegawa test (SH), expected likelihood weights (ELW), and the approximately unbiased (AU) test.

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	c-ELW	p-AU
hmmalign NT12							
1: Dung scarabs + Pleurosticti	-13896174	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-13896819	645.24	0	0	0	3.11E-134	0.00029
3: Hybosoridae + Dung scarabs	-13896694	519.73	0	0	0	3.74E-61	2.72E-05
hmmalign NT123							
1: Dung scarabs + Pleurosticti	-40490187	1800.4	0	0	0	0	4.53E-06
2: Hybosoridae + Pleurosticti	-40488386	0	1	1	1	1	1
3: Hybosoridae + Dung scarabs	-40490262	1875.5	0	0	0	0	1.55E-59
hmmalign AA							
1: Dung scarabs + Pleurosticti	-11819613	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-11820703	1090.2	0	0	0	9,97e-321	2.94E-06
3: Hybosoridae + Dung scarabs	-11820134	520.84	0	0	0	5.54E-51	1.97E-72
MAFFT NT12							
1: Dung scarabs + Pleurosticti	-12558587	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-12559281	693.31	0	0	0	2.08E-162	1.29E-06
3: Hybosoridae + Dung scarabs	-12559109	521.89	0	0	0	1.21E-84	4.85E-06
MAFFT NT123							
1: Dung scarabs + Pleurosticti	-36782425	1721.7	0	0	0	0	1.31E-52
2: Hybosoridae + Pleurosticti	-36780703	0	1	1	1	1	1
3: Hybosoridae + Dung scarabs	-36782464	1760.9	0	0	0	0	4.16E-06
MAFFT AA							
1: Dung scarabs + Pleurosticti	-10023091	0	1	1	1	1	1
2: Hybosoridae + Pleurosticti	-10024195	1104.1	0	0	0	0	3.72E-84
3: Hybosoridae + Dung scarabs	-10023664	572.39	0	0	0	1.79E-81	0.00017

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	c-ELW	p-AU
hmmalign NT12							
1: Adoretus + Anomala	-13896166	0	0.999	0.999	1	0.999	0.999
2: Adoretus + Dynastinae	-13896412	246.1	0.001	0.0007	0.0008	0.00106	0.00122
3: Anomala + Dynastinae	-13896558	391.65	0	0	0	3.93E-34	3.22E-05
hmmalign NT123							
1: Adoretus + Anomala	-40488384	0	0.989	0.989	1	0.989	0.99
2: Adoretus + Dynastinae	-40488602	217.46	0.0106	0.0115	0.017	0.0106	0.0106
3: Anomala + Dynastinae	-40488773	388.8	0	0	0	3.37E-19	3.80E-07
hmmalign AA							
1: Adoretus + Anomala	-11819613	0	0.646	0.641	1	0.646	0.638
2: Adoretus + Dynastinae	-11819645	31.7	0.354	0.359	0.501	0.354	0.362
3: Anomala + Dynastinae	-11819998	384.88	0	0	0	1.07E-68	7.50E-47

MAFFT NT12							
1: Adoretus + Anomala	-12558590	0	0.836	0.892	1	0.836	0.917
2: Adoretus + Dynastinae	-12558699	109.33	0.0747	0.0915	0.15	0.0749	0.111
3: Anomala + Dynastinae	-12558691	101.52	0.0891	0.108	0.179	0.089	0.132
MAFFT NT123							
1: Adoretus + Anomala	-36780719	0	0.63	0.699	1	0.63	0.754
2: Adoretus + Dynastinae	-36780770	50.741	0.246	0.301	0.44	0.246	0.348
3: Anomala + Dynastinae	-36780805	85.785	0.124	0.189	0.293	0.124	0.219
MAFFT AA							
1: Adoretus + Anomala	-10023128	37.129	0.328	0.338	0.477	0.329	0.336
2: Adoretus + Dynastinae	-10023091	0	0.67	0.662	1	0.669	0.687
3: Anomala + Dynastinae	-10023296	204.98	0.0019	0.0078	0.0131	0.00188	0.00412

Tree	logL	deltaL	bp-RELL	p-KH	p-SH	c-ELW	p-AU
hmmalign NT12							
1: Hopliini + Dynastinae etc.	-13896163	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-13897228	1064.9	0	0	0	9,88e-324	0.00032
3: Melolonthinae + Hopliini	-13896914	751.4	0	0	0	4.33E-189	3.21E-05
hmmalign NT123							
1: Hopliini + Dynastinae etc.	-40488268	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-40490887	2618.5	0	0	0	0	5.34E-05
3: Melolonthinae + Hopliini	-40490321	2052.5	0	0	0	0	2.04E-48
hmmalign AA							
1: Hopliini + Dynastinae etc.	-11819612	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-11820457	844.27	0	0	0	1.40E-213	1.57E-70
3: Melolonthinae + Hopliini	-11820241	628.59	0	0	0	7.87E-98	1.71E-37
MAFFT NT12							
1: Hopliini + Dynastinae etc.	-12558590	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-12559724	1133.6	0	0	0	0	2.87E-07
3: Melolonthinae + Hopliini	-12559436	845.57	0	0	0	1.33E-231	2.56E-57
MAFFT NT123							
1: Hopliini + Dynastinae etc.	-36780491	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-36783179	2688.9	0	0	0	0	3.67E-05
3: Melolonthinae + Hopliini	-36782640	2149.8	0	0	0	0	2.82E-43
MAFFT AA							
1: Hopliini + Dynastinae etc.	-10023092	0	1	1	1	1	1
2: Melolonthinae + Dynastinae etc.	-10024095	1003.6	0	0	0	2.89E-300	3.51E-44
3: Melolonthinae + Hopliini	-10024017	925.64	0	0	0	1.60E-254	1.47E-41