Phylogenomics of Phengodidae (Coleoptera: Elateroidea): towards a natural classification of a bioluminescent and paedomorphic beetle lineage, with recognition of a new subfamily


ABSTRACT

Phengodidae (Coleoptera: Elateroidea), commonly known as glowworm beetles, are a small family of bioluminescent and paedomorphic beetles. There are few phylogenetic studies of Phengodidae, and these are mostly discordant, especially when comparing morphology-based and molecular-based phylogenetic hypotheses. Here, we used the anchored hybrid enrichment approach to undertake the first phylogenomic analysis of Phengodidae (≤358 loci and 39 taxa) and evaluate the higher-level classification of the group. In agreement with previous molecular studies, we recovered Phengodidae as sister to Rhagophthalmidae, and the Old World Cydistinae as sister to all New World Phengodidae. In contrast to previous hypotheses, both Phengodinae and Mastinocerinae were each recovered as monophyletic. Cenophengus was found to be sister to Mastinocerinae, in contrast to some previous hypotheses that placed it as sister to all New World Phengodidae. Considering its morphological divergence, we here establish Cenophenginae subfam. nov. Despite the largest and most comprehensive sampling of Phengodidae in any molecular-based study to date, we had only limited success in revealing the relationships among genera within the most species-rich subfamily, Mastinocerinae. Further studies should focus on the phylogeny and classification of this taxonomically neglected subfamily, on the phylogenetic placement of enigmatic Penicillophorinae, and on seeking morphological support for the main clades of Phengodidae.

Keywords: anchored hybrid enrichment; bioluminescence; glowworms; lampyroid clade; neoteny; paedomorphosis; phylogeny; railroad worms

INTRODUCTION

Phengodidae (Coleoptera: Elateroidea) are a small but morphologically diverse family of bioluminescent beetles (Fig. 1). The family is composed of ~300 described extant species and 43 genera, which are classified in three or four subfamilies, depending on the adopted classification (e.g. Kundrata et al. 2014, 2019, Zaragoza-Caballero and Pérez-Hernández 2014, Zaragoza-Caballero and Zurita-García 2015). Most of...
Phengodidae diversity is found in the New World, especially in the Neotropical realm (Constantin 2014, Zaragoza-Caballero and Pérez-Hernández 2014, Roza et al. 2017, Roza and Mermudes 2019, 2020, Vega-Badillo and Zaragoza-Caballero 2019, Vega-Badillo et al. 2020, 2021a, b, c, Roza 2023); however, the subfamily Cydistinae, containing two genera and seven species, are distributed in West Asia (from Turkey to Iran; Kundrata et al. 2019). The origin of Phengodidae dates back to the Mesozoic Era, and the only known fossil is known from the mid-Cretaceous amber of northern Myanmar (Roza et al. 2023).

Aside from the bioluminescence found in adults and larvae of these beetles, giving them their common names of glowworm beetles or railroad-worm beetles (lights on the body of some larvae and females look like the lights from train windows at night; Tiemann 1970, Costa and Zaragoza-Caballero 2010), Phengodidae exhibit extreme sexual dimorphism. This is associated with paedomorphosis, a type of heterochrony in which the juvenile traits of an organism are retained into adulthood (Cicero 1988, Kundrata and Bocák 2019, Rosa et al. 2020, Ferreira and Ivie 2022). Phengodidae adult males have all the features that define an adult beetle, including elytra and hind wings (e.g. Constantin 2014, Roza et al. 2017, Roza and Mermudes 2019). Known adult females, in contrast, are extremely paedomorphic and strongly resemble their immature forms (Tiemann 1967, 1970, Costa et al. 1999).

Although morphologically diverse, most adult male Phengodidae can be identified by their weakly sclerotized integument, antennae with 12 antennomeres, with antennomere III short and double rami on antennomeres IV–VI (except in Penicillophorinae, which have antennae moniliform, serrate or uniflabellate), shortened elytra, and eight visible abdominal segments separated by membranes (Fig. 1). The most distinct character of adult male Phengodidae is the long flagellum extending from the median lobe of the aedeagus (Costa and Zaragoza-Caballero 2010). This character, however, has not been found in Asian Cydistinae, whose representatives have the median lobe divided into long dorsal and ventral lobes (Kundrata et al. 2019). Larvae and adult females can be distinguished from other families by the combination of elongate and slightly flattened body, head not covered by pronotum, mandible with an internal channel, surface of the mandibular base simple or slightly enlarged, and frontal suture absent (Costa et al. 1988). Although family-level diagnostic characters of Phengodidae are relatively homogeneous, there is much morphological variation within the family (Zaragoza-Caballero and Pérez-Hernández 2014, Kundrata et al. 2019). The morphology of body parts such as the antennae, mandibles, maxillary and labial palpi, pronotum, elytra, hind wing, and tarsi can vary dramatically across different genera, but in some cases are superficially very similar in distantly related groups (A.S.R., personal observation).

There are few studies on the phylogenetic relationships of Phengodidae. Results of these were often contradictory, especially when comparing morphology-based and molecular-based phylogenetic hypotheses. The family was traditionally classified into three subfamilies: Phengodinae, Mastinocerinae, and Penicillophorinae (Costa and Zaragoza-Caballero 2010). Zaragoza-Caballero and Zurita-García (2015) carried out the first morphology-based phylogenetic analysis of the family, and they challenged that traditional subfamily classification of Phengodidae. They did not recover Phengodidae as monophyletic, but instead with Penicillophorinae, as a monophyletic group nested inside the Telegeusidae (currently Ometidae: Telegeusinae; Kundrata et al. 2014, Cai et al. 2022; Fig. 2A). They also recovered Mastinocerinae and Phengodinae as paraphyletic, and the enigmatic Cenophengus LeConte, 1881 as sister to all other phengodids (Fig. 2A). Cenophengus has long evaded morphological placement by sharing a long prosternum with the West Asian Cydistus Bourgeois, 1885 and by a combination of head and genital characters that defined easy placement among genera from the Americas. Quintino (2017), in her unpublished PhD dissertation, used a modified version of the dataset compiled by Zaragoza-Caballero and Zurita-García (2015) to test the monophyly of Mastinocerinae. Her conclusions, like those of Zaragoza-Caballero and Zurita-García (2015), did not recover the subfamilies as monophyletic (Fig. 2B).

Although some molecular-based phylogenetic studies have tested and corroborated the monophyly of Phengodidae (Kundrata et al. 2014, McKenna et al. 2015), they were focused on interfamly relationships of beetles and included few phengodids. The study by Kundrata et al. (2019) is the most up-to-date molecular-based phylogenetic hypothesis for Phengodidae. There, the authors tested the placement of the West Asian Cydistus, previously classified as Elateriformia incertae sedis. In their analyses of two nuclear (18S and 28S) and two mitochondrial (rnlS and cox1) genes, Cydistinae + Phengodinae were sister to Rhaeophthalmidae. Hence, the formerly unplaced Cydistinae were transferred into a newly circumscribed Phengodinae. Phengodinae were monophyletic, and Mastinocerinae were recovered as paraphyletic because Cenophengus, historically in Mastinocerinae, was sister to all other Phengodinae and Mastinocerinae (Fig. 2C). The main limitations of that study were the lack of representatives of Penicillophorinae, and that most of the ingroup genera (data mined from GenBank) were identified only as ‘Phengodidae gen.’, making it impossible to infer within-family relationships.

André Roza’s (2020) unpublished PhD dissertation is the most comprehensive phylogenetic analysis of the family to date. In his dissertation, Roza (2020) produced a morphology-based phylogeny of Phengodinae, in which he included representatives of all described genera of the family. His results, like those of Kundrata et al. (2019), recovered Cydistinae as sister to New World Phengodinae, and Cenophengus as sister to Phengodinae + Mastinocerinae. Additionally, Penicillophorinae were found nested inside Mastinocerinae (Fig. 2D).
To sum up, the phylogeny of Phengodidae remained largely unresolved and, as a result, the classification controversial. In particular, the monophyly and limits of subfamilies remained unclear owing to unresolved intergeneric phylogenetic relationships. Seeking better phylogenetic resolution for a stable and complete classification of the group, we produced genomic data for Phengodidae. We implemented a target capture method via anchored hybrid enrichment, which uses short RNA probes as baits to target single-copy orthologous nuclear loci and which has been applied successfully to beetles (Haddad et al. 2017, Martin et al. 2019, Brunke et al. 2021, Douglas et al. 2021, 2023). The resulting dataset is the largest and most comprehensive ever produced for Phengodidae. We used it to test the limits and monophyly of Cydistinae, Phengodinae, and Mastinocerinae, to test the placement of Cenophengus, and to provide robust hypotheses of the subfamily- and genus-level relationships of Phengodidae.

MATERIALS AND METHODS

Taxon sampling and identification of material

Our sampling efforts aimed to provide a comprehensive coverage of the Phengodidae (sensu Costa and Zaragoza-Caballero 2010, Kundrata et al. 2014, 2019), with seven other Elateroidea families as outgroups. Based on current hypotheses (e.g. Kundrata et al. 2014, Douglas et al. 2021, Cai et al. 2022), Phengodidae are well nested within the superfamily Elateroidea. In all of these studies, Phengodidae were recovered as sister to Rhagophthalmidae, and that clade as sister to Elateridae (Kundrata et al. 2014) or Lampyridae (Douglas et al. 2021, Cai et al. 2022). These studies were used to guide our taxon sampling for outgroups.

For Phengodidae, we sampled 25 of 43 described genera, representing three of four currently recognized subfamilies (Phengodinae, Cydistinae, and Mastinocerinae) (Supporting Information, Tables S1 and S2). Although representatives of the fourth subfamily, Penicillophorinae, were available for study, we failed to recover sequence data from them. For outgroups, we sampled Rhagophthalmidae (three specimens), Elateridae (two), Lampyridae (two), Lycidae (two), Eucnemidae (one), Throscidae (one), and Artematopodidae (one). Our samples were dry-pinned and ethanol-preserved specimens (Supporting Information, Table S1). Although most of the data were generated anew in this study, we also incorporated public sequence data for six taxa generated using the same probe set from the study by Douglas et al. (2021) and recovered targets from two transcriptomes (Wang et al. 2017). For more information, see Supporting Information, Table S2.

Specimens were identified by A.S.R., M.A.I., S.Z.-C., R.K., V.S.F., and V.V.-B. using current literature on Phengodidae, i.e. Zaragoza-Caballero and Pérez-Hernández (2014), Roza et al. (2017, 2018), Kundrata et al. (2019), Roza and Mermudes (2019, 2020), Roza (2023), and Vega-Badillo et al. (2020, 2021a, c), and by direct comparison with museum specimens, including types. Voucher depositories are given in Supporting Information, Table S1.

Sample preparation and DNA extraction

Prior to DNA extraction, most specimens were severed at the connection between the prothorax and elytra, and they were used in their entirety for non-destructive immersion in proteinase K and lysis buffer. After DNA extraction, specimens were washed in 95% ethanol and mounted on card. For some specimens, only internal tissues were available for extraction, and the muscle remains were either discarded or dissolved by the proteinase K (see below). Owing to the soft-bodied nature and fragility of Phengodidae, some specimens were completely disarticulated during the extraction process. Specimens varied in age from 40 years (Neophengus sp., PD0129, collected in 1983) to <1 year from the date of collection to the time of DNA extraction (2021).

Extractions from ethanol-preserved or recently collected specimens (<5 years) were done using the DNeasy Blood and Tissue kit (Qiagen, Germantown, MD, USA). Pinned specimens with >5 years from the date of collection were extracted using a QIAamp DNA Micro Kit (Qiagen; standard protocol with carrier RNA added). For both kits, manufacturer’s protocols for tissue samples were followed, with a prolonged overnight (10–14 h) lysis period. For the ethanol-preserved samples, 4 µL of RNase A (100 mg/mL) was added after lysis, followed by a 2 min incubation interval at room temperature. For both kits, the elution buffer was preheated to ~60°C, and DNA was eluted in buffer EB after a 10 min incubation. This step was repeated twice, with 30 µL each time, for a final elution volume of 60 µL. All extractions were done in the molecular laboratory at the Zoological Museum of the Natural History Museum of Denmark. DNA extractions are permanently deposited at ~80°C in the DNA and tissue collections of the Senckenberg Deutsches Entomologisches Institut, Müncheberg, Germany.

Anchored hybrid enrichment: probe choice

For the anchored hybrid enrichment method, we used the probes ‘ElaterBaits’, developed by Douglas et al. (2021). The probes, aiming at 2260 target regions for Elateroidea, were not developed specifically for Phengodidae, but showed an adequate level of data recovery (753 loci) for one specimen of the family sampled by Douglas et al. (2021). The probe file, a FASTA file consisting of all the target regions, was submitted to Arbor Biosciences (Ann Arbor, MI, USA) for the production of a myBaits Custom probe kit.

Library preparation, hybridization, and sequencing

The DNA concentration of samples was quantified using a Qubit 3.0 fluorometer (Invitrogen, Burlington, ON, Canada), and the DNA fragment size range was measured by 2% agarose gel electrophoresis. Genomic libraries were prepared using an NEBNext DNA Ultra II FS Kit for Illumina (New England BioLabs, Ipswich, MA, USA). DNA was first sheared enzymatically to target an average length of 200–300 bp using incubation times of 1–15 min depending on the starting fragment size. Adapters were diluted to 0.6 µM for DNA input <50 ng, and adapter ligated inserts were eluted in 33 µL of 0.1× Tris-EDTA. Libraries were dual-indexed using corresponding NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1; New England BioLabs) and PCR amplified for 5–13 cycles (default amplification protocol) depending on the amount of original input DNA in each sample (highest cycles for the lowest DNA concentrations). Post-PCR, indexed libraries were quantified with another Qubit assay, and fragment size was measured for representative samples from each PCR group using 2% agarose gel electrophoresis. Post-PCR, indexed libraries were first grouped according to the DNA concentration (as a subjective indicator of quality), then pooled at equal concentrations with a target of ~500 ng input to the hybridization reaction, provided there was sufficient total DNA per sample.

Pools composed of degraded samples (those with low genomic DNA concentration and/or short fragment length) contained much less total input (100–400 ng, 10–40 ng per sample). Pooled libraries were reduced to ~7 µL volumes for hybridization using an incubator with the temperature set at 40°C. Reduced, pooled libraries were hybridized with the DNA probes using a myBaits Hyb Capture Kit (Arbor Biosciences) according to the myBaits v. 4.01 protocol, with 20–25 h of hybridization at 65°C, followed by the KAPA HiFi on bead PCR method (arborbiosci.com). Purified, hybridized libraries were amplified with 14–22 cycles of PCR, depending on the concentration. Quality control, pooling, and normalization of samples, in addition to sequencing,
were done at RAPiD Genomics (Gainesville, FL, USA) on a NovaSeq SP with 2 × 250 bp paired-end reads. Demultiplexed, raw-read FASTq files were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject PRJNA1104084.

Read assembly and orthology assessment pipeline
Bioinformatics were performed using Denmark’s National Life Science Supercomputing Center, ComputeRer v.2.0 (https://www.computerome.dk). The pipeline followed that described by Douglas et al. (2021) and Brunke et al. (2021). A bioinformatics pipeline, heavily based upon elements of the PHYLUCE package (Faircloth 2016), was developed using Snakemake (Köster and Rahmann 2012) to input raw Illumina reads and output aligned target loci for various target enrichment projects. This pipeline is available at https://github.com/AAFC-BiCoE/snake-make-partial-genome-pipeline (accessed 16 February 2023). Raw reads were initially adapter-trimmed using BBDuk, a component of the BBMAP package (sourceforge.net/projects/bbmap/; Bushnell et al. 2017). Single reads were then assembled de novo using three different assemblers, i.e. ABySS (Jackman et al. 2017), SPAdes (Bankevich et al. 2013), and rNA SPAdes (Bushmanova et al. 2019). Additionally, reads were merged using BBMerge, then assembled via a second run of ABySS.

The output of each of the four assembly methods, plus the probe sequences, was input separately to PHYLUCE, where assemblies were matched to target loci with minimum 65% identity and minimum coverage, respectively. Assemblies matching multiple target loci were filtered out with PHYLUCE, as were target loci with probes matching multiple assemblies considered to be different. The results of these four assembly methods were compared, and the longest fragment of each target locus was retained. Brunke et al. (2021) found that using multiple assemblers drastically increased the number of recovered targets, in agreement with the results of Hedin et al. (2018).

Alignment, trimming, and manual inspection
Alignments, trimming, and manual inspection largely followed Douglas et al. (2021, 2023) and Brunke et al. (2021), with modifications. Alignment and internal trimming were performed using elements of the PHYLUCE pipeline under default settings (Faircloth 2016), unless stated otherwise (see below). Alignment of each locus was performed in MAFFT (Katoh et al. 2002), with edge trimming turned off. Internal trimming of ambiguously aligned regions was performed in GBlocks (Talavera and Castresana 2007) using a less stringent approach [parameters: b1 0.5, b2 0.85, b3 8, b4 10; where: b1 = minimum number of sequences for a conserved position (as a percentage); b2 = minimum number of sequences for a flank position (as a percentage); b3 = maximum number of contiguous nonconserved positions; and b4 = minimum length of a block]. Trimmed, single-locus alignments were inspected manually in GENEIOUS PRIME (v.2022.2.2) to find the reading frame with the lowest number of stop codons and for removal of taxa with empty alignments (an artefact of earlier GBlocks steps), removal of taxa with very short sequences (<30 bp) as a result of trimming, and other alignment artefacts. Non-coding, flanking regions were trimmed, and only gaps divisible by three were allowed. Dubious gaps in coding probe regions and downstream nucleotides affected by the frameshift were converted to ambiguous (N’s). Non-orthologous sequences and obvious contamination not already filtered by PHYLUCE were removed after manual inspection. When two widespread paralogues were identified in an alignment, the most common was retained to preserve as much data as possible. Scripts and alignments are available at FigShare (https://doi.org/10.6084/m9.figshare.25650891.v3).

Phylogenetic analyses
Phylogenetic analyses and other tests were performed on two versions of our multiple sequence alignments composed of loci with ≥45% or ≥65% of taxa present (‘completeness’), respectively. For this, we used the PHYLUCE script ‘phyluce_align_get_only_loci_with_min_taxa’ (Faircloth 2016). The multiple sequence alignments were then concatenated and exported in NEXUS format using AMAS (Borowiec 2016). The phylogenetic analyses were conducted under three tree-building methods (optimality criteria) as follows: (i) concatenated analyses performed using the Bayesian inference (BI) method (Rannala and Yang 1996, Yang and Rannala 1997) via MrBayes v.3.2.7.a software (Ronquist et al. 2012); (ii) concatenated analyses performed using the maximum likelihood (ML) method (Felsenstein 1973, 1981) via IQ-TREE2 software (Minh et al. 2020a); and (iii) a summary method of species tree inference from gene trees, modelled under the multispecies coalescent model (MScM; Pamilo and Nei 1988, Rannala and Yang 2003) via ASTRAL-Pro2 software (Zhang and Mirarab 2022), with input gene trees being re-constructed individually under ML via IQ-TREE2 software.

Phylogenetic analyses were performed at the nucleotide level (45CPP-NT = 45% completeness matrix; 65CPP-NT = 65% completeness matrix) with either partitioned (45CPP-NT and 65CPP-NT) or unpartitioned (45CPU-NT and 65CPU-NT) variants (Table 1; Supporting Information, Table S3). Nucleotide datasets were partitioned by position and locus. We determined the evolutionary models for unpartitioned, partitioned, and individual gene trees analyses (MFP + MERGE command of IQ-TREE2) via MODEL FINDER software (Kalyaanamoorthy et al. 2017).

To perform an a priori evaluation of the quality of datasets, we used AliGROOVE software (Kück et al. 2014), an improvement of the ALISCORE software (Misof and Misof 2009). This approach (Misof and Misof 2009) uses a Monte Carlo resampling method to generate a ‘sliding window’ to aid visualization of heterogeneous sequence composition and similarity to random noise of individual taxa or groups of taxa. By identifying nodes with likely erroneous phylogenetic signals, we aimed to identify potential ‘rogue’ terminals among the generated multiple sequence alignments used here for phylogenetic analyses. Despite the general low phylogenetic signal in the evaluated datasets, we found a relatively consistent pattern in most pairwise sequence comparisons and did not identify any rogues among our terminals (Supporting Information, Fig. S1).

For the support measures of internal branches in ML analyses, we calculated the ‘ultrafast’ bootstrap (UFBoot; Hoang et al. 2018), the Shimodaira–Hasegawa-like approximate likelihood ratio test (SH-aLRT; Guindon et al. 2010) and the Bayesianlike transformation of aLRT (aBayes; Anisimova et al. 2011). For Bayesian inference analyses, we calculated the posterior probability (PP; Yang and Rannala 1997) of the clades of the
Table 1. Statistical support for internal branches measured for each phylogenetic reconstruction based on nucleotide (NT) datasets with 45% completeness. Abbreviations: CPU, unpartitioned analysis; CPP, partitioned analysis. For maximum likelihood (ML) analyses: aBayes, Bayesian-like transformation of aLRT; gCF, gene concordance factor; sCF, site concordance factor; SH-aLRT, Shimodaira–Hasegawa-like approximate likelihood ratio test; UFBoot, ultrafast bootstrap. For Bayesian inference (BI) analyses: PP, posterior probabilities of the clades of the majority-rule consensus tree of the posterior distribution generated via the Markov chain Monte Carlo (MRC; Margush and McMorris 1981) of the posterior distribution generated via the Markov chain Monte Carlo (Ronquist et al. 2012). For coalescent ASTRAL analyses, supports were calculated as local posterior probabilities (Local PP; Zhang et al. 2020).

Nodes with support values above ~95% (for UFBoot and aBayes) and above ~80% (for SH-aLRT) were considered strongly supported and therefore discussed as strong hypotheses of evolutionary relationships in the inferred phylogeny (Guindon et al. 2010, Anisimova et al. 2011, Hoang et al. 2018). The dominant interpretation of the PP as a support value is that a value above ~.95 approximates a 95% probability that a clade is true. However, this interpretation simplifies the vast (and varied) bibliography comparing phylogenetic support measures (e.g. Hillis and Bull 1993, Berry and Gascuel 1996, Alfaro et al. 2003, Anisimova et al. 2011). Additionally, PP values are more susceptible to model misspecification than are non-parametric bootstrap values (e.g. Douady et al. 2003, Erixon et al. 2003). In view of this, we consider values between ~.85 and ~.95 as only ‘medium supported’ (for a more in-depth discussion on this subject, see Berry and Gascuel 1996, Anisimova et al. 2011).

We also calculated other support measures used for phylogenomic studies, which are metrics based on the genealogical concordance of the genomic dataset: the gene concordance factor (gCF; Minh et al. 2020b), which computes the proportion of individual gene trees that support each specific node, properly corrected for variable terminal coverage; and the improved version of the site concordance factor (sCF; Minh et al. 2020b, Mo et al. 2023), which computes the proportion of alignment sites that support each specific node, in an ML framework. All support measures were computed via IQ-TREE2 software, except the PP, which was calculated via MrBayes v.3.2.7a software, and the Local PP, calculated via ASTRAL-PRO2 software. The threshold level of sCF and gCF followed Douglas et al. (2023), and we considered nodes with sCF of 37–39 to have weak phylogenetic signal, sCF 40–49 to have moderate phylogenetic signal, and those with ≥50 to have strong phylogenetic signal. For nodes with sCF < 37, phylogenetic signal was considered to be more or less equivocal.

To assert the statistical significance (i.e. P-value with α = 0.05 as criterion for statistical significance) for phylogenetic hypotheses, we applied non-parametric tree likelihood-based topology tests (Goldman et al. 2000): the two-tailed Kishino–Hasegawa (KH) test (Kishino and Hasegawa 1989), the one-tailed Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa 1999), and Shimodaira’s approximately unbiased (AU) test (Shimodaira 2002). This is a popular approach (e.g. Buckley 2002, Aris-Brosou 2003, Plant 2006, Ferreira et al. 2023) also applied in phylogenomic studies (e.g. Cui et al. 2013, Arcila et al. 2017, Díaz-Escandón et al. 2022, Dietz et al. 2023), including anchored hybrid enrichment datasets (e.g. Betancur-R et al. 2018). The results of tree topology tests are given in Supporting Information, Table S4.

Four-cluster likelihood mapping

To investigate the statistical support of alternative hypotheses in our reconstructed tree, we applied the method of four-cluster quartet likelihood mapping analysis (Strimmer and von Haeseler...
1997). With this method, hypothesized groups are organized into quartets (four-terminal sets), representing a simplified topology of the relationships to be tested. This procedure outputs a two-dimensional simplex plot that displays the statistical support of competing hypotheses via the proportions of quartets recovered from each possible topology or even inconclusive relationships. Four-cluster quartet likelihood mapping analyses were performed using IQ-TREE2 software.

RESULTS

Dataset and target capture

After Phyluce quality control filtering, we successfully recovered 66 (PD0126) to 1703 (SRR4045941) loci from each sample or archived data source (Supporting Information, Tables S1 and S2). This provided data for Phengodidae (recognition 66–1097; average 470 loci, \( N = 27 \)), Rhagophthalmidae (recognition 457–1703; average 1028 loci, \( N = 3 \)), Elateridae (recognition 1293–1538; average 1415 loci, \( N = 2 \)), Lampyridae (recognition 1060–1355; average 1207 loci, \( N = 2 \)), Lycidae (recognition 147–343; average 245 loci, \( N = 2 \)), Throsicidae (629, \( N = 1 \)), Eucnemididae (865, \( N = 1 \)), and Artematopodidae (847, \( N = 1 \)). Eleven of the 27 Phengodidae samples (40.7% of the Phengodidae dataset) were pinned museum specimens. For these samples, the lowest recovery was of 66 loci (PD0126, collected in 1996) and the highest of 776 (PD0125, collected in 2018); our oldest sample (PD0129, collected in 1983) recovered 518 loci. We kept Microphengodes sp. (PD0126) in the final dataset despite its lower number of recovered loci (66) because it represents one of only four genera in Phengodinae. The average of the recovery for pinned Phengodinae samples was of 434 loci. The probe set performed best in silico, recovering 1355 and 1703, respectively, for Lampyris noctiluca (Linnaeus, 1758) (SRR12432455) and Rhagophthalmus sp. (SRR4045941). After manual processing of single-locus alignments, the 45% and 65% completeness nucleotide datasets contained 358 loci (35 112 bp) and 90 loci (9234 bp), respectively. The final datasets are available at FigShare (https://doi.org/10.6084/m9.figshare.25650891.v3). According to all non-parametric tree topology tests performed (bp-RELL, KH, SH, WKH, WSH, and AU tests), the analyses performed via coalescent ASTRAL analyses were statistically less reliable (all \( P \)-values approach .00). All ML and BI analyses presented similar performance (all \( P \)-values > .066). Among analyses performed with the 45% completeness dataset, the partitioned ML analysis (ML-45CPP-NT) presented the highest score (all \( P \)-values > .351). Among analyses performed with the 65% completeness dataset, the unpartitioned ML analysis (ML-65CPU-NT) presented the highest score (all \( P \)-values > .438). For more details, see Supporting Information, Table S4.

All analyses of the 45% completeness datasets recovered Phengodidae + Rhagophthalmidae as monophyletic, usually with robust statistical support (except for the coalescent ASTRAL analyses, which had generally lower support for most recovered clades). The Rhagophthalmidae were monophyletic and maximally supported in all analyses. The Phengodidae as currently understood (i.e. including Cydistinae) were monophyletic with high to maximal support in all performed ML and BI analyses, but paraphyletic with Cydistinae as sister to Rhagophthalmidae in the coalescent analysis (Table 1; Fig. 3). Phengodidae received only weak support from sCF in the partitioned NT analysis. Cydistinae (represented by both known genera, Cydistus and Microcystis Kudrata et al., 2019) et al., were monophyletic in all analyses, with maximum support in all ML and BI analyses and with strong support from sCF (Table 1). Cydistinae were recovered as sister to all remaining (i.e. New World) Phengodidae in all but the coalescent analysis. The New World Phengodidae were also recovered as monophyletic in all analyses, with maximal support in all ML and BI analyses. This clade received weak support from sCF but was more reliably supported in the 65% CPP-NT analyses. The Phengodidae (represented by Phenogodes Illiger, 1807, Zarhipis LeConte, 1881, and Microphenogodes Wittmer, 1976) were monophyletic and sister to Cenophengus + Mastinocerinae in all analyses, with variable statistical support. The monophyly of Phengodinae was
Figure 3. A, four-cluster likelihood mapping test of alternative phylogenetic hypotheses showing the placement of Cenophengus in relationship to other Phengodidae subfamilies. B, phylogeny of Phengodidae inferred from the partitioned maximum-likelihood analysis of the 45% completeness nucleotide matrix (45CPP-NT), based on 358 loci and 35112 bp. Node circles correspond to selected individual analyses, as identified in the circle key on the left side of the figure. Abbreviations: 45CPP-NT-BI, Bayesian analysis of the partitioned 45% completeness nucleotide matrix; 45CPP-NT-ML, maximum likelihood analysis of the partitioned 45% completeness nucleotide matrix; 45CPU-NT-BI, Bayesian analysis of the unpartitioned 45% completeness nucleotide matrix; 45CPU-NT-ML, maximum likelihood analysis of the unpartitioned 45% completeness nucleotide matrix; PP, Bayesian posterior probabilities of the clades of the majority-rule consensus tree of the posterior distribution generated via the Markov chain Monte Carlo; SH-aLRT, Shimodaira–Hasegawa-like approximate likelihood ratio test; UFBoot, ‘ultrafast’ bootstrap. Habitus images represent the sampled species; numbers of the images correspond to numbers near tip labels (taxon names); images 1–25 by authors, 26 by Alexander S. Prosvirov, 27 by Amir Weinstein.
moderately to strongly supported by sCF. The clade of Phengodes and Zarhipis was always monophyletic and with maximal support in all BI and ML analyses. Cenophengus was always sister to Mastinocerinae with variable support (topology supported also by the four-cluster quartet likelihood mapping analysis; Fig. 3A), and the monophyletic Mastinocerinae received variable support across different analyses.

Within Mastinocerinae, Akamboja Roza et al., 2017 and Taximastinocerus Wittmer, 1963 formed a maximally supported clade in all analyses, and in all ML and BI analyses they were sister to Perthodius Gorham, 1881, although this was not supported statistically (Fig. 3B). Mastinocerus (Mastinocerus) Solier, 1849 and Mastinomorphus Wittmer, 1976 formed a 100% supported clade in all analyses. Paramastinocerus Wittmer, 1976 (currently a subgenus of Mastinocerus) + Phrixothrix Olivier, 1909 and Nephroroma Wittmer, 1976 + Brasilocerus Wittmer, 1963 were recovered in all analyses, usually with moderate to maximal support, and these two clades formed a monophylum in all ML and BI analyses. Neither Euryopa Gorham, 1881 nor Mastinocerus, each represented by two species, was found to be monophyletic in any analysis. Both Euryopa spp. (although non-monophyletic) formed a clade together with Eurymastinocerus Wittmer, 1976 and Pseudomastinocerus Wittmer, 1963 in all ML and BI analyses, usually with maximal support. In all BI and ML analyses, there was a clade formed by Steneuryopa Wittmer, 1986, Decamastinocerus Wittmer, 1988, Howdenia Wittmer, 1976, and Cleidella Roza & Mermudes, 2020, with low to maximal support. The genus Ivisiphengus Roza, 2023 was recovered as sister to all remaining Mastinocerinae in all analyses. Oxymastinocerus Wittmer, 1963 was the second basalmost branch in all BI and one ML analyses; however, it was found as a more terminal lineage in all remaining analyses.

SYSTEMATICS

Based on the results of our phylogenomic analysis (Fig. 3B) and taking into consideration results of previous morphology-based and Sanger data-based phylogenies (see Discussion), in addition to the disparate morphology of the genus Cenophengus (Fig. 4; Vega-Badillo et al. 2021a, b, c), we here establish a new subfamily for this genus within the family Phengodidae.

Additionally, for the first time, we provide an identification key to the families of the ‘lampyroid clade’ (i.e. Sinopyrophoridae, Lampyridae, †Cretophengodidae, Rhagophthalmidae, and Phengodidae) and a key to the subfamilies of Phengodidae. Both keys are based on the male adult characters only, because adult females of some groups are unknown or variously modified morphologically (paedomorphic, often larviform), in ways that are not yet defined at the family level. There is also no information on immature stages for most groups, e.g. Sinopyrophoridae, †Cretophengodidae, Cydistinae, Cenophengus, and Penicillophorinae (Paulus 1975, Cicero 1988, Costa and Zaragoza-Caballero 2010, Kawashima et al. 2010, Zaragoza-Caballero and Pérez-Hernández 2014, Bi et al. 2019, Kundrata et al. 2019, 2022, Roza 2020, Li et al. 2021).
Identification key to the families of the ‘lampyroid clade’ based on adult males

1. Body strongly sclerotized; posterior angles of pronotum distinctly produced and acute; prosternum elongate, much longer than width of procoxal cavity; prosternal process long, reaching the mesoventrite, with clicking mechanism fully developed ................................................................. Synopyrophorinae

1'. Body moderately to weakly sclerotized; posterior angles of pronotum not distinctly produced; prosternum moderately to extremely transverse, as long as to much shorter than width of procoxal cavity; prosternal process reduced, not reaching mesoventrite, clicking mechanism absent ................................................................. 3

2. Prosternum in front of coxae longer than the width of procoxal cavity; prosternal process reaching posterior edge of procoxae; abdomen with six apparently immovable ventrites; known only from Cretaceous amber ................................................................. †Cretophengodidae

2'. Prosternum in front of coxae as long as to much shorter than procoxal cavity; prosternal process not reaching as far as posterior edge of procoxae; indistinct in many; abdomen with eight freely movable ventrites in most (i.e. sternites II–IX; except for Lampyridae: Luciolinae) ........................................................................................................... 3

3. Head covered by pronotum in dorsal view in most (except for some Ototretinae); antennae with 7–62 antennomeres (usually with 11); mesoscutum divided, without notch, and separated from mesoscutellum by suture; abdominal tergites IX and X fused (except for some Ototretinae); tergite X flat ................................................................................................................................................... Lampyridae

3'. Head not covered by pronotum; antennae with 10–12 antennomeres (mostly with 12); mesoscutum undivided, with notch on anterior margin, fused with mesoscutellum; abdominal tergites IX and X not fused; tergite X conical, tube-like ................................................................. 4

4. Antennomere III with aedeagus similar to following antennomeres; elytra covering the entire abdomen (with some exceptions, e.g. Falsophrixothrix Pic, 1937); elytral surface with oblique elevated costae; paired tibial spurs present ........................................................................................................................................... 4

4'. Antennomere III wider than long as to wide as long, short, cup-shaped, unlike following antennomeres in shape; elytra not covering the entire abdomen; elytral surface without elevated costae; tibial spurs absent ................................................................................................. Phengodidae

Family Phengodidae LeConte, 1861

Subfamily Cenophenginae Roza, Vega-Badillo, Zaragoza-Caballero, Kundrata & Mermudes subfam. nov.

Type genus: Cenophengus LeConte, 1881.

urn:lsid:zoobank.org:act:FA65EC81-6500-40A0-9B43-15AC4DC9D2B0

Diagnosis and comparison: Cenophenginae can be recognized based on the following unique combination of characters: gular sutures sinuous and widely separated (Fig. 4C); tentorial pits widely separated (Fig. 4C); pronotum from as wide as long to longer than wide (Fig. 4A, B); prosternum moderately transverse, in front of procoxae usually longer than width of procoxal cavity (Fig. 4C); elytra with sides subparallel (Fig. 4A, B); aedeagus with median lobe not divided and with a distinct flagellum; and each paramere widened at mid-length in dorsal view, subparallel in lateral view (Fig. 4D, E).

Members of this subfamily can be distinguished easily from all other Phengodidae by the parameres, which are each subparallel over much of their length in lateral view (parameres tapered beyond mid-length or gradually narrowed in other subfamilies). Cenophenginae are similar to the Asian Cydistinae in having a relatively longer prosternum than all remaining New World phengodids; however, they differ from Cydistinae in having the median lobe of the aedeagus extremely bent at base, undivided, and with a distinct flagellum (straight at base, divided into a dorsal and a ventral lobe, and without any visible flagellum in Cydistinae). Additionally, Cenophenginae differ from all remaining New World phengodids (i.e. Mastinocerinae, Pencillophorinae, and Phengodinae) by having the gular sutures sinuous and widely separated, the posterior tentorial pits widely separated (differently shaped in these subfamilies), and by the prosternum moderately transverse, ~1.3–1.5 times wider than medial length, and in front of coxae ~0.5 times width of procoxal cavity (Fig. 4); extremely transverse, ~2.5–3.0 times wider than long medi ally, and in front of coxae ~0.3 or less times width of procoxal cavity in other New World subfamilies).

Composition and distribution: One genus, Cenophengus, with 30 described species from the Nearctic and Neotropical realms. Species are distributed from the south of the USA to Costa Rica, with the highest diversity in Mexico (Vega-Badillo et al. 2021a).

Identification key to the subfamilies of Phengodidae based on adult males

1. Antennae without rami ......................................................................................................................................................... Penicillophorinae

1'. Antennae with rami present on at least antennomeres IV–IX ................................................................................................. 2

2. Gular sutures sinuous and widely separated (by about the mentum width); posterior tentorial pits widely separated (by distance greater than the pit diameter); prosternum moderately transverse, 1.3–1.5 times wider than long medi ally, in front of coxae ≥0.5 times width of procoxal cavity ....................................................................................................................... 3
DISCUSSION

Phylogenetic position of Phengodidae in Elateroidea

Our anchored phylogenomics results agree with previous molecular studies in showing the Asian Rhagophthalmidae as the sister group of Phengodidae (e.g. Kundrata et al. 2014, 2019, McKenna et al. 2015, Douglas et al. 2021, Kusy et al. 2021). Despite never being recovered in morphology-based phylogenies (e.g. Branham and Wenzel 2001, Zaragoza-Caballero and Zurita-Garcia 2015), the close relationship between Phengodidae and Rhagophthalmidae was already hypothesized by Crowson (1972), based on morphological and ecological similarities between their larvae and in having paedomorphic females. Rhagophthalmidae were sometimes even placed within a widely delimited Phengodidae (Crowson 1972, Lawrence and Newton 1995, Bocak 2007). Although a reliable synapomorphy for the Phengodidae + Rhagophthalmidae clade is missing, there is currently little doubt that the two families are sister to each other and form a monophyletic group.

The phylogenetic position of clade Phengodidae + Rhagophthalmidae varies between Sanger and phylogenomics-based analyses. When analysing Sanger datasets, numerous studies have found a sister relationship of Phengodidae + Rhagophthalmidae with Elateridae (e.g. Bocakova et al. 2007, Amaral et al. 2014, Kundrata et al. 2014). Meanwhile, the much larger phylogenomic datasets indicate a sister relationship between Phengodidae + Rhagophthalmidae and Lampyridae (Zhang et al. 2018, Douglas et al. 2021, Kusy et al. 2021, Cai et al. 2022), which is in agreement with analyses using mitogenomes and RNA sequencing data (Amaral et al. 2016, 2019). Given this context, our finding that Phengodidae + Rhagophthalmidae was sister to Elateridae is novel in a phylogenomic framework. However, given that our analysis is focused mainly on the higher classification of Phenogidae (whereas the above-mentioned analyses investigated Elateroidea or even whole Coleoptera), sampling of other elateroid outgroups (including Lampyridae and Elateridae) was small, and inferred relationships outside of the Phengodidae + Rhagophthalmidae clade should be taken with caution.

On a broader scale, Phengodidae belong to the ‘lampyroid clade’ defined by Kusy et al. (2021). This clade was originally proposed to accommodate Lampyridae, Phengodidae, Rhagophthalmidae, and a recently described enigmatic bioluminescent genus, Sinopyrophorus Bi & Li in Bi et al. (2019), which was originally placed in Elateridae (Bi et al. 2019) but was treated as a separate family by Kusy et al. (2021). Later, Li et al. (2021) suggested that the newly described Mesozoic Cretophengodidae also belonged to the ‘lampyroid clade’, hypothesizing its close relationships with Phengodidae and Rhagophthalmidae. Given that the interrelationships between the Elateridae and ‘lampyroid’ group of families are not yet fully understood (Douglas et al. 2021, Kusy et al. 2021), further phylogenomic studies are needed with an expanded taxon sampling of all above-mentioned families. Although morphology-based phylogenies are generally less well supported and resolved (see Solis and Solis 2003) than those based on genomic data, partly because of fewer characters available for analysis, larval characters might also provide further evidence regarding the monophyly of the ‘lampyroid clade’ (see e.g. Phengodidae + Lampyridae in the paper by Beutel 1995).

Phengodidae monophyly and subfamilial classification

Our preferred topology showed the currently defined Phengodidae (i.e. including Cydistinae) as monophyletic with robust statistical support. Cydistinae were sister to the remaining Phengodidae, and Phengodinae were sister to Cenopenginae + Mastinocerinae (Fig. 3).

The Asian Cydistinae include two genera and seven species (Kundrata et al. 2019). They were earlier classified by some authors close to genera currently in Rhagophthalmidae (Crowson 1972; for a review, see Kundrata et al. 2019). The Rhagophthalmidae are also distributed in Asia, but further east without distributional overlap (Kundrata et al. 2022). Our results reject a sister-group relationship between Cydistinae and Rhagophthalmidae (Fig. 3). Instead, Cydistinae were recovered as the sister group of the New World phengodids, in concordance with the findings of Kundrata et al. (2019) and Roza (2022), based on the analyses of Sanger data and morphological characters, respectively. Morphologically, Cydistinae are like Cenopengus in having the prothorax longer than wide and the prosternum that is 1.3–1.5 times wider than long medi ally. However, Cydistinae can be distinguished easily from all New World phengodids by the male genitalia, which are characterized by the straight median lobe divided into dorsal and ventral lobes (extremely bent in the anterior portion and undivided in all other phengodids) and absence of the flagellum (present in all other phengodids). Cydistinae can be differentiated further from Mastinocerinae, Pencillophorinae, and Phengodinae by having the gular sutures sinuous and widely separated, posterior
tentorial pits widely separated (differently shaped in these subfamilies), and by the longer prosternum, which is in front of coxae, as long as or even longer than the width of the procoxal cavity (extremely transverse, in front of coxae usually ≤0.3 times the width of procoxal cavity in these subfamilies; A.S.R. and R.K., personal observations).

The subfamily Phengodinae contains four genera and 61 species distributed in Neartic and Neotropical realms, from southern Canada to northern Argentina (Zaragoza-Caballero and Pérez-Hernández 2014). Previous morphology-based studies have recovered Phengodinae in various phylogenetic positions, i.e. nested within Mastinocerinae (Quintino 2017), being non-monophyletic with respect to Mastinocerinae (Zaragoza-Caballero and Zurita-García 2015), or being monophyletic and sister to Mastinocerinae (including Penicillophorinae; Roza 2022). Kundrata et al. (2019) found monophyletic Phengodinae sister to Mastinocerinae (without Cenophengus; for an overview, see Fig. 2). Our results newly show Phengodinae as monophyletic and generally sister to Cenophengus + Mastinocerinae. All our analyses showed Phengodinae to be monophyletic, with Microphengodes sister to Zarhipis + Phengodes. The close relationship between Zarhipis and Phengodes was also previously recovered by earlier analyses (e.g. Zaragoza-Caballero and Zurita-García 2015; Roza 2022). Further support for the monophyly of Phengodinae was provided by sCF, with moderate to strong support. Phengodinae can be distinguished easily from other subfamilies by the presence of antennal tubercles (absent in other subfamilies except for Mastinocerinae: Paraphrixothrix), apical maxillary palpomeres acuminate or subcylindrical (elongate-secuform to secuform in other subfamilies), and elytra slightly to strongly dehiscent (subparallel in other subfamilies).

Cenophenginae are established here for a single genus, Cenophengus, with 30 described species distributed in the area from the south of the USA to Costa Rica, with the largest number of species reported from Mexico (Vega-Badillo et al. 2021a). Cenophengus was classified earlier in Mastinocerinae, but morphology-based analyses (Zaragoza-Caballero and Zurita-García 2015; Roza 2022) and Sanger data-based ML analyses (Kundrata et al. 2019) suggested that this genus was the sister group of all other New World Phengodinae, and therefore, not a member of Mastinocerinae. Although there are some plesiomorphic characters shared by Cenophengus and Cydistinae (Zaragoza-Caballero and Zurita-García 2015, Kundrata et al. 2019; Roza 2022), no sister-taxon relationship between them has been recovered by any phylogenetic analysis. Here, we found Cenophengus to be sister to Mastinocerinae, a position previously found only in the Bayesian analysis by Kundrata et al. (2019). An overwhelming majority of the phylogenetic signal within our nucleotide dataset provided strong support for this placement (Fig. 3A). We here established a subfamily rank for this genus, which differs morphologically from all other New World Phengodinae (details in Results section and Vega-Badillo et al. 2021a, b). This taxonomic change will also enable much easier diagnosis of Mastinocerinae, which differ from Cenophengus in several characters on the head, thorax, and male genitalia (see below).

The subfamily Mastinocerinae contains 30 genera and 233 described species distributed from the southern USA to central Chile and Argentina, with especially high diversity reported from Brazil and Mexico (Zaragoza-Caballero and Pérez-Hernández 2014, Roza 2023). This is by far the most taxonomically problematic lineage of Phengodidae. Mastinocerinae hold the bulk of phengid species diversity and much morphological variability also. Therefore, it is not surprising that their monophyly has been questioned repeatedly. Genus Cenophengus, which differs morphologically and was found outside the core Mastinocerinae (Fig. 3B), is here placed in a separate subfamily. Additionally, the morphology-based analysis by Zaragoza-Caballero and Zurita-García (2015) recovered Mastinocerinae as non-monophyletic because it included some Phengodinae (i.e. Microphengodes + Pseudophengodes); however, this topology has never been found in any molecular-based study, including ours. In his morphology-based analyses, Roza (2022) found Penicillophorinae inside Mastinocerinae, and the enigmatic genus Paraptorthodius Schaeffer, 1904 outside Mastinocerinae. Unfortunately, none of these findings could be tested using our anchored hybrid enrichment dataset, because we failed to obtain data for those crucial taxa. However, our topologies showed strong support for a monophyletic Mastinocerinae (Fig. 3B; Table 1).

Although the newly circumscribed Mastinocerinae (i.e. without Cenophengus) were consistently recovered as monophyletic here, their internal relationships were weakly resolved (see the next subsection). Owing to the enormous morphological plasticity of Mastinocerinae, it is extremely difficult to characterize them morphologically as a group. For example, antennae might be composed of 12, 11, or 10 antennomeres, labial palps might have one, two, or three palpomeres, and wing venation is also extremely variable (Roza 2022). Although there are some distinct characteristics, such as tarsal combs (shared with some Penicillophorinae; Zaragoza-Caballero and Pérez-Hernández 2014) and membranous projections of abdominal ventrite VIII (Roza 2023), those are also variable and absent in many genera.

Given that there are currently no known synapomorphies for this group, future morphology-oriented studies are needed to improve the diagnosis of this subfamily. Mastinocerinae can, however, be distinguished from the Cenophenginae and the Cydistinae by their gular sutures and posterior tentorial pits, which are both either fused or closely approximate (both widely separated in Cenophenginae and Cydistinae), and by the distinctly transverse prosternum, which is ~2.5–3.0 times wider than long medially (much less transverse, being 1.3–1.5 times wider than long medially in Cenophenginae and Cydistinae). Additionally, Mastinocerinae do not have subparallel-sided palpomeres as found in Cenophenginae (Vega-Badillo et al. 2021a, b) and do not have the straight median lobe, which is divided into a dorsal and a ventral lobe, which is found in Cydistinae (Kundrata et al. 2019). Compared with Phengodinae, the members of Mastinocerinae can be distinguished by the absence of antennal tubercles (present in Phengodinae but also in Mastinocerinae: Paraphrixothrix), by the apical maxillary palpomere elongate-secuform to secuform (acuminate or subcylindrical in Phengodinae), and by the subparallel elytra (slightly to strongly dehiscent in Phengodinae). Finally, all known Penicillophorinae differ from Mastinocerinae by the moniliform to serrate antennae without rami (antennal rami from antennomere IV to the subapical antennomere in all Mastinocerinae).
The last subfamily, Penicillophorinae, contains five genera and six species described from Mexico, Guatemala, Dominican Republic, and Colombia (Zaragoza-Caballero and Pérez-Hernández 2014). It is a morphologically diverse group, in which several characters, such as the number of antennomeres, the gular pits, tarsal combs, and wing venation are extremely variable. However, the group can be distinguished easily from all other Phengodidae subfamilies by the moniliform to serrate antennae without rami. Unfortunately, we failed to obtain anchored hybrid enrichment data from the available specimens; therefore, the placement of this highly morphologically modified lineage remains a mystery. In their morphology-based analysis, Zaragoza-Caballero and Zurita-García (2015) recovered Penicillophorinae as monophyletic but outside the Phengodidae, in a close relationship with Telegeusinae, which are currently in Omethidae, i.e. one of the basalmost groups in Elateroidea (Kundrata et al. 2014). Not surprisingly, Roza (2022) recovered Penicillophorinae nested within the Mastinocerinae based on morphological evidence. In fact, as currently defined, these groups differ diagnostically only by the non-ramose antennae of penicillophorines, hence the hypothesis that the penicillophorines are within Mastinocerinae is plausible.

**Intergeneric relationships within the subfamilies of Phengodidae**

Our study clearly contradicts previous morphological analyses by confirming both Phengodinae and Mastinocerinae as monophyletic. It also contradicts both morphological and Sanger-based analysis by positioning *Cenophengus* as the sister to Mastinocerinae. However, we believe that our new findings are correct given the robustness of the phylogenomic analyses presented here, both in the larger taxon sampling and in the higher statistical support of recovered clades (Fig. 3; Table 1).

The internal relationships of Mastinocerinae, however, are far from being resolved. When comparing analyses, the position of two Mastinocerinae genera varied the most: *Iviphengus* and *Oxymastinocerus*. Both genera were recovered either among basal splits or in more internal positions of Mastinocerinae. It is important to highlight that two genera (*Mastinocerus* and *Euryopa*) were found to be non-monophyletic in our analyses. This is a reflection of the poorly resolved taxonomy of the family and of how much the Phengodidae genera need revisions and redefinitions. It is evident from all morphology-based analyses of Phengodidae (Zaragoza-Caballero and Zurita-García 2015, Quintino 2017; Roza 2022) that characters usually used to delimit the genera (e.g. interantennal distance, number of labial palpi, presence of tentorial pits, and tarsal combs) are highly homoplastic. Therefore, Phengodidae taxonomy, and that of Mastinocerinae in particular, requires revision and a search for new reliable characters.

**CONCLUSION**

Using anchored hybrid enrichment, we successfully reconstructed the most extensive molecular phylogeny of Phengodidae to date. We used this phylogeny to evaluate the higher-level classification of the group. Phengodidae were recovered as sister to Rhagophthalmidae, and the Old World Cydistinae as sister to the remaining Phengodidae from the New World, in agreement with previous molecular and morphological studies. In contrast to most of the previous hypotheses, however, both Phengodinae and Mastinocerinae were recovered as monophyletic. Contrary to most previous hypotheses, which placed *Cenophengus* as sister to all New World Phengodidae, we found this genus with high support as sister to the remaining Mastinocerinae. Considering its phylogenetic placement and morphological disparity in comparison to other Mastinocerinae, we here establish *Cenophenginae* subfam. nov.

We present the largest and most comprehensive sampling of Phengodidae in a molecular-based study so far, but we had limited success in resolving the generic relationships of the diverse Mastinocerinae. Most internal nodes were poorly supported and topologically unstable between analyses. It is clear that more studies are needed to enlighten the evolution and classification of this morphologically diverse and species-rich group.

Future phylogenetic research, with the inclusion of taxa unavailable here and with the combination of genomic, mitogenomic, and comprehensive morphological data, could test the position of *Cenophengus* further, resolve the internal topology of Mastinocerinae, and place the Penicillophorinae. We hope that our study will accelerate efforts towards a stable natural classification of Phengodidae and to our understanding of the relationships between the main clades and all genera. We hope this will, in turn, allow better biological insights into this ecologically important and evolutionarily fascinating group of beetles.

**SUPPLEMENTARY DATA**

Supplementary data is available at Zoological Journal of the Linnean Society online.

**ACKNOWLEDGEMENTS**

We are grateful to Jim McClarin (Ecuador), Karlyn Lewis (USA), Matt Bertone (USA), Ben Theodore (Israel), Mike Quinn (USA), Tyler Bishop (USA), Norton Santos (Brazil), Chris Carlton (USA), Dusty_in_vilas-tx (iNaturalist, USA), Alexander S. Prosvirov (Russia), and Amir Weinstein (Israel) for allowing us to use their photographs. We thank Milada Bocakova (Czech Republic) and Alexander S. Prosvirov (Russia) for some of the samples used in our study.

**CREDIT STATEMENT**

Vinicius S. Ferreira (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing—original draft) [equal], André S. Roza (Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Visualization, Writing—original draft) [equal], Felipe F. Barbosa (Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Visualization, Writing—review and editing), Viridiana Vega-Badillo (Investigation, Resources, Writing—review and editing), Santiago Zaragoza-Caballero (Investigation, Resources), José Ricardo M. Mermudes (Funding acquisition, Investigation, Writing—review and editing), Michael A. Ivie (Investigation, Resources, Writing—review and editing), Aslak Hansen (Formal analysis, Funding acquisition, Investigation, Methodology, Writing—review and editing), Hume B. Douglas (Investigation, Methodology, Writing—review and editing), Alexey Soledovnikov (Investigation, Methodology, Resources, Writing—review and editing), Robin Kundrata (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Visualization, Writing—original draft).
Conflict of Interest

None declared.

Funding

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 101018841 (postdoctoral fellowship of V.S.F.) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of Brazil (grant agreement 202659/2015-7). A.S.R. was supported by the CNPq (process 150747/2022-5) and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Education Ministry of Brazil (CAPES). J.R.M.M. was funded by CNPq—Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—FAPERJ; 205.818/2022 and 205.819/2022). F.F.B. was supported by the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ; 205.821/2022). A.S.R. was supported by the CNPq (process 150747/2022-5) and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, process 150747/2022-5).

Data Availability

Demultiplexed, raw-read FASTq files were deposited in the NCBI SRA under BioProject PRJNA1104084. Scripts, alignments, and phylogenetic trees are available at FigShare (https://doi.org/10.6084/m9.figshare.25650891.v3). All data underlying this article are available upon reasonable request from one of the corresponding authors.

References


Vega-Badillo V, Morrone JJ, Zaragoza-Caballero S. Revision of the genus *Cenophengodes* LeConte, 1881 (Coleoptera, Phengodidae), with the description of four new species, new geographic records and a new synonymy. ZootKeys 2021a; 1068:73–148. https://doi.org/10.3897/zootkeys.1068.70295
Zaragoza-Caballero S, Pérez-Hernández CX. Sinopsis de la Familia Phengodidae (Coleoptera): Tracterzos, Bigotudos, Glow-Worms, Rail-Road Worms, or Bessouros trem de ferro. Mexico: Universidad Nacional Autónoma de Mexico (UNAM), 2014.
Zaragoza-Caballero S, Zurita-García ML. A preliminary study on the phylogeny of the family Phengodidae (Insecta: Coleoptera). Zootaxa 2015; 3947:527–42. https://doi.org/10.11646/zootaxa.3947.4.4