

Phylogenomic analyses clarify the pattern of evolution of Adephaga (Coleoptera) and highlight phylogenetic artefacts due to model misspecification and excessive data trimming

ALEXANDROS VASILIKOPOULOS¹, MICHAEL BALKE²,
SANDRA KUKOWKA¹, JAMES M. PFLUG³,
SEBASTIAN MARTIN¹, KAREN MEUSEMANN⁴,
LARS HENDRICH², CHRISTOPH MAYER¹,
DAVID R. MADDISON³, OLIVER NIEHUIS⁴, ROLF G.
BEUTEL⁵ and BERNHARD MISOF⁶

¹Centre for Molecular Biodiversity Research, Zoological Research Museum Alexander Koenig, Bonn, Germany, ²Department of Entomology, SNSB-Bavarian State Collections of Zoology, Munich, Germany, ³Department of Integrative Biology, Oregon State University, Corvallis, OR, U.S.A., ⁴Department of Evolutionary Biology and Animal Ecology, Institute of Biology I (Zoology), Albert-Ludwig University of Freiburg, Freiburg, Germany, ⁵Institut für Zoologie und Evolutionsforschung, Friedrich-Schiller-Universität Jena, Jena, Germany and ⁶Zoological Research Museum Alexander Koenig, Bonn, Germany

Abstract. Adephaga is the second largest suborder of Coleoptera and contains aquatic and terrestrial groups that are sometimes classified as Hydradephaga and Geadephaga, respectively. The phylogenetic relationships of Adephaga have been studied intensively, but the relationships of the major subgroups of Geadephaga and the placement of Hygrobiidae within Dytiscoidea remain obscure. Here, we infer new DNA-hybridization baits for exon-capture phylogenomics and we combine new hybrid-capture sequence data with transcriptomes to generate the largest phylogenomic taxon sampling within Adephaga presented to date. Our analyses show that the new baits are suitable to capture the target loci across different lineages of Adephaga. Phylogenetic analyses of moderately trimmed supermatrices confirm the hypothesis of paraphyletic ‘Hydradephaga’, with Gyrinidae placed as sister to all other families as in morphology-based phylogenies, even though quartet-concordance analyses did not support this result. All analyses conducted with site-heterogeneous models suggest Trachypachidae as sister to a clade Carabidae + Cicindelidae in congruence with results from morphological studies. Haliplidae is inferred as sister to Dytiscoidea, while a clade of Noteridae (+ most likely Meruidae) is inferred as sister to all remaining Dytiscoidea. A strongly supported clade Hygrobiidae + (Amphizoidae + monophyletic Aspidytidae) is inferred in most analyses of moderately trimmed supermatrices when a site-heterogeneous model is used. In general, we find that stringent trimming of supermatrices results in reduced deviation from model assumptions but also in reduction of phylogenetic information. We also find that site-heterogeneous C60 models provide greater stability of phylogenetic relationships of Adephaga across analyses of different amino-acid supermatrices than site-homogeneous models. Thus, site-heterogeneous C60

Correspondence: Alexandros Vasilikopoulos, Centre for Molecular Biodiversity Research, Zoological Research Museum Alexander Koenig, 53113 Bonn, Germany. E-mail: alexvasilikop@gmail.com
Bernhard Misof, Zoological Research Museum Alexander Koenig, 53113 Bonn, Germany. E-mail: b.misof@leibniz-zfmk.de

[Corrections added on 30 July 2020, after the first online publication: Copyright line has been updated]

© 2021 The Authors. *Systematic Entomology* published by John Wiley & Sons Ltd on behalf of Royal Entomological Society.

991

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

models can potentially reduce incongruence in phylogenomics. Lastly, we show that gene-tree errors are prominent in the data, even after sub-sampling genes to reduce these errors, but we also show that subsampling genes based on the likelihood mapping criterion in summary coalescent analyses results in higher topological congruence with the concatenation-based tree. Overall, our analyses demonstrate that moderate alignment trimming strategies, application of site-heterogeneous models and mitigation of gene-tree errors should be routinely included in the phylogenomic pipeline in order to more accurately infer the phylogeny of species.

Introduction

Beetles (Coleoptera) are the most speciose insect order and their phylogeny has been the focus of attention for many decades (Crowson, 1960; Lawrence & Newton, 1982; Hunt *et al.*, 2007; Lawrence *et al.*, 2011; Beutel *et al.*, 2019a, 2020; McKenna *et al.*, 2019). Polyphaga is the largest beetle suborder with numerous phytophagous species (McKenna *et al.*, 2019) but also many other feeding habits. Adephaga, which mostly includes predacious species, is the second largest beetle suborder with more than 45 000 species assigned to 11 families (Beutel *et al.*, 2020; Duran & Gough, 2020). The family-level phylogenetic relationships of Adephaga have been extensively debated but scientists are now reaching a consensus on the most likely scenario of their evolution (McKenna *et al.*, 2019; Beutel *et al.*, 2020; Gustafson *et al.*, 2020). Despite this, open questions remain, such as the phylogenetic relationships of the major terrestrial groups, the phylogenetic position of Hygrobiidae within Dytiscoidea and the intra-familial relationships within Carabidae, Cicindelidae and Dytiscidae (Michat *et al.*, 2017; Vasilikopoulos *et al.*, 2019; Beutel *et al.*, 2020; Gustafson *et al.*, 2020). In addition, previous analyses of family-level relationships in Adephaga have suggested that some results of previous studies might be artefacts due to systematic errors (Cai *et al.*, 2020). In this study, we address these unresolved issues by combining newly generated exon-capture sequence data with transcriptomic sequence data to infer the phylogeny of Adephaga based on extensive sampling of species.

The majority of species diversity in Adephaga belong to the terrestrial family Carabidae (ground beetles, >35 000 extant species), whereas the closely related family Cicindelidae is a medium-sized terrestrial group (tiger beetles, >2400 extant species). Trachypachidae is another terrestrial family with only six extant species (Beutel *et al.*, 2020; Duran & Gough, 2020; Lorenz, 2020). These families have been collectively referred to as 'Geadephaga' (Crowson, 1960). The monophyly of this unit has been disputed in the past based on analyses of morphological characters (Burmeister, 1976; Beutel & Roughley, 1988), but most recent morphological and molecular analyses suggest a single origin of the terrestrial adephagan groups (Beutel *et al.*, 2006, 2020; Maddison *et al.*, 2009; McKenna *et al.*, 2019; Gustafson *et al.*, 2020). In contrast to this, the phylogenetic relationships among Carabidae, Cicindelidae and Trachypachidae remain controversial as different phylogenomic analyses

have produced different topologies. Phylotranscriptomic analyses have placed Trachypachidae as sister to Carabidae + Cicindelidae (McKenna *et al.*, 2019). In contrast, analyses of mitochondrial genomes suggested a weakly supported clade of Cicindelidae + Trachypachidae as sister to Carabidae (López-López & Vogler, 2017), whereas analyses of ultraconserved elements (UCEs) suggested Cicindelidae + (Carabidae + Trachypachidae) (Gustafson *et al.*, 2020). It should be noted, however, that the taxon sampling of previous phylogenomic studies was not sufficient to test the monophyly of Carabidae and Cicindelidae and to robustly infer the phylogenetic position of the small family Trachypachidae (Zhang *et al.*, 2018b; McKenna *et al.*, 2019; Gough *et al.*, 2020; Gustafson *et al.*, 2020). In addition, the results of some molecular analyses do not agree with results of morphological studies that suggest Trachypachidae as sister to Carabidae + Cicindelidae (Beutel *et al.*, 2020). Therefore, a re-evaluation of the relationships of Geadephaga with careful examination of potential sources of systematic error and increased species sampling is needed.

The species of the remaining eight families of Adephaga (Amphizoidae, Aspidytidae, Dytiscidae, Haliplidae, Hygrobiidae, Meruidae, Noteridae and Gyrinidae) occur primarily in aquatic or semi-aquatic habitats (Jäch & Balke, 2008; Short, 2018). Most species of Dytiscidae, Gyrinidae, Hygrobiidae and Noteridae are strictly aquatic. Species of Amphizoidae are also aquatic, whereas Aspidytidae and Meruidae occur in hygropetric habitats (Kavanaugh, 1986; Balke *et al.*, 2003; Spangler & Steiner, 2005; Vasilikopoulos *et al.*, 2019). Crowson (1960) suggested that all the aquatic and semi-aquatic groups constitute a monophylum to which he referred to as 'Hydradephaga'. Only a few molecular phylogenetic studies have supported this concept (Shull *et al.*, 2001; Ribera *et al.*, 2002; McKenna *et al.*, 2015; López-López & Vogler, 2017), whereas the monophyly of this group has been refuted in more comprehensive studies based on analyses of morphological characters and genomic data (Beutel & Roughley, 1988; Beutel *et al.*, 2006, 2020; Baca *et al.*, 2017a; Gustafson *et al.*, 2020; McKenna *et al.*, 2019). More specifically, the placement of Gyrinidae as sister to all other Adephaga is currently a well-accepted scenario (Baca *et al.*, 2017a; Beutel *et al.*, 2020; Beutel & Roughley, 1988; Gustafson *et al.*, 2020; but see Freitas *et al.*, 2021). In addition, most analyses suggest a sister group relationship of Haliplidae to the superfamily Dytiscoidea (which includes Amphizoidae, Aspidytidae,

Dytiscidae, Hygrobiidae, Meruidae and Noteridae) and a clade Meruidae + Noteridae as sister to all remaining families of Dytiscoidea (Beutel *et al.*, 2006; Baca *et al.*, 2017a; Vasilikopoulos *et al.*, 2019; Gustafson *et al.*, 2020). Despite this, the phylogenetic position of the family Hygrobiidae (squeak beetles) within Dytiscoidea remains contentious (Toussaint *et al.*, 2016; Baca *et al.*, 2017a; Vasilikopoulos *et al.*, 2019, 2021; Cai *et al.*, 2020; Gustafson *et al.*, 2020).

The issue of model and data selection has received considerable attention in the context of the phylogeny of insects and other groups (Misof *et al.*, 2013; Lanfear *et al.*, 2014; Song *et al.*, 2016; Feuda *et al.*, 2017; Ballesteros & Sharma, 2019; Cai *et al.*, 2020; Kapli & Telford, 2020; Evangelista *et al.*, 2021). Specifically, several studies have demonstrated that using unrealistic models of molecular evolution might result in spurious phylogenetic estimates (Lartillot *et al.*, 2007; Song *et al.*, 2010, 2016; Wang *et al.*, 2019; Crotty *et al.*, 2020; Kapli & Telford, 2020). It has also been suggested that selecting sites or genes with reduced deviation from model assumptions might be beneficial (Philippe *et al.*, 2017; Simion *et al.*, 2020). In contrast, other authors have shown that it is difficult to remove systematic bias from the data without removing phylogenetic signal at the same time (Mongiardino Koch & Thompson, 2021). Such issues relating to model misspecification and data selection were also recently discussed in the context of the phylogeny of Adephaga (Vasilikopoulos *et al.*, 2019, 2021; Cai *et al.*, 2020). Heterogeneous composition of amino acids and nucleotides across taxa or across alignment sites, systematic bias resulting from hypervariable alignment sites, and deficient taxon sampling are among the potential factors affecting the internal phylogeny of the superfamily Dytiscoidea, including the monophyly of Aspidytidae (Baca *et al.*, 2017a; Vasilikopoulos *et al.*, 2019; Cai *et al.*, 2020; Gustafson *et al.*, 2020). Furthermore, it has been observed that summary coalescent and concatenation-based phylogenetic analyses often deliver incongruent topologies within Adephaga (Baca *et al.*, 2017a; Gustafson *et al.*, 2020; Freitas *et al.*, 2021). However, the factors that contribute to this incongruence remain poorly understood (Baca *et al.*, 2017a; Freitas *et al.*, 2021). In particular, the extent of gene-tree errors in previous summary coalescent analyses of Adephaga and their effect on species-tree estimation remain uncertain (Baca *et al.*, 2017a; Vasilikopoulos *et al.*, 2019; Gustafson *et al.*, 2020). Thus, two issues are imperative for a thorough assessment of the phylogenetic relationships of Adephaga in the light of increased taxon sampling: (i) evaluating the extent of gene-tree errors in summary coalescent analyses and (ii) using biologically realistic models in concatenation-based analyses.

The issue of data-collection strategies in phylogenetics has also been extensively discussed (McCormack *et al.*, 2013; Young & Gillung, 2020) and several hybrid-enrichment (or sequence-capture) approaches for phylogenomics have been developed (Faircloth *et al.*, 2012; Lemmon *et al.*, 2012; Bragg *et al.*, 2016; Mayer *et al.*, 2016). The use of UCEs (Faircloth *et al.*, 2012) is the only sequence-capture approach that has been applied to the phylogeny of Adephaga so far (Baca *et al.*, 2017a; Gustafson *et al.*, 2020). However, some authors have suggested the use of other sequence-capture or transcriptomic approaches

in addition to or independent of the UCE approach (Bank *et al.*, 2017; Karin *et al.*, 2020) in an attempt to validate and compare results among studies (see also Vasilikopoulos *et al.*, 2021). In this sense, hybrid-enrichment of protein-coding exons (Bank *et al.*, 2017; Sann *et al.*, 2018; Mayer *et al.*, 2021) is another sequence-capture method that can provide complementary or independent evidence for testing the validity of previously suggested phylogenetic hypotheses of Adephaga.

Concerning the utility of the exon-capture approach across different scales of molecular divergence, previous research suggests it is only effective for investigating taxonomic clades characterized by small to moderate levels of molecular divergence (Bi *et al.*, 2012; Bragg *et al.*, 2016; Mayer *et al.*, 2016). Nevertheless, if transcriptomic resources are available for a broad set of species within the group of interest, they can be used for testing the applicability of exon-specific DNA-hybridization baits at deeper evolutionary timescales with higher levels of molecular divergence. Additionally, recently developed bioinformatic approaches are able to automatically detect suitable regions for bait design in aligned DNA sequence data, including protein-coding data, by minimizing overall bait-to-target distances (Mayer *et al.*, 2016). Therefore, these bioinformatic approaches offer a promising solution to the problem of designing probes with broad phylogenetic applicability (Lemmon & Lemmon, 2013). Transcriptomic and genomic resources for adephagan beetles have increased considerably in the last few years (Gustafson *et al.*, 2019; McKenna *et al.*, 2019; Vasilikopoulos *et al.*, 2019). Combined with the above-mentioned bioinformatic approaches, these new data make it now possible to test the applicability and efficiency of exon capture for deep-level phylogenetics in Adephaga.

In this study, we develop a new set of DNA-hybridization baits specifically tailored to capture hundreds of single-copy protein-coding genes across adephagan lineages and generate new hybrid-capture data to infer the phylogeny of Adephaga. We test the efficiency of this set of baits for locus recovery in a large number of specimens. We also combine the newly generated hybrid-capture data with transcriptomes to generate the most species-rich phylogenomic dataset for adephagan beetles presented to date. In order to avoid biased estimates of phylogeny of Adephaga, we take measures to minimize phylogenetic artefacts by: (i) employing biologically realistic models of sequence evolution and (ii) by reducing potentially biasing factors in the data, using data-filtering strategies that select conserved alignment sites. We evaluate the effects of model misspecification and excessive data trimming both on the results of phylogenetic tree reconstructions and on quartet-based analyses of phylogenetic incongruence in an attempt to acquire a more detailed view of phylogenetic signal, conflict and bias in the backbone phylogeny of Adephaga. We also explore whether or not gene-tree discordance (GTD) can be explained by gene-tree estimation errors and suggest possible strategies for selecting informative genes that may increase congruence with concatenation-based analyses. Lastly, we discuss our results in the context of the morphological evolution of Adephaga.

Materials and methods

Taxon sampling

We combined 38 transcriptomes from 23 species of Adephaga and 15 outgroup species (File S1: Table S1) with newly generated exon-capture sequence data from 95 species of Adephaga (File S1: Table S2, note that two specimens of *Hydrocanthus oblongus* were initially processed but only one was included in the present study). Our initial taxon sampling comprised data from 118 species of Adephaga representing all families, except the monotypic Meruidae, and 21 outgroup species (two terminals of Hymenoptera, three of Mecoptera, two of Strepsiptera, four of Neuropterida, two of Myxophaga, two of Archostemata and six of Polyphaga). The initial taxon sampling includes the six reference species of the ortholog set (see below).

Inference of bait nucleotide sequences for hybrid enrichment of protein-coding exons

We used 24 transcriptomes of Adephaga as a basis to build codon-based nucleotide multiple sequence alignments (MSAs) of orthologous genes and search for MSA regions that are suitable for bait design within Adephaga (see File S1: Table S1 and File S2). First, we used a custom ortholog gene set consisting of 3085 ortholog clusters of single-copy genes (COGs) at the hierarchical level of Holometabola (Vasilikopoulos *et al.*, 2019) to assign orthologous transcripts from each transcriptome to each COG. Orthology assignment of transcripts to each COG was performed with Orthograph v. 0.6.1 (Petersen *et al.*, 2017). Subsequently, we followed procedures for amino-acid MSA, alignment refinement, outlier-sequence removal and removal of reference taxa before generating codon-based nucleotide MSAs (see supplementary information of Misof *et al.*, 2014a for details on these procedures). We then used Baitfisher v. 1.2.7 (Mayer *et al.*, 2016) to screen the codon-based MSAs for regions that are appropriate for bait design within the Adephaga clade (File S2). We conducted seven different tiling design experiments, corresponding to different lengths of bait regions, bait offsets and total number of baits in order to capture as many promising coding exons as possible while accounting for variable exon length, possibly large amount of missing data or hypervariable regions in some parts of the MSAs (see Mayer *et al.*, 2016 for details of the procedure used by Baitfisher, File S1: Table S3). In order to exclude baits targeting multiple genomic regions in adephagan genomes, we filtered the resulting baits (separately for each tiling design experiment) by conducting a blast search against a draft genome assembly of the beetle *Bembidion corgenoma* (Gustafson *et al.*, 2019, as *Bembidion haplogonum*, see File S2). We then selected only one bait region per coding exon in each tiling design experiment: the one that required the minimum number of baits (Mayer *et al.*, 2016). Subsequently, for those exons that were captured in multiple tiling design experiments only the longest bait region among experiments was considered (see File S2). In total, we inferred 49 786 120 bp-long bait sequences for targeting 923

protein-coding exons from 651 genes. For the sake of simplicity, we refer to our approach as ‘exon capture’ in this study instead of ‘coding-exon capture’, even though in our procedure we intended to include and analyse only the protein-coding regions of the targeted exons (i.e., excluding 3′ and 5′ untranslated regions).

Tissue preservation, total genomic DNA extraction, next-generation sequencing library preparation and hybrid enrichment of protein-coding exons

Most specimens used for hybrid-enrichment of target genomic DNA (gDNA) were freshly collected and preserved in 96% ethanol but we also used a few dry pinned museum specimens (File S1: Table S4). Total gDNA was extracted from 95 specimens of Adephaga (File S1: Tables S2, S4) using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) and eluted in 100 µL nuclease-free water. Whenever available, voucher material has been deposited at the Zoologische Staatssammlung München (Zoological State Collection in Munich, Germany, see File S1: Table S4). Quality and quantity of the extracted gDNA were assessed with a Fragment Analyzer (Agilent Technologies Inc., Santa Clara, CA, U.S.A.) and a Quantus Fluorometer (Promega, Fitchburg, WI, U.S.A.). Whenever sufficient amount of extracted DNA was available, we used 100 ng of DNA diluted in 10 µL for fragmentation before library preparation, otherwise less than 100 ng were used. First, gDNA was sheared into fragments of 150–400 bp using a Bioruptor Pico sonication device (Diagenode s.a., Seraing, Belgium). Multiple shearing steps were performed for each sample until at least ~90% of fragments were within the desired length interval. The quality and quantity of the fragmented gDNA were assessed with a Fragment Analyzer at the end of each shearing step. For library preparation, we followed the SureSelectXT2 Target Enrichment System Protocol for Illumina Paired-End Multiplexed Sequencing (Version E1 published in June 2015 by Agilent Technologies Inc.) with some minor modifications (Bank *et al.*, 2017). Specifically, in the library preparation steps ‘End Repair’ and ‘A-tailing’, we reduced the reaction volume specified in Agilent’s protocol (pages 43–49 for 100 ng DNA samples) by 50% as described by Bank *et al.* (2017). Subsequently, adapter ligation was performed with the NEBNext Quick Ligation Module and the adapters from the NEBNext Multiplex Oligos for Illumina (Dual Index Set1) kit. Next-generation sequencing (NGS) library PCR was then performed with the NEBNext Multiplex Oligos for Illumina and the NEBNext Q5 HotStart HiFi PCR Master Mix, to dual-index the libraries. Cycles of the NGS library PCR were adjusted as follows (due to the concentration measurements after ‘A-tailing’): 98°C for 30 s, followed by 8–10 cycles of 98°C for 10 s and 65°C for 75 s, followed by 5 min at 65°C followed by 4°C until the samples were removed from the thermocycler. Subsequently, all steps of the hybrid enrichment followed the protocol given by Bank *et al.* (2017) with modifications adjusted to the number of library pools and volume concentrations in our study (see File S2).

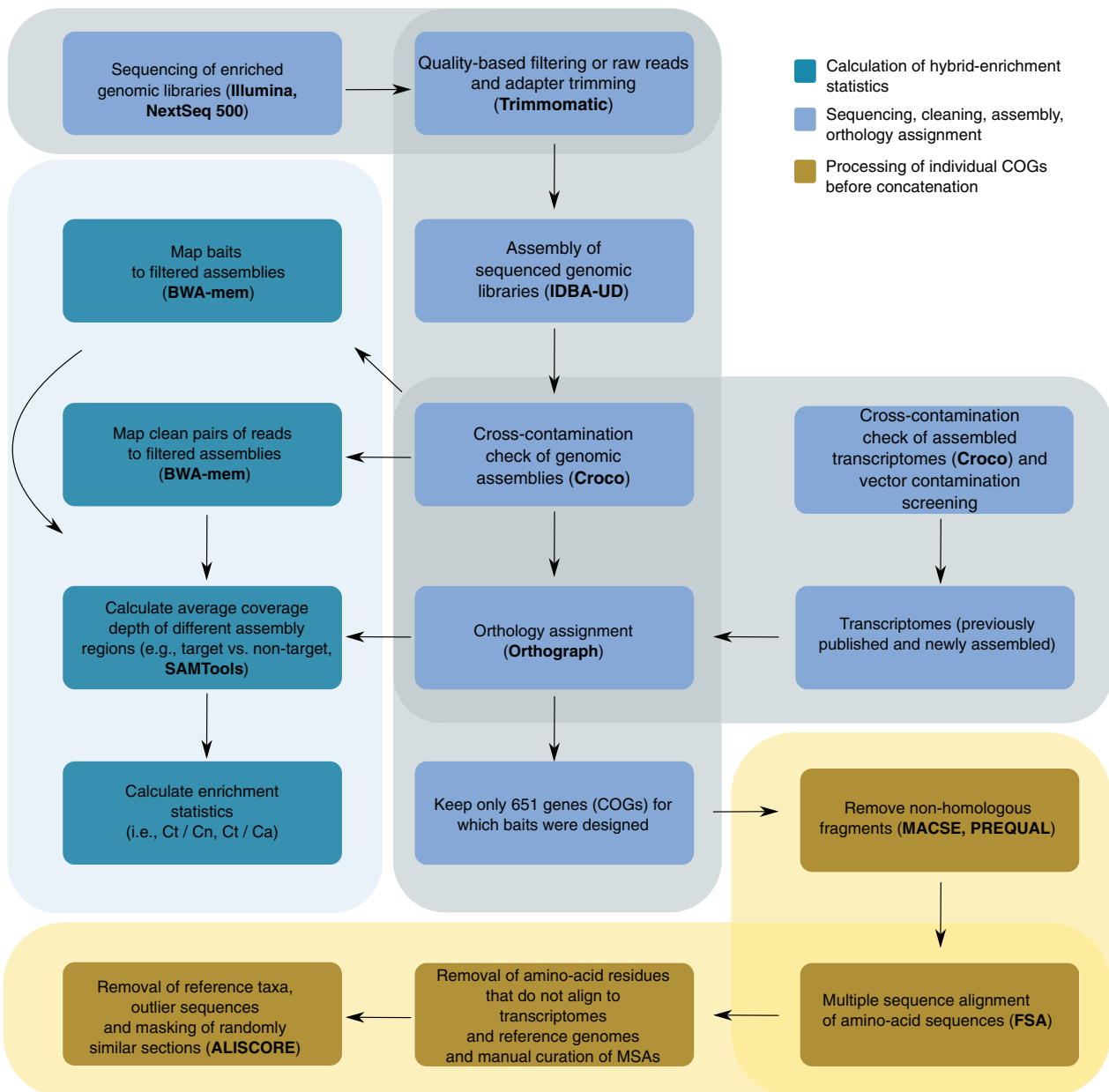


Fig. 1. Summarized workflow of the steps that were followed to sequence, clean, assemble and combine the hybrid-capture sequence data with transcriptomes and to generate individual COGs. A short workflow for calculating the hybrid-enrichment statistics is also provided.

Sequencing and assembly of the enriched genomic libraries

The enriched genomic libraries for the 95 samples of Adephaga were paired-end sequenced (150 bp) on a single flow cell of an Illumina NextSeq 500 sequencer (Illumina Inc., San Diego, CA, U.S.A., Fig. 1). Sequenced raw reads of each genomic library were trimmed to remove Illumina adapter sequences and low quality reads with Trimmomatic v. 0.38 (Bolger *et al.*, 2014, see File S2 for options). Only full pairs of trimmed reads were used for *de novo* assembly of the enriched genomic libraries (File S1: Table S2). *De novo* assembly of each genomic

library was performed with the software IDBA-UD v. 1.1.3 (see File S2 and Fig. 1) that is optimized to assemble genomic data with highly unequal coverage depth (Peng *et al.*, 2012).

Calculation of hybrid-enrichment statistics

We calculated the ratio of average per-base coverage depth of target regions (Ct) divided by the average per-base coverage depth of the nontarget regions (Ct/Cn, File S1: Table S2 and Fig. 2) as an approximate measure of the enrichment

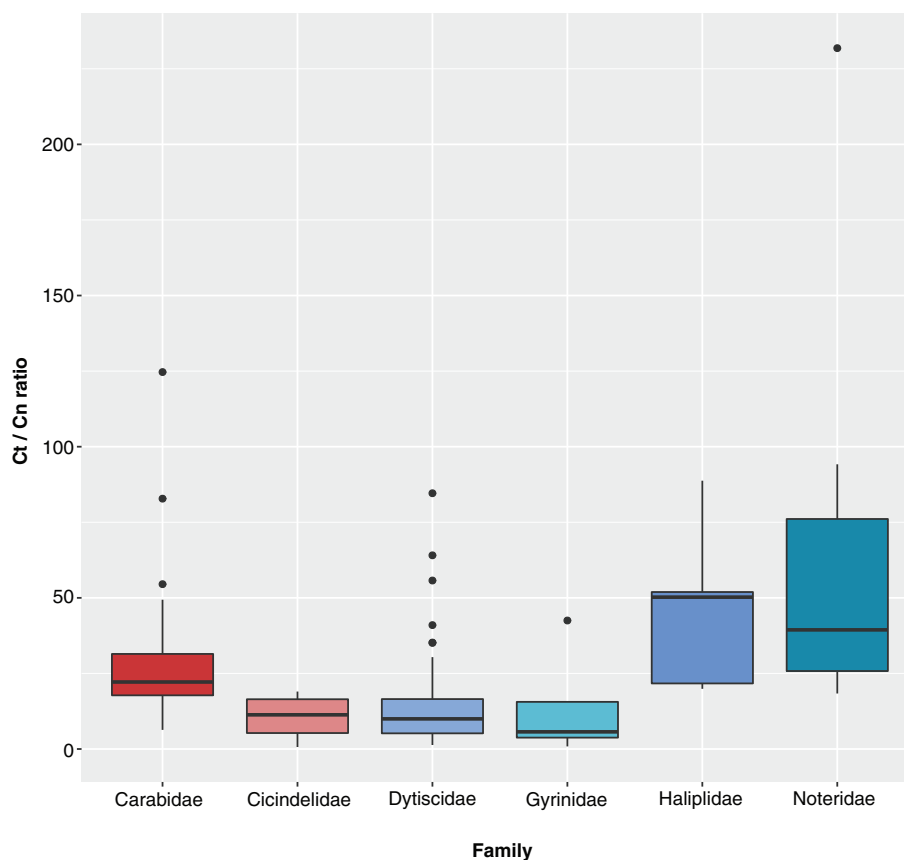


Fig. 2. Box-plots of Ct/Cn ratios inferred separately for each family of Adephaga. The plots were calculated by pooling the ratios for species of the same family into the same box-plot.

success for each genomic library in our analyses. To identify the target regions, we first identified bait-binding regions in each assembled genomic library by mapping the bait nucleotide sequences to the clean assembly files (i.e., after putative cross-contaminated contigs had been removed) using the software BWA-mem v. 0.7.17 (Li & Durbin, 2009). Subsequently, we separately mapped the trimmed reads to the assemblies with the same version of BWA-mem. A summarized file with the coverage depth of each assembly position in each assembly was generated with SAMtools v. 1.7 (Li *et al.*, 2009). We used a custom Python script and the IDs of the contigs that contained orthologous sequence (contigs assigned to any of the 651 target COGs, see below) to calculate the average per-base coverage depth of the bait-binding regions but only on those contigs that contained orthologous sequence (i.e., target regions, Ct, Fig. 1). We subsequently calculated the average per-base coverage depth of all remaining regions in the assembly for each genomic library (i.e., nontarget regions, Cn). Lastly, we calculated the average per-base coverage depth of the whole assembly for each assembled genomic library (Ca). Positions with zero coverage depth were excluded from the above calculations to avoid the inflation of enrichment statistics. We considered the statistics: Ct/Cn and Ct/Ca as approximate measures of the enrichment success for each of the 95 genomic libraries (File

S1: Table S2 and Figs 1, 2). We generated box-plots of these statistics separately for each adephagan family and performed pairwise Mann–Whitney–Wilcoxon tests between families in order to assess whether or not the values for different families were drawn from the same underlying distribution. The pairwise statistical tests were performed in R v. 3.6.3 (File S1: Table S5; R Core Team, 2020).

Cross-contamination checks and orthology assignment

Putative cross-contaminated sequences or sequences of ambiguous origin within the assembled sequence-capture data were identified with the software package CroCo v. 1.1 (Simion *et al.*, 2018). CroCo is primarily designed to screen RNA-seq data for contamination but can also potentially identify cross-contaminants from genomic data based on the assumption that the coverage of the contaminated contigs differs between the source library of contamination and the contaminated library respectively (see Simion *et al.*, 2018 and also Mayer *et al.*, 2016 for a similar approach). We considered contigs that were 99% similar over a fragment of 200 nucleotides as suspicious for cross-contamination (option: -tool K and otherwise default options). Contigs that were identified as putative contaminants

as well as those of ambiguous origin were deleted from the assemblies before downstream analyses (see File S1: Table S6 and File S2 for cross-contamination checks applied for some of the transcriptomes).

Orthology assignment of genomic fragments to each of the COGs of the ortholog set was performed with Orthograph v. 0.6.3 (Petersen *et al.*, 2017). From the 3,085 COGs of the ortholog set, we conservatively chose to analyse only the 651 COGs for which we had originally designed baits (File S1: Tables S1, S2). Orthograph-reporter script was run with the 'protein2dna' exonerate model for all hybrid-capture data (File S1: Table S2), whereas the default 'protein2genome' model was used for all transcriptomes in the dataset (File S1: Table S1, see also File S2 for additional options).

Data filtering, MSA, outlier-sequence removal and masking of randomly similar sections

The output of Orthograph could still possibly contain non-exonic residues due to random extension of open reading frames beyond the protein-coding regions (Bank *et al.*, 2017). Therefore, we followed additional procedures for filtering sequences within each COG. Specifically, we used the software MACSE v. 2.03 (Ranwez *et al.*, 2018, option: -trimNonHomologous) to remove long individual sequence fragments that shared no homology with other sequences in each COG, such as those of possibly unidentified intronic fragments (Ranwez *et al.*, 2018). The software PREQUAL v. 1.02 was subsequently used to remove shorter nonhomologous fragments such as those resulting from assembly artefacts or annotation errors (default parameters, Whelan *et al.*, 2018). These filtering steps were applied at the nucleotide sequence level, and the resulted COGs (amino-acid COGs: aaCOGs, nucleotide COGs: nCOGs) were used for further downstream filtering. We used the software FSA v. 1.15.9 (option: -fast) to infer amino-acid MSAs for each filtered COG (Bradley *et al.*, 2009). We selected the software FSA because it shows higher accuracy (i.e., lower false-positive alignment rate) than other MSA software and tends to leave nonhomologous amino-acid residues unaligned (Bradley *et al.*, 2009). By aligning the amino-acid sequences with FSA, we greatly reduced the possibility of aligning nonhomologous fragments to each other. Subsequently, we filtered the aligned aaCOGs so that amino-acid residues from hybrid-enrichment data that did not align to amino-acid residues of at least one reference species (i.e., official gene set) and at least one transcriptome were masked with an 'X'. Transcriptomic amino-acid residues that did not align to the protein-coding sequences of at least one reference taxon were also masked with an 'X'. As a last quality check, we manually curated all aligned aaCOGs to mask putative nonhomologous amino-acid fragments (see File S2). We used these filtered amino-acid alignments as a blueprint to generate corresponding codon-based nucleotide alignments with a modified version of PAL2NAL (Suyama *et al.*, 2006) as described by Misof *et al.* (2014a). A custom Python script was then used to mask all corresponding codons of the previously

masked amino acids with 'NNN'. We performed additional identification and removal of individual outlier sequences in each aligned aaCOG based on BLOSUM62 expected distances among taxa (see Dietz *et al.*, 2019 and File S2). We then removed all sequences of the reference taxa, except for the sequences of the two hymenopteran species (*Harpegnathos saltator*, *Nasonia vitripennis*) and those of *Tribolium castaneum* that we included as outgroups. Lastly, alignment sections of random similarity within each aaCOG were identified with ALIScore v. 1.2 (Misof & Misof, 2009; Kück *et al.*, 2010), as described by Vasilikopoulos *et al.* (2019), and were subsequently removed with ALICUT v. 2.31 (<https://github.com/PatrickKueck/ALiCUT>, access 16 June 2020) both at the amino-acid and the nucleotide sequence levels. The filtered and aligned aaCOGs were finally concatenated into a supermatrix with FASconCAT-G v. 1.04 (Kück & Longo, 2014).

Supermatrix evaluation and optimization for phylogenetic analyses

We opted for an informative subset of the above-described amino-acid supermatrix by using the software MARE v. 0.1.2rc and by removing partitions with an information content of zero (IC = 0, Misof *et al.*, 2013). After careful visual inspection of the resulted supermatrix (supermatrix A, Table 1) we observed that it still contained hypervariable alignment blocks. In addition, supermatrix A contained a large proportion of missing data (~50%, Table 1), which can bias phylogenetic reconstructions if missing characters are not randomly distributed (Lemmon *et al.*, 2009; Misof *et al.*, 2014b). Additionally, supermatrix A showed evidence for deviation from the assumption of stationarity, reversibility and homogeneity (SRH) as measured with the Bowker's and Stuart's tests of symmetry in SymTest v. 2.0.47 (Bowker, 1948; Stuart, 1955; Misof *et al.*, 2014a, see Table 1). Therefore, we chose to filter supermatrix A by applying strategies designed to select conserved alignment sites, reduce the degree of missing data and the potential effects of model violations in phylogenetic reconstructions (Misof *et al.*, 2001; Sharma *et al.*, 2014; Laumer *et al.*, 2019). First, we identified and removed individual gene partitions that deviate from model assumptions using the -symtest option in IQ-TREE v. 2.0.4 (Naser-Khdour *et al.*, 2019; Minh *et al.*, 2020). The resulting filtered amino-acid supermatrix was then trimmed with the software BMGE v. 1.12 (h = 0.5, amino-acid replacement matrix: BLOSUM62) to remove hypervariable alignment sites (resulting in supermatrix D). We selected the software BMGE because it selects informative sites by inferring biologically realistic variability for each column of the alignment (Crisuolo & Gribaldo, 2010; Cai *et al.*, 2020). We also generated five additional and independent amino-acid supermatrices by directly trimming supermatrix A or the partitions of supermatrix A with BMGE in order to examine the effects of progressively more aggressive filtering on the phylogenetic results (see Table 1). Additional supermatrices were generated by using three degrees of stringency (h = 0.5, h = 0.4 and h = 0.3, see Table 1 and File S3: Fig. S1).

Table 1. Summarized statistics and description for each generated and analysed amino-acid supermatrix (see File S3: Fig. S1). Saturation statistics of each supermatrix (adjusted R² and slope) based on the patristic and *p*-distances are also presented. Saturation of each supermatrix was also measured with the average pairwise lambda score (see text).

Amino-acid supermatrix ID	No. of species	No. of alignment sites	P.I. sites	Percent. (%) of P.I. sites	Average pairwise λ (lambda) score	Adjusted R ² (SHETU)	Slope (SHETU)	Adjusted R ² (SHOMU)	Slope (SHOMU)	Adjusted R ² (SHOMP)	C _a
A	136	200 017	104 221	52.1%	0.163	–	–	–	–	–	0.504
B	136	49 468	21 917	44.3%	0.118	0.425	0.126	0.486	0.213	0.479	0.831
C	136	55 521	26 220	47.2%	0.135	0.369	0.111	0.403	0.182	0.405	0.790
D ^a	136	49 797	21 401	43.0%	0.116	0.451	0.133	0.512	0.226	N/A	0.846
D - recoded ^a	136	49 797	12 699	25.5%	0.069	–	–	–	–	–	0.846
E	136	50 614	21 773	43.0%	0.116	0.454	0.133	0.515	0.227	N/A	0.846
F ^a	136	36 511	14 143	38.7%	0.095	0.510	0.155	0.569	0.256	N/A	0.882
G ^a	120	36 511	10 879	29.8%	0.079	0.396	0.230	0.393	0.272	N/A	0.880
H ^a	100	36 511	9658	26.5%	0.074	0.570	0.247	0.575	0.306	N/A	0.892
I ^a	136	29 361	11 711	39.9%	0.104	0.418	0.135	0.480	0.225	N/A	0.857
J ^a	136	23 442	7684	32.8%	0.069	0.556	0.177	0.642	0.299	N/A	0.911

Amino-acid supermatrix ID	Average <i>p</i> -dist	Median pairwise P value to the Bowker's test	Median pairwise P value to the Stuart's test	IC	Percent. (%) of pairwise P-values < 0.05. Bowker's test	Percent. (%) of pairwise P-values < 0.05. Stuart's test	Description
A	0.154	2.14E-02	7.38E-05	0.672	58.92%	82.94%	Concatenated supermatrix of masked genes with ALISCOPE after partitions with IC = 0 had been removed
B	0.111	1.07E-01	1.15E-02	0.620	37.44%	64.07%	Trimmed each gene partition of supermatrix A with BMGE, BLOSUM62, h = 0.4, keep only genes with length ≥ 50 amino-acid sites
C	0.127	9.46E-02	6.73E-03	0.599	40.10%	68.27%	Trimmed each partition of supermatrix A with BMGE, BLOSUM62, h = 0.5, keep only genes with length ≥ 80 amino-acid sites and $\leq 30\%$ missing data
D ^a	0.109	1.26E-01	1.19E-02	N/A	34.69%	64.11%	Removed genes that fail symmetry tests (IQ-TREE) from supermatrix A. Subsequently, trimmed resulting supermatrix with BMGE (h = 0.5, BLOSUM62)
D - recoded ^a	0.052	2.16E-01	–	N/A	24.67%	–	Dayhoff-6 recoded version of supermatrix D
E	0.109	1.22E-01	1.14E-02	N/A	35.02%	64.19%	Trimmed supermatrix A with BMGE, BLOSUM62, h = 0.5
F ^a	0.089	1.99E-01	4.15E-02	N/A	24.98%	51.94%	Trimmed supermatrix A with BMGE, BLOSUM62, h = 0.4
G ^a	0.074	2.27E-01	6.99E-02	N/A	20.35%	45.27%	Removed distantly related outgroup species from supermatrix F
H ^a	0.070	2.34E-01	8.53E-02	N/A	18.85%	41.92%	Removed fast evolving ingroup species (20 ingroup species with highest LB scores) from supermatrix G
I ^a	0.098	1.75E-01	4.51E-02	N/A	25.59%	50.94%	Removed 50% of genes with the highest RCFV value from supermatrix A. Trimmed resulting supermatrix with BMGE, BLOSUM62, h = 0.5
J ^a	0.065	2.96E-01	1.51E-01	N/A	13.97%	35.21%	Trimmed supermatrix A with BMGE, BLOSUM62, h = 0.3

^a Analysed under the Bayesian site-heterogeneous model CAT+GTR + G4 (BSHETU).

P.I.: parsimony informative, C_a: Overall alignment completeness scores, IC: information content (MARE), *p*-dist: observed pairwise distances, N/A: Not applicable, SHETU: site-heterogeneous unpartitioned, SHOMU: site-homogeneous unpartitioned, SHOMP: site-homogeneous partitioned.

Among-species compositional heterogeneity is a potential source of systematic error that is frequently associated with fast evolving sites (Foster, 2004; Kocot *et al.*, 2017). We generated two amino-acid supermatrices by using two different approaches for reducing among-species compositional heterogeneity (i.e., Dayhoff6-recoding and removal of genes with high relative composition frequency variation, RCFV, see Table 1 and File S2). We also tested whether the removal of distantly related

outgroup species or the removal of long-branched ingroup taxa (based on long-branch scores, LB scores, see File S2 and File S1: Table S7) affected the phylogenetic relationships.

We performed a large number of statistical tests on each generated supermatrix in order to evaluate its suitability for phylogenetic reconstruction (Table 1). First, we inferred substitution saturation plots for most analysed supermatrices (Table 1 and File S2, see Misof *et al.*, 2001; Nosenko *et al.*, 2013) by

calculating pairwise amino-acid p -distances and pairwise patristic distances. We also inferred an alternative measure of substitution saturation that is independent of the patristic distances of the inferred trees; the average lambda score for each supermatrix (i.e., λ , ranging from 0.0 to 1.0) that was recently introduced for pairs of aligned sequenced data (higher values indicate higher degree of saturation, Jermini & Misof, 2020). All pairwise λ scores in each supermatrix were calculated with the software SatuRation v. 1.0 (available from: <https://github.com/ljermini/SatuRation.v1.0>, last access: 5 January 2021, Jermini & Misof, 2020). We also measured the overall deviation from SRH conditions with the software SymTest v. 2.0.47 (current version available at <https://github.com/ottmi/symtest>, last access 20 April 2020, see Misof *et al.*, 2014a) for each filtered supermatrix and for the original supermatrix A by applying the Bowker's and Stuart's tests of symmetry (Table 1). Additionally, we calculated the overall completeness scores of the analysed supermatrices and generated heatmaps of pairwise completeness scores with AliStat v. 1.11 (Wong *et al.*, 2020, Table 1). Lastly, we screened each generated supermatrix for taxa with heterogeneous sequence divergence by generating heatmaps of pairwise mean similarity scores with ALIGROOVE v. 1.06 (Kück *et al.*, 2014).

Concatenation-based phylogenetic analyses of amino-acid supermatrices

Modelling site-specific propensities of amino acids has been shown to be more important than modelling partition-wise heterotachy in concatenation-based phylogenomic analyses (Feuda *et al.*, 2017; Wang *et al.*, 2019). In order to account for site-specific amino-acid preferences in the supermatrices, we analysed most amino-acid supermatrices under the site-heterogeneous model CAT + GTR + G4 (Bayesian site-heterogeneous model, BSJETU) using the software Phylobayes MPI v. 1.8 (Table 1, Lartillot *et al.*, 2013). Two independent MCMC chains were run for each dataset until more than 20 000 samples were collected or until convergence (File S1: Table S8).

We also analysed the amino-acid supermatrices using a maximum likelihood approach (ML) with IQ-TREE v. 1.6.12 (Nguyen *et al.*, 2015). We first selected the best-fitting substitution models in ModelFinder based on the AICc criterion on the unpartitioned matrices (File S1: Table S9; Akaike, 1974; Kalyaanamoorthy *et al.*, 2017). In order to test the relative fit of site-heterogeneous versus site-homogeneous models, we also included empirical site-heterogeneous profile mixture models in our model-selection procedure (i.e., C20, C40, C60, Quang *et al.*, 2008). In total, more than 270 models were tested on each of supermatrices B–J (unpartitioned data) except for the recoded dataset, which was only analysed with the BSJETU model. For the partitioned supermatrices (B, C, Table 1), we also calculated an optimal partitioning scheme using an edge-linked partition model using the same version of IQ-TREE (File S2, Chernomor *et al.*, 2016; Lanfear *et al.*, 2014). For these supermatrices, we assessed the relative model fit of site-homogeneous unpartitioned (SHOMU), site-homogeneous partitioned (SHOMP) and

site-heterogeneous unpartitioned (SHETU) models by using a fixed neighbour-joining tree (File S1: Table S10 and File S2). Phylogenetic tree inference was performed for each supermatrix with the SHOMU, SHETU, PMSF (posterior mean-site frequency profile model as an approximation to the C60 SHETU model, File S2, Wang *et al.*, 2018) and SHOMP models (where applicable). This was done in order to explore the extent to which using a suboptimal model affected phylogenetic reconstructions (File S1: Tables S9, S10). Statistical branch support of the inferred relationships in all concatenation-based ML analyses was estimated based on 2,000 ultrafast bootstrap (UFB) replicates (Hoang *et al.*, 2018). As a complementary measure of support, we inferred quartet-concordance scores (QC) with Quartet Sampling v. 1.3.1 (Pease *et al.*, 2018, option: -nreps 150) on the tree that resulted from the SHETU-based analysis of supermatrix D (Fig. 3). For inferring QC, we used a site-heterogeneous but less complex model than the one used to infer the tree (i.e., JTT + C20 + F + R8 instead of JTT + C60 + F + R8 due to computational limitations and using the same version of IQ-TREE, File S3: Fig. S2). Lastly, we calculated pairwise RF distances among the inferred trees under the same model (SHOMU, SHETU and PMSF) for amino-acid datasets with full taxon sampling using ETE v. 3.1.1 (Huerta-Cepas *et al.*, 2016).

Phylogenetic analyses of nucleotide sequence data

To assess the stability of phylogenetic results among analyses of different types of data, we also generated and analysed four supermatrices at the nucleotide sequence level (File S1: Table S11). Analyses of these supermatrices were performed with the same version of IQ-TREE and by selecting best-fitting SHOMP and SHOMU models (see File S2 and File S1: Table S11). We also inferred phylogenetic relationships using a model that accounts for heterotachy among sequences but has only been extensively tested in analyses of nucleotide sequence data (see File S1: Table S11, Crotty *et al.*, 2020).

Estimating alternative and confounding signals in supermatrices via four-cluster likelihood mapping and data permutations

In addition to the quartet-concordance measure, we applied the four-cluster likelihood mapping approach (FcLM, Strimmer & von Haeseler, 1997) to assess the robustness of phylogenetic results, and to measure the strength of alternative phylogenetic signals with respect to specific phylogenetic hypotheses that resulted from the analyses of supermatrix D (Fig. 3 and Table 2). The hypotheses that we tested were the following: (i) Hygrobiidae is sister to a clade of Amphizoidae + Aspidytiidae (hypothesis 1) and (ii) Cicindelidae is the sister group of Carabidae (hypothesis 2). FcLM analyses were performed on different amino-acid supermatrices that were trimmed with different degrees of stringency and were based on both SHETU and SHOMU models, in an attempt to assess whether model misspecification affected the phylogenetic signal in favour

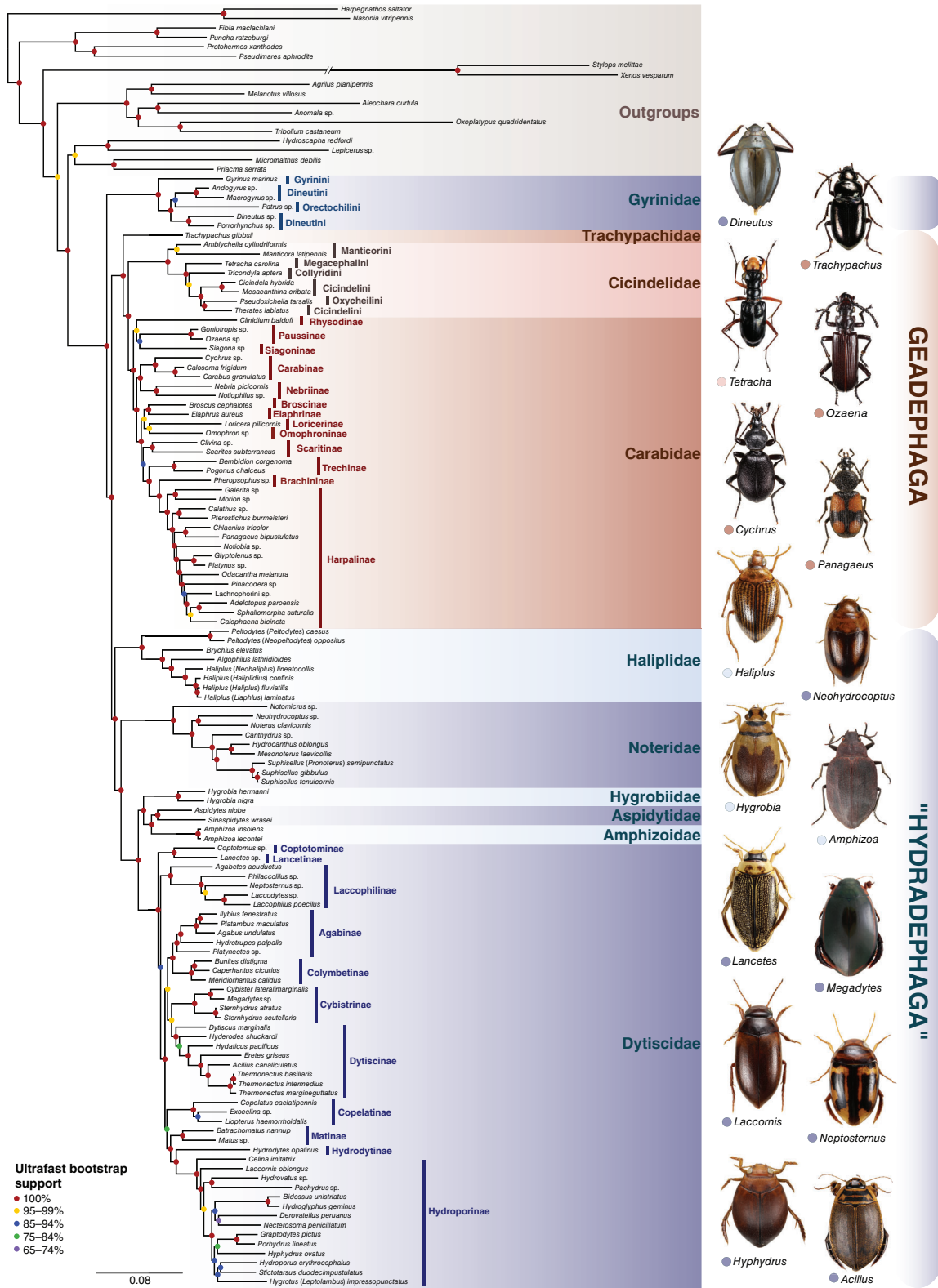


Fig. 3. Phylogenetic relationships of Adephaga as they resulted from the analysis of supermatrix D under the JTT + C60 + F + R8 site-heterogeneous model (i.e., SHETU model). Circles on tree nodes indicate branch support based on 2,000 ultrafast bootstraps (UFB). All beetle photos were provided by M. Balke.

Table 2. Detailed results of the four-cluster likelihood mapping analyses (FcLM) for the two examined phylogenetic hypotheses. Results (i.e., percentages) are shown only for the fully resolved quartets (i.e., quartets falling within the corner areas of the triangular Voronoi diagrams, see Strimmer & von Haeseler, 1997).

	SHETU model (permuted data)				SHOMU model (original data)			
	Given topology supermatrix D	Alternative topology 1 supermatrix D	Alternative topology 2 supermatrix D	Total resolved quartets (%)	Given topology supermatrix D	Alternative topology 1 supermatrix D	Alternative topology 2 supermatrix D	Total resolved quartets (%)
Hypo1 (25/296 quartets)	(Hyg. + Amp. + Asp.), (Dyt. + Rem.)	(Hyg. + Rem.), (Dyt. + Amp. + Asp.)	(Hyg. + Dyt.), (Rem. + Amp. + Asp.)	96.70%	(Hyg. + Amp. + Asp.), (Dyt. + Rem.)	(Hyg. + Rem.), (Dyt. + Amp. + Asp.)	97.70%	
Supermatrix D	59.80%	28.80%	8.10%	96.70%	65.30%	43.60%	97.70%	
Supermatrix E	58.80%	29.60%	8.30%	96.70%	64.40%	48.00%	96.20%	
Supermatrix F	53.70%	29.00%	11.70%	94.40%	61.50%	34.50%	89.50%	
Supermatrix J	44.00%	36.00%	10.60%	90.60%	51.00%	2.40%	9.60%	
Hypo2 (30/912 quartets)	(Cici. + Cara.), (Tr. + Rem.)	(Cici. + Cici.), (Tr. + Rem.)	(Tr. + Cara.), (Cici. + Rem.)	95.00%	(Cici. + Cara.), (Tr. + Rem.)	(Tr. + Cara.), (Cici. + Rem.)	96.20%	
Supermatrix D	76.50%	5.90%	12.60%	95.00%	67.20%	28.60%	96.20%	
Supermatrix E	75.90%	6.20%	12.80%	94.90%	66.80%	24.20%	95.90%	
Supermatrix F	68.10%	7.30%	16.60%	92.00%	60.90%	22.00%	88.80%	
Supermatrix J	50.20%	13.80%	21.90%	85.90%	48.30%	3.50%	17.20%	

Note: for the FcLM analyses we only included supermatrices that are comparable with respect to the effects of data trimming because they resulted from direct trimming of supermatrix A. Supermatrix D resulted from trimming a slightly different version of supermatrix A from which only 12 genes had been removed.

Abb.: Amphizoidae, Asp.: Aspidytidae, Hyg.: Hygrobiidae, Dyt.: Dytiscidae, Rem.: remaining species, Cici.: Cicindelidae, Cara.: Carabidae, Tr.: Trachypachidae, SHETU: site-heterogeneous unpartitioned, SHOMU: site-homogeneous unpartitioned.

of specific hypotheses (Table 2). In addition, FcLM analyses under the better-fitting SHETU models were performed with permutations of data (i.e., randomization of phylogenetic signal, permutation no. I, see Misof *et al.*, 2014a), in order to assess whether or not the FcLM support for a particular inferred relationship under the SHETU models resulted from misleading signal (Table 2, Misof *et al.*, 2014a).

Summary coalescent phylogenetic analyses

To explore the sensitivity of our concatenation-based analyses to the putative effects of incomplete lineage sorting (ILS), we conducted summary coalescent phylogenetic analyses (SCAs) with ASTRAL III v. 5.7.3 (Zhang *et al.*, 2018a). As SCAs are prone to gene-tree estimation errors (Mirarab *et al.*, 2016; Sayyari *et al.*, 2017) we took steps to reduce these effects on our analyses. Alignment trimming methods have been shown to be detrimental in phylogenetic inference of gene trees (Tan *et al.*, 2015) and therefore we selected the unmasked amino-acid alignments for these analyses (before trimming with ALISCORE, Fig. 1, File S3: Fig. S1). However, in order to reduce the negative effects of fragmentary sequences (Sayyari *et al.*, 2017), which are common for hybrid-capture data (Hosner *et al.*, 2016), we (i) removed alignment sites with more than or equal to 50% ambiguous characters, and then (ii) removed sequences for which more than 75% of sequence length contained ambiguous characters. Finally, we kept only genes that had a length of at least 150 amino acids and less than 50% missing data. The filtering tasks were performed with custom Perl scripts. In total, 348 filtered gene alignments were used for SCA. Gene trees were inferred after selecting the best-fitting models (SHOMU models) with the same version of IQ-TREE (see File S2). Branch support of individual gene trees was calculated based on 10 000 SH-aLRT replicates (Guindon *et al.*, 2010; Simmons & Kessenich, 2020). SCAs were then conducted with ASTRAL after collapsing weakly supported branches (<50% SH-aLRT support) with ETE v. 3.1.1.

As SCA resulted in different topologies from the concatenation analyses, we explored whether or not selecting genes with the highest levels of phylogenetic information resulted in higher congruence with the concatenation-based analyses. Potential phylogenetic information of each of the 348 filtered genes was assessed based on three criteria: (a) average SH-aLRT branch support of inferred gene trees (SH), (b) percentage of fully resolved quartets by likelihood mapping (see File S2, LM, Strimmer & von Haeseler, 1997) and (c) number of parsimony-informative sites per gene (PI). Subsets of genes with the highest scores were then obtained for downstream analyses (i.e., with values larger than the median for criteria a and b and larger or equal to the median for criterion c, Fig. 4C, D, E, F). Subsequently, SCAs were repeated for all selected subsets of genes as well as for the overlaps of genes that were selected by different approaches (Fig. 4F, G again after collapsing weakly supported branches with lower than 50% support). In order to evaluate gene-tree support for competing hypotheses

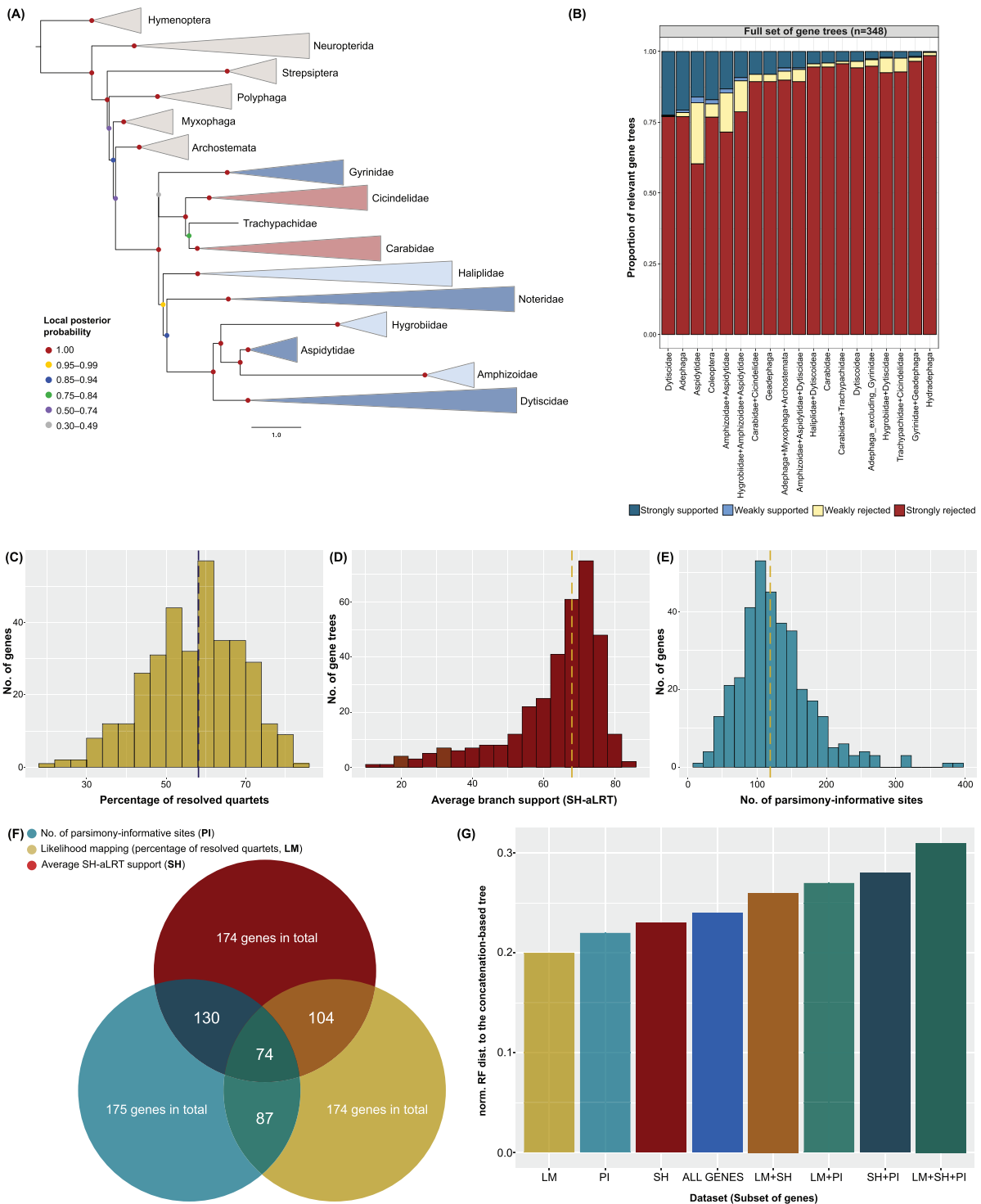


Fig. 4. Results of summary coalescent (SCA) and gene-tree discordance (GTD) analyses. (A) Summarized phylogram that resulted from the SCA analyses with all genes ($n = 348$), (B) GTD analyses showing proportion of gene trees that support or reject different relationships of Aedephaga and outgroups when the full set of gene trees was used, (C) distribution of the percentage of resolved quartets among different genes in our dataset ($n = 348$), (D) distribution of the average SH-aLRT branch support of inferred gene trees among different genes in our dataset ($n = 348$), (E) distribution of the number of parsimony-informative sites among different genes in our dataset ($n = 348$), (F) Venn diagram showing number of genes selected based on different criteria (LM, SH, PI) and number of overlapping genes among selected subsets (i.e., LM + SH = 104, LM + PI = 87, SH + PI = 130), (G) normalized RF distances of the trees inferred under different subset of genes in Fig. 4F to the concatenation-based species tree in Fig. 3. Dashed lines in histograms of Fig. 4C, D, E indicate median values.

and to assess whether or not gene-tree error might have contributed to the conflicting phylogenetic results and low branch support values in the SCA, we performed GTD analyses with DiscoVista v. 1.0 (Sayyari *et al.*, 2018). GTD was separately performed for each subset of gene trees (i.e., full set of gene trees and for the three selected subsets with higher phylogenetic information) using a branch support threshold of 70% for clades to be considered strongly accepted or rejected (Sayyari *et al.*, 2018). GTD was calculated for custom phylogenetic hypotheses but also for clades that are generally accepted based on previous analyses of molecular and morphological data (e.g., monophyletic Coleoptera, monophyletic Adephaga, a clade Amphizoidae + Aspidytidae, see Fig. 4B). We postulated that the concatenation-based tree under the better-fitting SHETU model (Fig. 3) provides a good approximation of the true familial relationships of Adephaga, because it is highly congruent with morphology-based phylogenies and latest molecular phylogenetic analyses of the group (Baca *et al.*, 2017a; Beutel *et al.*, 2020; Gustafson *et al.*, 2020). Based on that premise, we calculated normalized Robinson–Foulds (RF) distances (Robinson & Foulds, 1981) between this tree and the different species trees that resulted from the SCA analyses under various gene-subsampling strategies using ETE v. 3.1.1 and visualized the results using R v. 3.6.3. The statistical properties of the different subsets of genes (e.g., median length of genes, median number of parsimony-informative sites per gene, median proportion of missing data per gene, File S1: Table S12) were calculated either using BaCoCa v. 1.105 (RCFV values, Kück & Struck, 2014), or directly from the output of IQ-TREE (number of parsimony-informative sites), or using a custom Python script (number of sequences, gene length and missing data statistics, File S1: Table S12).

Results

Sequencing, assembly, cross-contamination check and orthology assignment for the hybrid-capture sequence data

On average, we generated 1 670 754 pairs of sequence reads per genomic library (File S1: Table S2). Quality and adapter trimming resulted in the removal of 239 210 paired reads from each sequenced genomic library on average. After assembly and cross-contamination removal, each of the clean assemblies contained 28 923 contigs on average. The summarized results of the orthology assignment for the sequence-capture data show that more than half of the 651 genes of the bait set were identified in the species of each family of Adephaga (File S3: Fig. S3, median values: Cicindelidae = 523, Carabidae = 547.5, Dytiscidae = 532, Gyrinidae = 497, Haliplidae = 596, Noteridae = 549.5). On average, 534 genes were identified in the orthology assignment step in each genomic assembly (median = 542, max. = 642, min. = 177, File S1: Table S2). Results of the orthology assignment for the transcriptomes are separately presented in File S1: Table S1 (no. of orthologous transcripts: mean = 640.5, median = 650, max. = 651, min. = 533).

Statistics of the hybrid enrichment

The results show that the overall Ct/Cn ratio is much higher than the value of one for the majority of species, which in turn suggests that the enrichment of the target regions was successful for the majority of species in our dataset (File S1: Table S2, Ct/Cn median values: Carabidae = 22.163, Noteridae = 39.414, Haliplidae = 50.231, Gyrinidae = 5.682, Cicindelidae = 11.312, Dytiscidae = 9.976, Fig. 2). The same applies for the Ct/Ca ratio (File S1: Table S2 and File S3: Fig. S4). However, the calculated statistics showed that the enrichment was potentially more successful for some adephagan families than others (Fig. 2 and File S3: Fig. S4). For example, Noteridae and Haliplidae have the best overall Ct/Cn scores that are statistically significantly greater than values for Gyrinidae, Dytiscidae and Cicindelidae (Fig. 2 and File S1: Table S5). The calculated enrichment statistics for Carabidae suggest that the enrichment was potentially more successful for this family than for the species in Cicindelidae and Dytiscidae, although not statistically different from the species of Gyrinidae, Haliplidae and Noteridae (Fig. 2 and File S1: Table S5).

Family-level phylogenetic relationships of Adephaga

Most concatenation-based analyses delivered a congruent picture on the evolution of adephagan beetles irrespective of the data type used or the model applied (Fig. 3). Specifically, a clade of Archostemata + Myxophaga as sister to Adephaga was recovered in all analyses under the best-fitting SHETU models in a ML framework (Fig. 3 and File S3: Figs S5–S12), in most BSSETU analyses of amino-acid data (File S3: Figs S13–S19) but also in the analyses of nucleotide sequence data under site-homogeneous models and models that account for heterotachy (File S3: Figs S20–S24). The family Gyrinidae was inferred as sister to all other Adephaga in all concatenation-based analyses under full taxon sampling except for the unconverged BSSETU analyses of the Dayhoff6-recoded supermatrix D (File S3: Fig. S14). Interestingly, removal of distantly related outgroups from supermatrix F (i.e., supermatrix G) without also removing long-branched ingroup taxa (i.e., supermatrix H) resulted in the equivocal placement of Gyrinidae (i.e., polytomy) under the BSSETU model, whereas SHOMU, SHOMP, SHETU and PMSF models consistently recovered Gyrinidae as sister to all other adephagan groups for these datasets (File S3: Figs S5–S44). Despite the overall agreement across concatenation-based analyses of different data types and models concerning the placement of Gyrinidae, quartet-concordance analyses did not support the monophyly of Adephaga excluding Gyrinidae (File S3: Fig. S2, QC = -0.16).

Geadephaga were consistently inferred as monophyletic and sister to Haliplidae + Dytiscoidea under concatenated analyses of different models and data types (Fig. 3 and File S3: Figs S5–S44). Within Dytiscoidea, the family Noteridae was inferred as sister to all other dytiscoid families, and Amphizoidae was inferred as sister to Aspidytidae in all concatenation-based analyses of amino acids and nucleotides

Table 3. Branch support statistics for the two most controversial clades of Adephaga under different models in all analysed supermatrices. In those cases that the clade Hygrobiidae + (Amphizoidae + Aspitytidae) was not inferred, a clade Dytiscidae + (Amphizoidae + Aspitytidae) was inferred instead (as sister to Hygrobiidae) with low branch support (i.e., lower than 95% ultrafast bootstrap support or lower than 0.95 posterior probability). In all cases that a clade Cicindelidae + Carabidae was not inferred, a clade Trachypachidae + Carabidae was recovered instead (as sister to Cicindelidae) with low branch support (i.e., lower than 95% ultrafast bootstrap support).

Dataset	No. of species	Clade									
		Hygrobiidae + (Amphizoidae + Aspitytidae)					Carabidae + Cicindelidae				
		SHOMU	SHOMP	SHETU	PMSF	BSHETU	SHOMU	SHOMP	SHETU	PMSF	BSHETU
Supermatrix B	136	70	N.I.	96	98	–	100	100	100	100	–
Supermatrix C	136	N.I.	77	96	N.I.	–	100	100	100	100	–
Supermatrix D ^a	136	96	N.A.	100	100	1.00	100	N.A.	100	100	1.00
Supermatrix E	136	94	N.A.	100	100	–	100	N.A.	100	100	–
Supermatrix F ^a	136	N.I.	N.A.	96	90	0.99	100	N.A.	100	100	1.00
Supermatrix G	120	N.I.	N.A.	93	87	N.I.	100	N.A.	100	100	1.00
Supermatrix H	100	95	N.A.	98	99	0.83	100	N.A.	100	100	1.00
Supermatrix I ^a	136	N.I.	N.A.	96	94	N.I.	100	N.A.	100	100	1.00
Supermatrix J	136	N.I.	N.A.	92	86	N.I.	N.I.	N.A.	61	N.I.	0.88

^aThe BSHETU analyses of supermatrices D and I did not reach convergence (i.e., maxdiff. > 0.30), whereas analyses of supermatrix F have reached the convergence value of maxdiff. = 0.307 (considered acceptable in our study).

N.I.: not inferred, N.A.: not applicable, SHETU: site-heterogeneous unpartitioned, SHOMP: site-homogeneous partitioned, SHOMU: site-homogeneous unpartitioned, BSHETU: Bayesian CAT+ GTR + G4 model.

(Fig. 3 and File S3: Figs S20–S24). Within Geadephaga, the monophyly of tiger beetles (Cicindelidae) and their placement as sister to monophyletic ground beetles (Carabidae) were inferred in most analyses under the site-heterogeneous models (BSHETU, SHETU, PMSF, Table 3; File S3: Figs S5–S19, S34–S42) and was also supported by analyses of nucleotide sequence data (File S3: Figs S20–S24). In contrast, a clade Trachypachidae + Carabidae was only obtained in analyses of supermatrix J and only under conditions of model misspecification (i.e., SHOMU model) or under the PMSF approximation, yet with no strong statistical branch support (Table 3; File S3: Figs S33, S42).

Concerning the inferred position of the family Hygrobiidae, all ML analyses under the better-fitting SHETU models supported a clade of Hygrobiidae + (Amphizoidae + Aspitytidae) and most of them with strong UFB support (e.g., Fig. 3). QC score also strongly supported this clade (File S3: Fig. S2, QC = 0.33). UFB support in favour this clade under SHETU models was lower when more stringent trimming criteria were applied, but the inference of this clade remained robust to the selection of dataset when a SHETU model was applied (Table 3). On the other hand, analyses under the SHOMU and SHOMP models were inconsistent regarding this hypothesis (Table 3). Specifically, SHOMU analyses of the most stringently trimmed supermatrix under full taxon sampling (supermatrix J) supported a clade Dytiscidae + (Amphizoidae + Aspitytidae) as sister to Hygrobiidae but not with strong statistical branch support (Table 3). In general, progressive trimming with more stringent criteria resulted in shift from a strongly or moderately supported Hygrobiidae + (Amphizoidae + Aspitytidae) clade (supermatrix D and E) to a poorly supported Dytiscidae + (Amphizoidae + Aspitytidae) clade (supermatrix F and J) but only in conditions of model misspecification (SHOMU models). This pattern is also observed under BSHETU model but only

for the most stringently trimmed supermatrix (supermatrix J, Table 3). Phylogenetic analyses with the PMSF approximation to the SHETU model (using a SHOMU-based guide tree with a different topology, Table 3) restored the monophyly of Hygrobiidae + (Amphizoidae + Aspitytidae) for most supermatrices (e.g., supermatrices F, G, I, J, but not for supermatrix C). This suggests that the clade Dytiscidae + (Amphizoidae + Aspitytidae) inferred under SHOMU models for these supermatrices is likely an artefact due to model misspecification. Overall, a clade that includes Amphizoidae, Aspitytidae and Dytiscidae is never strongly supported even in the few instances that it is inferred under a site-heterogeneous model (BSHETU or PMSF, Table 3; File S3: Figs S16, S18, S19, S35, S43).

Additional support for a clade Hygrobiidae + (Amphizoidae + Aspitytidae) comes from the results after removing distant outgroups and long-branched ingroup taxa from supermatrix F. Specifically, removing only distantly related outgroup taxa did not result in strong UFB support for this clade under the SHETU model (93%, Table 3) but when long-branched ingroup taxa were also removed, UFB support for the above-mentioned clade increased under the same model (98%). Additionally, the topology flipped from the clade Dytiscidae + (Amphizoidae + Aspitytidae) to the clade Hygrobiidae + (Amphizoidae + Aspitytidae) under the SHOMU and BSHETU models when long-branched ingroup species were removed (although not with strong support under the BSHETU, Table 3). This suggests that removal of distant outgroups without also accounting for branch-length heterogeneity of the ingroup might result in erroneous topology even when a site-heterogeneous model is used. Phylogenetic analyses of the Dayhoff6-recoded matrix D recovered unexpected and poorly supported clades with respect to the internal phylogeny of Dytiscoidea and more generally Adephaga (e.g., Gyrinidae + Geadephaga and Amphizoidae + Dytiscidae with low support, File S3: Fig. S14). Although the BSHETU

analyses of the recoded matrix did not reach convergence (File S3: Fig. S14 and File S1: Table S8, maxdiff = 0.49, more than 29 000 samples per MCMC chain), these observations suggest that amino-acid data-recoding might be detrimental when excessive alignment trimming and data filtering have been applied before recoding the data.

Internal phylogeny of Carabidae, Cicindelidae, Dytiscidae and Gyrinidae based on analyses of concatenated sequence data

Analyses of amino-acid and nucleotide supermatrices in a concatenation framework yielded the monophyly of all subfamilies in Dytiscidae (Fig. 3). However, phylogenetic relationships among these dytiscid subunits were unstable and not consistently resolved in all analyses except for a few cases. For instance, the subfamily Hydrodytinae was always inferred as sister to Hydroporinae with strong UFB and QC support (Fig. 3 and File S3: Fig. S2). The subfamilies Coptotominae and Lancetinae were always inferred as sister groups (Fig. 3; File S3: Figs S5–S19). In addition, all concatenation-based analyses resulted in a clade that includes all subfamilies of Dytiscidae excluding Coptotominae, Laccophilinae and Lancetinae with strong UFB and QC support (Fig. 3 and File S3: Figs S2, S5–S44). Specifically, most analyses with the SHETU models recovered Lancetinae + Coptotominae as sister to Laccophilinae + remaining Dytiscidae (Fig. 3; File S3: Figs S5–S12). In addition, most analyses of amino-acid supermatrices suggested the placement of Copelatinae as sister to a clade Matinae + (Hydrodytinae + Hydroporinae) (Fig. 3; File S3: Figs S5–S19, S25–S44). Lastly, the clades Agabinae + Colymbetinae and Cybistrinae + Dytiscinae were inferred consistently with strong support (Fig. 3; File S3: Figs S5–S44).

Concerning the internal phylogeny of Cicindelidae, the tribe Manticorini was inferred as sister to all other subfamilies in all concatenation-based analyses (Fig. 3; File S3: Figs S5–S44). This result received high QC or high UFB support across concatenation-based analyses (File S3: Figs S2, S5–S44). Although a paraphyletic Manticorini was inferred in a few instances, this result was likely an artefact due to the extremely high degree of missing data for the species *Manticora latipennis* (File S1: Table S2). The tribe Megacephalini was placed as sister to all remaining Cicindelidae except Manticorini, whereas the tribe Collyridini was inferred as sister to a clade that included Cicindelini and Oxycheilini (Fig. 3 and File S3: Figs S5–S19, S25–S44). In contrast to Cicindelidae, the internal phylogeny of the megadiverse Carabidae remained largely unstable across analyses of different supermatrices and models (File S3: Figs S5–S44). However, some relationships were robustly inferred. For instance, the subfamily Trechinae was always inferred as sister to Brachininae + monophyletic Harpalinae, whereas the subfamilies Paussinae, Rhysodinae and Siagoninae were placed in a monophyletic group close to the base of the tree of Carabidae in analyses of amino-acid supermatrices (Fig. 3 and File S3: Figs S5–S19, S25–S44). Lastly, Carabinae was inferred as

sister to Nebriinae in most phylogenetic analyses of amino-acid supermatrices (Fig. 3 and File S3: Figs S5–S19, S25–S44).

Within Gyrinidae, a strongly supported clade Dineutini + Orectochilini (as sister to Gyrinini) was inferred in different concatenation analyses of different types of data and models (Fig. 3 and File S3: Figs S5–S44). Dineutini was inferred as paraphyletic with respect to Orectochilini in analyses of amino-acid supermatrices but not always with strong UFB support (Fig. 3). Additionally, the inferred QC score did not support a paraphyletic Dineutini (QC = -0.1, File S3: Fig. S2). Analyses of nucleotide sequence data mostly suggested a monophyletic Dineutini as sister to Orectochilini, but monophyly of Dineutini was only strongly supported in one analysis of nucleotide sequence data (supermatrix D_nt, File S1: Table S11 and File S3: Fig. S24).

Comparison of different schemes of evolutionary modelling and the predictability of substitution saturation

In total, 277 models were tested on each unpartitioned amino-acid supermatrix with ModelFinder. The results show that SHETU models significantly outperformed the best SHOMU models for all supermatrices in an unpartitioned context (File S1: Table S9). All the best-fitting SHETU models included 60 categories of fixed empirical amino-acid frequencies (i.e., C60 site-heterogeneous models) suggesting that the most complex SHETU models fitted the data better even for the most stringently trimmed supermatrices (e.g., supermatrices F and J, File S1: Table S9). Comparison of the optimal partitioning schemes (SHOMP) for supermatrices B and C with the complex SHETU models showed that site-heterogeneous models (SHETU) fitted these datasets better than both partitioned and unpartitioned site-homogeneous models (SHOMP and SHOMU, File S1: Table S10). Based on the observation that SHETU models fit the data better, the inferred saturation statistics showed that using a site-homogeneous model (SHOMP or SHOMU) resulted in underestimation of the amount of substitution saturation in the amino-acid supermatrices when a measure that is dependent on patristic distances was used (i.e., adjusted R^2 , Table 1).

Stability of inferred relationships of Adephaga across analyses with different evolutionary models

We calculated all pairwise normalized RF distances among trees inferred under the same model (SHOMU, SHETU or PMSF) for those amino-acid datasets with full taxon sampling (seven trees per model, supermatrices B, C, D, E, F, I, J, Fig. 5). We assessed whether topological distances between inferred trees differ when using different evolutionary models. Although RF distances of inferred trees did not significantly differ between PMSF and SHOMU models (P value = 0.237, Mann–Whitney–Wilcoxon test with continuity correction) or between PMSF and SHETU models (P value = 0.136, Mann–Whitney–Wilcoxon test with continuity correction),

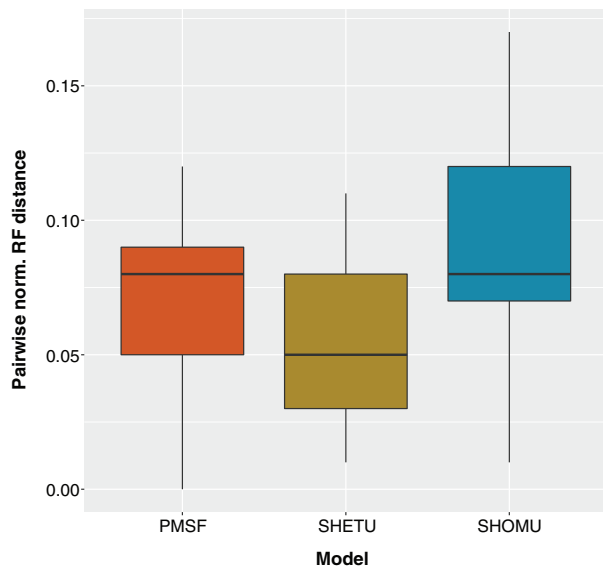


Fig. 5. Box-plots of all pairwise normalized Robinson–Foulds distances among trees that were inferred from different amino-acid supermatrices under the same type of model (normalized distances, only maximum likelihood analyses). We only included distances among trees that were inferred with full taxon sampling (i.e., supermatrices: B, C, D, E, F, I, J). SHETU: site-heterogeneous unpartitioned model, PMSF: posterior mean-site frequency profile model, SHOMU: site-homogeneous unpartitioned model.

RF distances of inferred trees were lower in analyses of SHETU models when compared with the SHOMU models (P value = 0.013, Mann–Whitney–Wilcoxon test with continuity correction, Fig. 5). This result is congruent with the consistent inference of the clade Hygrobiidae + (Amphizoidae + Aspitytidae) under SHETU models that were instead not consistently inferred under the SHOMU models, and constitutes further evidence that full site-heterogeneous empirical mixture models (C60, ML-based) result in greater stability of the inferred relationships than the less complex SHOMU models (Table 3 and Fig. 5).

Effects of removing hypervariable sites, distantly related outgroups and long-branched taxa on the statistical properties of amino-acid supermatrices

Removal of hypervariable sites had a positive impact on the statistical properties of amino-acid supermatrices in terms of eliminating potential confounding factors (Table 1). In particular, trimming the supermatrices with BMGE resulted in reduction of total and pairwise missing data (Table 1 and File S3: Figs S45–S54) and reduced deviation from SRH conditions as indicated by the reduced percentage of pairwise comparisons that failed the corresponding symmetry tests in the analysed supermatrices (Table 1, Bowker's test: 35.02%, 24.98% and 13.97% failed tests in supermatrices E, F and J respectively, see File S3: Figs S55–S65). Additionally, progressive removal of hypervariable sites resulted in progressively increasing

completeness of the supermatrices (C_a scores: 0.846, 0.882 and 0.911 for supermatrices E, F and J respectively, Table 1, File S3: Figs S45–S54). Supermatrices D and E did not significantly differ when comparing their statistical properties because only 12 genes from supermatrix A failed the symmetry tests in IQ-TREE and had therefore been removed before trimming (Table 1). Pairwise alignment similarity scores of taxa and indices for substitution saturation also improved with BMGE trimming (supermatrices D, E, F and J, Table 1 and File S3: Figs S66–S95), suggesting that progressively removing hypervariable sites results in progressively less saturated supermatrices (supermatrices D, E, F and J). The average λ scores within each supermatrix also showed that progressive removal of hypervariable sites resulted in supermatrices with less decay of potential historical signal (i.e., lower average λ scores, supermatrices D, E, F and J in Table 1). On the other hand, progressively more aggressive trimming of hypervariable sites resulted in progressive reduction of total parsimony-informative sites and reduced percentage of parsimony-informative sites (from 43.00% in supermatrix E to 32.80% in supermatrix J, Table 1). In a similar fashion, Dayhoff6-recoding resulted in removal of 40.66% of parsimony-informative sites from supermatrix D (Table 1).

Removal of distantly related outgroups from supermatrix F resulted in a less saturated supermatrix according to average λ score, whereas the linear regression of p - and patristic distances under the SHOMU and SHETU models showed reduced adjusted R^2 value (i.e., suggesting higher saturation) compared with the dataset before removing distantly related outgroups (i.e., supermatrix F). Comparisons of saturation statistics among datasets and models showed that conventional statistics of substitution saturation (R^2 and slope of regression) are dependent on the applied model (Table 1). Despite this, removal of distantly related outgroups from supermatrix F resulted in reduced proportion of failed pairwise symmetry tests (Bowker's test: 24.98%, 20.35%, 18.85% failed tests in supermatrices F, G, H respectively). Removal of long-branched ingroup taxa (see File S1: Table S7) resulted in further decrease in potential deviations from SRH conditions and also in further reduction in the degree of saturation (Bowker's test: 24.98%, 20.35%, 18.85% failed tests, λ scores: 0.095, 0.079, 0.074 in supermatrices F, G, H, respectively).

Effects of removing hypervariable sites on the branch support statistics of well-established adepagan relationships

We examined how removing hypervariable sites with BMGE using different degrees of stringency affected phylogenetic branch support for previously well-established clades of Adephaga and their outgroups. A clade that includes all adepagan families except Gyrinidae was strongly supported when using a moderate trimming strategy (supermatrices D, E, Fig. 6B) but UFB support for this relationship decreased with more aggressive trimming of the data under the SHETU and PMSF models (SHETU: 93% and 87% support in supermatrices F and J respectively, Fig. 6B). This pattern is also observed under the

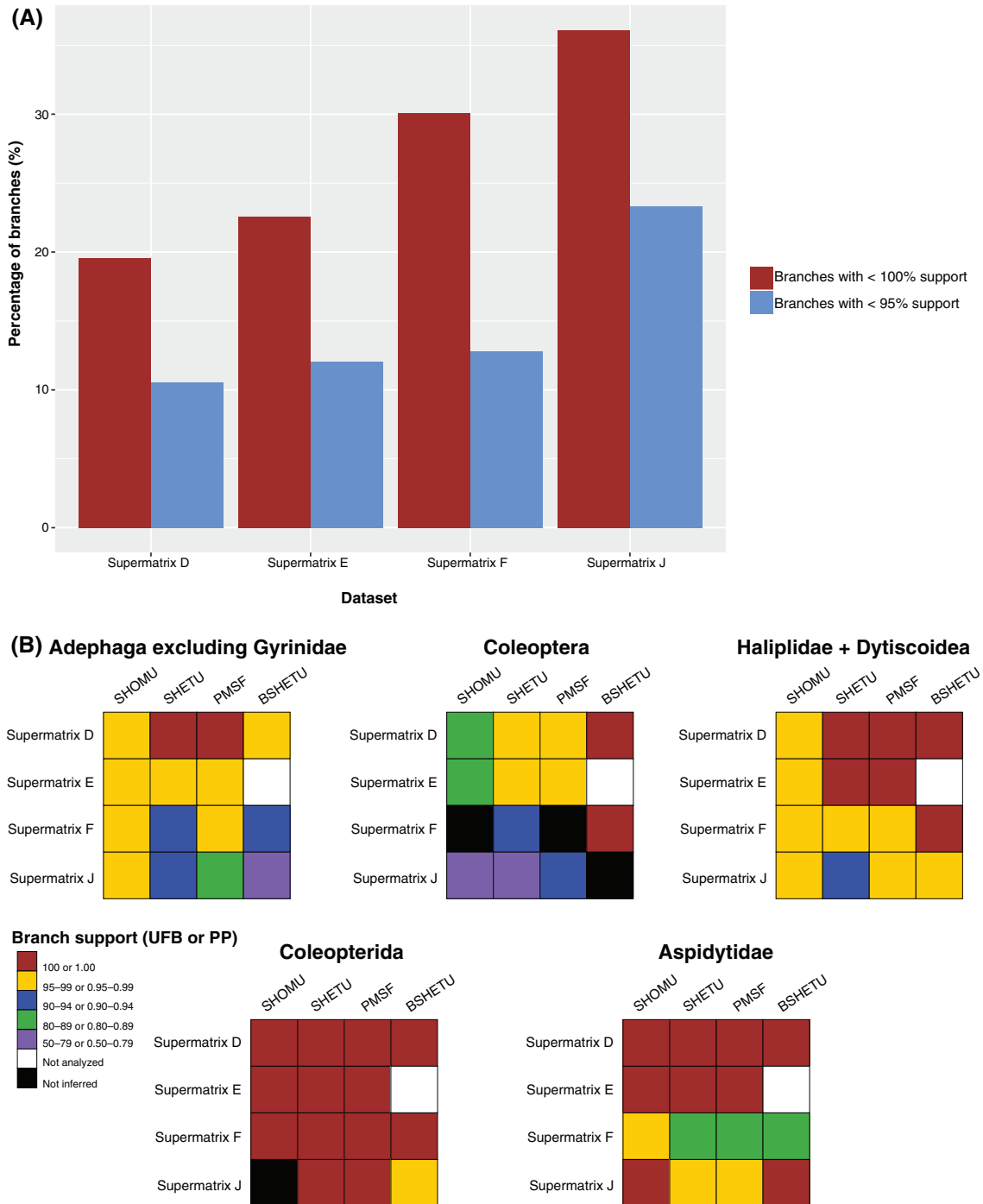


Fig. 6. Effects of removing hypervariable sites on the branch support statistics when different degrees of trimming stringency were applied (i.e., $h = 0.5$ for supermatrices D and E, $h = 0.4$ for supermatrix F, $h = 0.3$ for supermatrix J). For these comparisons we only included supermatrices that are directly comparable because they resulted from direct trimming of supermatrix A. Supermatrix D resulted from trimming a slightly different version of supermatrix A from which only 12 genes had been removed. (A) Percentage of branches with UFB support lower than 100% (red bars) and lower than 95% (blue bars) under the SHETU model in analyses of selected amino-acid supermatrices. (B) Branch support (UFB or posterior probability) for specific well-established phylogenetic clades of Adephaga and outgroups (based on morphology and other molecular phylogenetic studies) depending on the dataset that was analysed. SHOMU: site-homogeneous unpartitioned, SHETU: site-heterogeneous unpartitioned, PMSF: posterior mean-site frequency profile model, BSHETU: Bayesian site-heterogeneous CAT+GTR+G4 model. The BSHETU analyses of supermatrix D did not reach convergence ($\text{maxdiff.} = 1$), whereas BSHETU analyses of supermatrix F have reached the value of $\text{maxdiff.} = 0.307$ (considered here as marginally acceptable, see File S1: Table S8). BSHETU analyses of supermatrix J have also reached the acceptable convergence value $\text{maxdiff.} = 0.293$.

complex BSSETU model (0.94 and 0.78 posterior probability in supermatrices F and J, respectively, Fig. 6B), whereas analyses under a mis-specified model (SHOMU) still gave strong support for this relationship (98% in supermatrix J). A similar pattern is observed for the monophyly of a clade Halipilidae + Dytiscoidea, which is inferred under all models but receives lower support in the analyses of supermatrices that were trimmed more aggressively (99% and 92% UFB support in supermatrices F and J under the SHETU model respectively, Fig. 6B). Additionally, excessive trimming of the supermatrix A resulted in very low UFB support for the monophyly of Coleoptera under the better-fitting SHETU model and even resulted in nonmonophyletic beetles in cases of model misspecification (Fig. 6B, supermatrix F). The monophyly of Aspidytidae is also less well-supported in the analyses of supermatrices that were produced by very stringent trimming (supermatrices F and J, 81% and 99% respectively under the SHETU model, Fig. 6B). Lastly, trimming of the data with progressively more stringent criteria resulted in the increase of the proportion of clades that are poorly supported under the better-fitting SHETU models (total proportion of branches with <95% UFB support, Fig. 6A).

Measuring alternative and confounding signals using a combination of FcLM and data permutations

Overall, more aggressive trimming of hypervariable sites (i.e., $h = 0.4$, $h = 0.3$) resulted in a reduction of the total number of resolved quartets for the two tested hypotheses under SHOMU models and even more profoundly for the better-fitting SHETU models (Table 2). More specifically, for hypothesis 2, less than 90% of the total number of quartets were fully resolved after applying the most stringent trimming regime under the SHETU models (85.90% in the analyses of supermatrix J, Table 2). Concerning the position of Hygrobiidae (hypothesis 1) there was moderate to strong support for a clade Hygrobiidae + (Amphizoidae + Aspidytidae) in the analyses of moderately trimmed matrices (D and E, 59.80% and 58.80%) without detectable confounding signal (see FcLM of permuted data, Table 2). However, the signal in favour of this clade was reduced in the supermatrices that were trimmed more aggressively and a shift in phylogenetic support for the other two alternatives was observed (Table 2). The same pattern was observed in hypothesis 2 in which a clade Carabidae + Cicindelidae was strongly supported in the FcLM analyses of supermatrices D and E (76.50% and 75.90% respectively), whereas there was reduction in support for this hypothesis when more stringent criteria were applied (68.10% and 50.20% respectively). The absence of detectable confounding signal supporting the original results of tree reconstructions in the moderately trimmed matrices (D and E, permuted data) suggests that the shift in support from the original strongly supported hypotheses to the other two alternatives is likely not due to removal of potentially confounding signal when trimming the data but likely due to removal of genuine phylogenetic signal.

Summary coalescent phylogenetic analyses

The SCA from the analyses of all genes produced topologies that were mostly congruent with concatenation-based analyses concerning the interfamilial relationships of Adephaga and outgroups with some exceptions (Fig. 4A). Despite this, SCA resulted in weakly supported well-established clades (e.g., Coleoptera and Dytiscoidea, Fig. 4A). The clade Archostemata + Myxophaga was disrupted in most SCAs that instead resulted in a clade Archostemata + Adephaga but with low branch support (Fig. 4A and File S3: Figs S96–103). In addition, SCA did not recover Gyrinidae as sister to all other adephagan families, but the exact phylogenetic position of the family differed based on the subset of genes that were analysed (Fig. 4A and File S3: Figs S96–103). SCA with either all genes included or with an LM-based optimal subset of genes did not reject the monophyly of the families of Adephaga excluding Gyrinidae, because the branch length of the inferred clade Gyrinidae + Geadephaga was estimated as zero resulting in a polytomy (Fig. 4A and File S3: Figs S96, S97). Other conflicts between concatenation-based analyses and SCA concern a clade Trachypachidae + Carabidae with low branch support in the latter (Fig. 4A and File S3: Figs S96–S103), and also differences in relationships within adephagan families that are generally poorly supported in all SCA (File S3: Figs S96–S103). GTD analyses showed that the vast majority of the best gene trees strongly reject the tested clades of Adephaga and outgroups and the monophyly of some generally accepted groups (e.g., Coleoptera, Halipilidae + Dytiscoidea, Carabidae and Dytiscidae). In addition, considering only the most unstable interfamilial relationships of Adephaga (e.g., position of Gyrinidae, Hygrobiidae and Trachypachidae), the vast majority of relevant gene trees reject all possible alternative topologies with respect to the placement of these families, which in turn suggests that gene-tree error is prominent in the gene trees. Additionally, the distribution of potential phylogenetic signal among the sampled genes by any applied criterion (LM, PI, SH) shows that many of the genes were highly uninformative (Figs 4C, D, E; median LM = 58.14, median SH = 67.97, median PI = 119) and therefore unlikely to have produced accurate gene trees.

We tested whether choosing genes with higher potential phylogenetic information results in higher topological congruence with the concatenation-based species-tree and whether GTD for well-established hypotheses of Adephaga is reduced when applying gene-subsampling strategies. Overall, the different SCAs delivered different topologies, suggesting that the applied summary coalescent method is sensitive to the set of gene trees used (Figs 4F, G and File S3: Figs S96–S103). Selecting subsets of gene trees based on the phylogenetic information of genes (here measured with LM, PI and SH criteria, Fig. 4F) resulted in higher topological congruence with the selected concatenation-based tree (Fig. 3) than the analyses utilizing all gene trees (Fig. 4G). Nevertheless, SCA of the SH- and PI-based gene subsets failed to recover some interfamilial relationships of Adephaga (such as the placement of Gyrinidae as sister to all other Adephaga and the monophyly of Dytiscoidea; Fig. 4G and File S3: Figs S98, S99). Overall, subsampling genes

based on the LM criterion resulted in lower RF distances to the concatenation-based tree (Fig. 3) than subsampling based on the PI and SH criteria (Fig. 4G). SCA of the LM-subset of genes resulted in familial relationships identical to the SCA with all genes (Fig. 4A and File S3: Figs S96, S97) but with higher overall topological congruence to the concatenation-based tree (i.e., lower RF distance, Fig. 4G). Despite these observations, GTD analyses on the LM-, SH- and PI-selected subsets of genes showed that gene-tree error was still very prominent even for the analyses of selected gene subsets as the vast majority of gene trees strongly rejected all tested phylogenetic hypotheses similarly to the GTD analysis performed for the full set of gene trees (Fig. 4B and File S3: Figs S104–S106). It is also noteworthy that different criteria of potential phylogenetic informativeness produced different predictions on which genes are the most informative, with SH- and PI-based selected gene subsets showing a greater overlap of selected genes (Fig. 4F).

We performed SCA analyses of overlapping subsets of genes among different subsampling approaches. When overlapping sets of genes were analysed, the subsets LM + PI and LM + SH (87 and 104 genes respectively) resulted in higher topological congruence to the concatenation-based tree than the SCA of the PI+SH subset (130 genes) despite the lower number of genes analysed (Fig. 4F). These observations further support that subsampling genes based on LM may be superior to subsampling based on the other two criteria.

Discussion

A universally applicable set of DNA-hybridization baits for evolutionary studies in Adephaga

We tested the applicability of the exon-capture approach for locus recovery in many different species of Adephaga. Our orthology assignment results show that the new set of baits can be used to capture the majority of target loci in different species of the suborder. Our calculated hybrid-enrichment statistics confirm this result, as they suggest that the coverage depth of the target regions is generally higher than the coverage depth of nontarget regions. This in turn indicates that the recovery of the target loci was not due to random sequencing, but due to successful enrichment of the target loci in the species of interest. However, it should be noted that the calculated hybrid-enrichment statistics could have been potentially inflated due to inability of the assembler to include regions of low coverage, resulting in low coverage regions being potentially underrepresented in the assembly relative to high coverage regions. Despite this, we used a genomic assembler potentially robust to uneven coverage depth among different genomic regions (Peng *et al.*, 2012). In addition, potential off-target binding of baits is expected to reduce the actual differences in coverage depth, therefore balancing out the potential inflation of the calculated enrichment statistics. Hence, our calculated statistics do not provide an exact quantification of the DNA target enrichment in each species, but they rather constitute an approximate comparison of coverage depth between target and

nontarget regions in the assemblies, which was used here as a proxy for evaluating DNA target enrichment success.

Notwithstanding the success of the hybrid enrichment in all families of Adephaga, the statistics show that the baits may be more successful for enriching the target loci in some families than others. Specifically, the values of the enrichment statistics were higher for species of Noteridae, Haliplidae and Carabidae than for those in Dytiscidae, Cicindelidae and Gyrinidae. The observed differences are difficult to interpret and could be due to technical factors such as specimen quality and processing of the samples of species in some families, but also due to biological factors such as the smaller evolutionary distances among the species of these families for the genes analysed here. It should be noted that the taxon sampling in some families (e.g., Gyrinidae, Noteridae) was too small to provide conclusive evidence on the relative success of hybrid enrichment, and therefore our results should be further corroborated in future studies with increased species sampling. In summary, our results show that the new set of DNA-hybridization baits constitutes a valuable resource for future phylogenomic and potentially other evolutionary genomic studies of Adephaga. Lastly, given that Adephaga likely originated in the Permian (Gustafson *et al.*, 2017b; Beutel *et al.*, 2020), our results show that when available transcriptomic resources are sampled broadly within the clade of interest, they can be utilized to successfully infer exon-specific baits that are useful for phylogenetics at deep evolutionary timescales.

Resolving the phylogeny of Adephaga by combining transcriptomic and exon-capture sequence data

The presented interfamilial relationships of Adephaga, as inferred from our concatenation-based analyses, are generally highly congruent with the most recent phylogenomic studies based on analyses of UCEs and transcriptomes (Baca *et al.*, 2017a; McKenna *et al.*, 2019; Gustafson *et al.*, 2020). This constitutes further evidence for the phylogenetic utility of our baits at deep evolutionary timescales and helps to further consolidate the phylogeny of Adephaga. In addition, the results of analyses of concatenated data largely confirm the pattern of morphological evolution outlined by Beutel *et al.* (2020). The first split into the highly specialized surface-swimming Gyrinidae and the remaining families of Adephaga, suggested for the first time by Beutel & Roughley (1988), is well supported by transformations of larval and adults features and confirms the parphyly of ‘Hydradephaga’ as previously suggested by concatenation-based analyses of UCEs (Baca *et al.*, 2017a; Gustafson *et al.*, 2020). Recent SCA analyses that tentatively suggested monophyletic ‘Hydradephaga’ are very implausible from a morphological perspective. Moreover, strong clade support for this hypothesis was not provided (Freitas *et al.*, 2021). The reason why our quartet-concordance analyses did not support a sister group relationship between Gyrinidae and the remaining Adephaga is unknown. Given that Adephaga excluding Gyrinidae was also previously supported by phylogenetic analyses of other large genomic and morphological datasets (Baca *et al.*, 2017a; McKenna *et al.*, 2019; Beutel

et al., 2020; Gustafson *et al.*, 2020), and based on our extensive concatenation-based analyses using different models and data types, we consider the placement of Gyrinidae as sister to the rest of Adephaga as the most likely evolutionary scenario. The sampling of Gyrinidae was limited in our study as it did not include *Spanglerogyrus* and *Heterogyrus*, the sister group of the remaining family and of the large subfamily Gyrinae respectively (Miller & Bergsten, 2012; Gustafson *et al.*, 2017a,b; Beutel *et al.*, 2019b, 2017). Nevertheless, a clade comprising Orectochilini and Dineutini, previously suggested by analyses of morphological data (Beutel & Roughley, 1993; Beutel *et al.*, 2006) and by combined analyses of morphological and molecular data (Gustafson *et al.*, 2017b), was confirmed in our concatenation-based analyses (but see Miller & Bergsten, 2012 for an alternative hypothesis). The placement of Orectochilini with respect to Dineutini (i.e., sister to or nested within Dineutini) is not resolved in our study and should be investigated using more extensive taxon sampling.

All of our analyses confirm the monophyly of Geadephaga that is mainly supported by the presence of a specific protibial antenna cleaner and a dense antennal pubescence (Beutel *et al.*, 2006). The clade composed of monophyletic Carabidae and Cicindelidae is well supported by morphological apomorphies, notably by various larval features (Beutel *et al.*, 2020). This is in agreement with analyses of transcriptomes (McKenna *et al.*, 2019), but not with analyses of UCE data that suggested a clade Trachypachidae + Carabidae (Gustafson *et al.*, 2020). Here, we find that a clade Trachypachidae + Carabidae is only inferred in cases of excessive alignment trimming and suboptimal models or in SCA analyses but it is never strongly supported. The sister group relationship between Trachypachidae and the clade Cicindelidae + Carabidae indicates that a broad procoxal process and broad prothoracic postcoxal bridge are autapomorphies of tiger beetles, in addition to numerous derived features of the highly specialized ambush predatory larvae. However, the interpretation of the prothoracic features remains somewhat ambiguous, as the same supposedly derived conditions occur in the wood-associated Rhysodinae (Beutel, 1992a). Evolutionary changes in larvae and adults of Carabidae have been outlined in several studies (Beutel, 1992a,b; Dressler & Beutel, 2010). Despite this, robust molecular phylogeny with a dense taxon sampling of Carabidae is required for a solid reconstruction of the character evolution in this megadiverse lineage. Concerning the phylogeny of Cicindelidae, our inferred tribal relationships are mostly congruent to previous phylogenetic hypotheses of the family, with Mantikorini placed as sister to all other tribes (Gough *et al.*, 2019, 2020; Duran & Gough, 2020).

The placement of Haliplidae as sister to Dytiscoidea is in agreement with recent large-scale phylogenomic and morphological studies (McKenna *et al.*, 2019; Beutel *et al.*, 2020; Gustafson *et al.*, 2020), and with Beutel *et al.* (2013), a study that included extant and extinct lineages of Adephaga. Morphological support for this clade is sparse, but an important implication is that the common ancestor invaded the aquatic environment for a second time after Gyrinidae. Aquatic habits in the ground plan of Adephaga would be equally parsimonious,

but this appears unlikely given the very different adaptations of aquatic larvae in these families. The phylogenetic pattern recovered within Haliplidae is consistent with Beutel & Ruhau (1990), with *Peltodytes* placed as sister to the rest of the family, and *Brychius* as sister to all remaining Haliplidae except *Peltodytes* (see also van Vondel, 2019).

Dytiscoidea are characterized by many well-defined morphological synapomorphies (see Beutel *et al.*, 2020) and our analyses corroborate results of previous morphological and molecular analyses (McKenna *et al.*, 2019; Gustafson *et al.*, 2020). The sister group relationship between a clade Noteridae + Meruidae (note: Meruidae was not included in our study) and the remaining Dytiscoidea is robust (Beutel *et al.*, 2006; Balke *et al.*, 2008), supported for instance by elongate caudal tentorial arms and an entire series of ventral pharyngeal dilators. This concept is also corroborated by our results and by other recent phylogenomic studies (Baca *et al.*, 2017a; McKenna *et al.*, 2019; Vasilikopoulos *et al.*, 2019; Gustafson *et al.*, 2020). A placement of *Notomicrus* as sister to all other Noteridae, as inferred in our analyses, is compatible with older studies based on morphology (Beutel & Roughley, 1987; Belkaceme, 1991; Miller, 2009) and with molecular phylogenetic studies based on a few genes (Maddison *et al.*, 2009; Baca *et al.*, 2017b). However, the taxon sampling of Noteridae in our study is not sufficient for a reconstruction of the character evolution in this family.

Resolving the phylogenetic relationships of the small families of Dytiscoidea (i.e., Amphizoidae, Aspdytidae, Hygrobiidae) has proven an extremely difficult task (Vasilikopoulos *et al.*, 2019, 2021; Cai *et al.*, 2020). The notoriously difficult phylogenetic placement of these families, dating back to the Mesozoic (see Hawlitschek *et al.*, 2012 and Toussaint *et al.*, 2017 for ages of families of Dytiscoidea) may be due to the accumulation of multiple substitutions along their phylogenetic branches in combination with relatively closely spaced ancestral speciation events (see Hawlitschek *et al.*, 2012). Such phenomena are known to impede modelling of evolutionary processes (Rodríguez-Ezpeleta *et al.*, 2007; Irisarri & Meyer, 2016; Kapli *et al.*, 2020), but the validity of this hypothesis should be tested using divergence-time analyses of Adephaga that include both representatives of Aspdytidae. Concerning the placement of Hygrobiidae, our SCA analyses agree with most of our concatenation-based analyses under the best models and therefore incongruencies due to ILS do not seem very likely. Our analyses confirm the monophyly of Aspdytidae and their sister group relationship to Amphizoidae (Vasilikopoulos *et al.*, 2019; Cai *et al.*, 2020; Gustafson *et al.*, 2020), whereas most concatenated and all SCAs suggest Hygrobiidae as sister to Amphizoidae + Aspdytidae in agreement with analyses of UCE data (Gustafson *et al.*, 2020) and with analyses based on a few genes (McKenna *et al.*, 2015; Toussaint *et al.*, 2016). Despite this, the clade comprising Amphizoidae, Aspdytidae and Hygrobiidae is not supported by any solid morphological evidence so far. It implies that the reduction of the duplicatures of the metacoxal plates occurred independently in Dytiscidae and *Hygrobia*, and also the independent acquisition of prothoracic defensive glands in these groups (Beutel *et al.*, 2020). Nevertheless, Forsyth (1970) already pointed out that prothoracic glands

might have evolved independently in members of Dytiscidae and Hygrobiidae (see also Gustafson *et al.*, 2021). The presence of large and sclerotized epipharyngeal sensorial lobes is a shared feature of Dytiscidae, Aspdytidae and Amphizoidae (Dressler & Beutel, 2010). A clade Dytiscidae + (Aspdytidae + Amphizoidae) was previously inferred in analyses of specific subsets of transcriptomic data (Vasilikopoulos *et al.*, 2019) or under conditions of incomplete taxon sampling in analyses of UCE data (Baca *et al.*, 2017a; Gustafson *et al.*, 2020). In the present study, a clade Dytiscidae + (Aspdytidae + Amphizoidae) is only inferred under conditions of suboptimal models or in some analyses of stringently trimmed supermatrices under the complex BSHEU model but is never strongly supported.

As suggested by Vasilikopoulos *et al.* (2019), our analyses confirm that removing sites that deviate from the model assumptions (i.e., compositionally heterogeneous genes or hypervariable sites) result in a shift from a strongly or moderately supported Hygrobiidae + (Amphizoidae + Aspdytidae) clade to a less well-supported Dytiscidae + (Amphizoidae + Aspdytidae) clade. This change in topology is here observed only for the less-fitting site-homogeneous models (SHOMU) and for the most stringently trimmed dataset under the BSHEU model. In contrast to previous phylotranscriptomic analyses (Vasilikopoulos *et al.*, 2019), all but one phylogenetic reconstructions under the PMSF model resulted in a clade Hygrobiidae + (Amphizoidae + Aspdytidae). The same applies for most analyses under the better-fitting SHETU models, but also BSHEU models in analyses of moderately trimmed supermatrices with full taxon sampling.

A shift in phylogenetic signal between these two hypotheses is evident in FcLM analyses of both SHETU and SHOMU models when more aggressive trimming is applied. We postulate that this shift in FcLM support under the SHETU model is likely due to elimination of phylogenetic information because no biasing factors in favour of Hygrobiidae + (Amphizoidae + Aspdytidae) were detected in the quartet analyses of permuted data. Given these observations, we suggest that a clade Dytiscidae + (Aspdytidae + Amphizoidae) seems less likely and probably stems from model mis-specification or excessive data removal.

Within the Dytiscidae, we recover all subfamilies as monophyletic and also confirm the major phylogenetic patterns within the subfamilies as suggested by adult and larval morphology (Miller, 2001; Michat *et al.*, 2017), or by Sanger sequencing data combined with morphology (Miller & Bergsten, 2014; Désamoré *et al.*, 2018). Copelatinae and *Hydrotrupes* were previously assumed to be early branches in the tree of Dytiscidae based on their mandibles with open mesal grooves (Beutel, 1994). As in other molecular analyses (Ribera *et al.*, 2008; Miller & Bergsten, 2014), their separate position in the inferred trees implies that larval mandibular sucking channels are an apomorphy of Dytiscidae and were secondarily lost in these two subordinate groups. This scenario of character reversal is very likely linked to shifts in larval feeding behaviour. Furthermore, our analyses establish Coptotominae as sister to Lancetinae. This hypothesis was supported by an evaluation of larval morphology that placed the clade Coptotominae + Lancetinae as the sister to Dytiscinae + Cybistrinae (Michat *et al.*, 2017). In most

analyses under the best models, we instead recovered Coptotominae + Lancetinae as sister to all remaining diving beetles, as opposed to other studies with Matinae placed as sister to the rest of Dytiscidae (Miller, 2001; Désamoré *et al.*, 2018). The placement of Matinae as sister to all other Dytiscidae is morphologically established based on their female genital structure (Miller, 2001). Here, we retrieved Matinae nested within the family and as the sister taxon of Hydrodytinae + Hydroptorinae. In terms of morphological characters, this could possibly imply a reversal from closer metacoxal lines to more widely separated ones, and also the reversal in the case of the separated bursa copulatrix and vagina (Miller, 2001). These observations corroborate Nilsson's (1989) claim that 'dytiscid phylogeny will most probably be difficult to reconstruct, because of the widespread convergent evolution' (of morphological characters). This scenario was also discussed in detail by Michat *et al.* (2017) based on analysis of 303 larval characters. Overall, previous morphological analyses of Dytiscidae recovered the same major clades as in our study, but identified widespread character homoplasy and ambiguity along the backbone nodes of the tree.

It should be noted that we refrained from estimating divergence times of Adephaga and we instead focused on thorough exploration of phylogenetic relationships of these groups as this constitutes a crucial first step in order to reliably assign fossil calibrations for molecular dating analyses. Given that subfamilial relationships in the large groups of Adephaga (i.e., Carabidae and Dytiscidae) have proven difficult to resolve, we suggest that a thorough investigation of these intra-familial relationships should be conducted before performing divergence-time estimation analyses in future studies.

Excessive trimming of supermatrices results in reduced resolution of phylogenetic relationships

Our sensitivity analyses to remove hypervariable sites with different degrees of stringency show that there is a clear trade-off between removing sites that potentially violate the model assumptions and removing sites that contain phylogenetic information. The negative effect of excessive alignment trimming on the phylogenetic reconstructions has been demonstrated before (Talavera & Castresana, 2007; Tan *et al.*, 2015; Portik & Wiens, 2021). However, the authors of these studies examined the loss of phylogenetic information when analysing single genes or loci and not when trimming phylogenomic alignments. Our FcLM analyses for specific phylogenetic hypotheses of Adephaga show that the number of resolved quartets but also the number of parsimony-informative sites decreases in the supermatrices that are trimmed with high degree of stringency. Moreover, overall branch support in trees inferred under the better-fitting SHETU model also decreases in the analyses of the stringently trimmed datasets. Lastly, when looking at the branch support for specific hypotheses of Adephaga and outgroups, it is obvious that very stringent trimming results in poor support or even nonmonophyly of some generally accepted insect clades, such as the clades Coleoptera and Dytiscoidea

+ Haliplidae. In general, we confirm recent analyses that suggest that using BMGE for removing hypervariable sites with very stringent thresholds results in reduced phylogenetic accuracy (Steenwyk *et al.*, 2020). Low and conflicting branch support is a well-documented phenomenon in analyses of shorter MSAs and results from stochastic error (Gontcharov *et al.*, 2004; Phillips *et al.*, 2004; Delsuc *et al.*, 2005). In our study, even the most stringently trimmed supermatrices are long enough (i.e., F and J, >20 000 amino-acid sites) to be considered phylogenomic datasets, yet the proportion of well-supported clades in their inferred trees is drastically reduced in comparison to less stringently trimmed datasets. These observations suggest that a balance between removing data-driven bias and phylogenetic information should be pursued in phylogenomic analyses.

Site-heterogeneous models result in greater stability of phylogenetic relationships of Adephaga

Models that account for site-specific amino-acid propensities in the supermatrices, by incorporating heterogeneity in the amino-acid equilibrium frequencies among sites, have been shown to provide a better fit to the data than site-homogeneous models (partitioned or unpartitioned, Feuda *et al.*, 2017). Our analyses confirm these results although our model selection procedure was not performed in a Bayesian framework to include the most complex site-heterogeneous models (i.e., CAT, Lartillot & Philippe, 2004). Nevertheless, recent research shows that when the number of amino-acid equilibrium frequency categories is fixed (e.g., C60 models), the model can potentially describe heterogeneous processes in the data as well as the unconstrained CAT model (Li *et al.*, 2021). Therefore, the use of an unconstrained number of amino-acid equilibrium frequency categories in phylogenetic analyses is not justified (Li *et al.*, 2021). An interesting outcome of our study is that C60 site-heterogeneous models result in more stable phylogenetic relationships than unpartitioned site-homogeneous models. Specifically, we observed that irrespective of the inferred phylogenetic position of Hygrobiidae under SHOMU model, analyses under the SHETU model (and most analyses under the PMSF model) resulted in a clade Hygrobiidae + (Amphizoidae + Aspidytidae). In addition, comparison of the pairwise RF distances of inferred trees among different models suggests that SHETU models result in more stable phylogenetic relationships of Adephaga and are potentially less affected by the trimming or gene selection regimes. Due to computational limitations, we were not able to test this hypothesis for the CAT + GTR + G4 model as not all analyses reached convergence and given that we were not able to perform BSHETU analyses for all datasets. Nevertheless, we suggest that SHETU models may help to reduce incongruence in analyses of different amino-acid supermatrices. Lastly, we corroborate previous claims that site-homogeneous models underestimate substitution saturation (Song *et al.*, 2016; Lozano-Fernandez *et al.*, 2019) for a wide selection of amino-acid datasets and trimming regimes.

GTD analyses and locus-subsampling strategies highlight gene-tree errors in the data

GTD analyses on the complete set of loci but also on the selected subsets of loci suggest that our inferred gene trees are characterized by widespread gene-tree errors. The vast majority of gene trees strongly rejected any given well-known clade in Adephaga or in their outgroup but also any alternative phylogenetic hypothesis for the controversial groupings of Adephaga. Further indirect evidence for the extent of gene-tree errors in our dataset is provided by observing the distribution of phylogenetic information among the inferred gene trees. It is frequently assumed that GTDs are mainly due to biological factors such as ILS (Linkem *et al.*, 2016; Cloutier *et al.*, 2019). Despite this, we consider unlikely that ILS has affected all possible deep nodes in the phylogeny of Adephaga and their outgroups and therefore suggest that the observed GTD patterns are very likely due to gene-tree errors. This is more apparent when considering that our GTD analyses mostly show strongly rejected alternative phylogenetic hypotheses in the vast majority of relevant gene trees, rather than strongly supported discordance among alternative phylogenetic hypotheses. Our results confirm the views of other authors who suggest that the biasing effects of biological GTD is possible but might be less important than other biasing factors such as model misspecification and gene-tree errors at deep evolutionary timescales (Gatesy & Springer, 2014; Bryant & Hahn, 2020). Although there is no direct evidence from our analyses that the errors affect specific branches of our inferred species tree, our observations suggest that the results of the different SCAs cannot be trusted with confidence. This is further corroborated from comparing the distances of the selected concatenation-based tree to the trees inferred with SCA using different subsets of genes. These comparisons show that the coalescent method is sensitive to the set of input gene trees. It is, however, encouraging that the SCA recovered many well-established relationships of Adephaga when all genes are sampled, although some of them with low support.

It should be noted that the inability of the SCA to infer congruent results with the concatenation-based tree or strongly supported results might also be due to the small number of genes in the selected gene subsets. Specifically, we observed that species trees inferred using the four smallest subsets of genes had the highest topological distance from the concatenation-based tree. In addition, a recent study showed that the ASTRAL method can infer species trees more accurately when thousands of loci are sampled (Tilic *et al.*, 2020). In our study, we investigated whether using genes with higher phylogenetic information reduces potential gene-tree estimation error, yet the potential of increasing the accuracy of SCA by reducing systematic error has to be explored (e.g., by using empirical site-heterogeneous models, such as C10, for inferring individual gene trees Quang *et al.*, 2008). Despite this, our results show that selecting genes based on the likelihood mapping criterion may be a better approach than selecting genes based on number of parsimony-informative sites or the average branch support when aiming at reducing incongruence between SCA and

concatenation-based analyses. This result is in accordance with previous research that suggests likelihood mapping may be a good *a priori* estimator of phylogenetic informativeness (Klopfstein *et al.*, 2017).

Conclusions

We provide a new set of DNA-hybridization baits that show great promise in recovering protein-coding exons for evolutionary genomic investigations in Adephaga. Using an extensive sampling of species, by combining hybrid-capture sequence data and transcriptomes, we are able to clarify the phylogenetic relationships of the major groups such as the sister group relationship of Gyrinidae to all other families, a clade Haliplidae + Dytiscoidea, and the sister group relationship of Trachypachidae to a clade Carabidae + Cicindelidae. Furthermore, our extensive analyses under different trimming regimes and models shed light on the evolution of the families in Dytiscoidea. We show that moderate supermatrix trimming and a better-fitting site-heterogeneous model place Hygrobiidae as sister to a clade Amphizoidae + monophyletic Aspidytidae. Excessive removal of hypervariable sites using stringent trimming strategies should be avoided as it can lead to potential reduction in phylogenetic signal and reduced resolution of phylogenetic relationships. Site-heterogeneous models fit the data better but most importantly our results show that analyses with C60 site-heterogeneous models result in increased stability of inferred phylogenetic relationships of Dytiscoidea and Adephaga in general. Hence, incongruence between analyses of different subsets of amino-acid supermatrices may be ameliorated by using C60 models. Moreover, our analyses of a carefully curated set of genes suggest that gene-tree errors are prominent in the data and possibly responsible for poorly supported or incongruent species trees in SCA or for incongruent results between concatenation and SCA. Thus, our results show that scientists should take measures to eliminate or minimize gene-tree errors before attributing GTD and phylogenomic incongruence to other factors (e.g., ILS). As we have shown, a promising solution for reducing incongruence between coalescent-based and concatenation-based analyses is to select informative genes based on the likelihood mapping criterion.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

File S1. All supplementary tables (Tables S1–S12) that include: (1) an overview of transcriptomes used in the study and results of orthology assignment for transcriptomes, (2) detailed statistics of the hybrid-enrichment data including orthology assignment results, (3) results of tiling design experiments for bait design, (4) collection and voucher information for the insect samples processed in this

study, (5) Mann-Whitney-Wilcoxon tests for pairwise tests of hybrid-enrichment statistics, (6) summarized results of cross-contamination checks for the hybrid-enrichment data, (7) LB scores of species in supermatrix G, (8) convergence statistics of the Bayesian phylogenetic analyses, (9) model selection results for all amino-acid supermatrices, (10) model comparison statistics for partitioned amino-acid supermatrices based on a fixed neighbor-joining tree, (11) description of analyzed nucleotide supermatrices and their inferred statistics, (12) summarized statistics of the different subsets of genes that were used in SCA.

File S2. Supplementary experimental procedures that are not described in detail in the main materials and methods section.

File S3. All supplementary figures (Figs S1–S106) that include: (1) the summarized workflow for generating amino-acid supermatrices after the MARE step, (2) quartet-concordance results for the tree resulted from the SHETU-based analysis of supermatrix D, (3) box-plots of number of recovered loci per family for the hybrid-enrichment data, (4) box-plots of Ct/Ca ratio for each family of Adephaga, (5) all inferred phylogenetic trees with branch support values, (6) AliStat, SymTest and ALIGROOVE heatmaps for each amino-acid dataset, (7) all saturation plots inferred under different models, (8) all trees inferred from SCA and (9) GTD analyses for analyzed subsets of genes (LM, PI and SH subsets).

Acknowledgements

MB, ON, RGB and BM acknowledge the German Research Foundation (DFG) for funding the project ‘Die Integration von Phylogenomik, Sammlungsbeständen, innovativer Morphologie und umfangreicher paläontologischer Daten - Phylogenie und Evolution der Adephaga (Coleoptera) als Modellfall’ (BA 2152/24-1, NI 1387/7-1, BE 1789/11-1, MI 649/19-1). MB acknowledges support from the SNSB-Innovative scheme. We thank Wendy Moore (University of Arizona, U.S.A.), for granting access to the transcriptome of *Metrius contractus* before its official release. The authors would like to thank Claudia Etzbauer, Panagiotis Provataris, Jan Philip Oeyen, Malte Petersen and Lars Podsiadlowski for helpful discussions in various steps of the analyses. The authors would also like to thank Dr. Alexandr Prokin for providing valuable background information on fossils of Adephaga. The authors declare that there are no conflicts of interest.

Author contributions

MB, ON, RGB and BM conceived the initial project. MB, DRM, ON, RGB and BM contributed to funding acquisition. AV, MB, ON, RGB and BM designed the study. MB and LH collected and provided insect samples. AV and SK performed all

molecular laboratory experiments. AV and JMP performed assembly and cross-contamination checks for transcriptomes. SM and JMP performed NCBI sequence submissions. AV performed assembly, contamination checks and further processing of combined hybrid-enrichment data and transcriptomes. AV and KM performed phylogenetic analyses. CM provided bioinformatic methods for outlier detection. AV, MB and RGB drafted the manuscript with AV taking the lead. All authors contributed with comments in later versions of the manuscript.

Data availability statement

The datasets supporting the conclusions of this article have been deposited in the figshare digital repository (doi: <https://doi.org/10.6084/m9.figshare.14838390>, ortholog set, bait nucleotide sequences, assemblies of hybrid-enrichment data, filtered multiple sequence alignments, supermatrices, treefiles and custom scripts). New hybrid-enrichment genetic data are deposited in GenBank (Bioproject ID: PRJNA645047, see also File S1: Table S2). Raw reads for the transcriptome of *Chlaenius tricolor* have been deposited in GenBank (SRA: SRR13633634, see also File S1: Table S1). Open Access funding enabled and organized by Projekt DEAL.

References

- Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.
- Baca, S.M., Alexander, A., Gustafson, G.T. & Short, A.E.Z. (2017a) Ultraconserved elements show utility in phylogenetic inference of Adephaga (Coleoptera) and suggest paraphyly of ‘Hydradephaga’. *Systematic Entomology*, **42**, 786–795.
- Baca, S.M., Toussaint, E.F.A., Miller, K.B. & Short, A.E.Z. (2017b) Molecular phylogeny of the aquatic beetle family Noteridae (Coleoptera: Adephaga) with an emphasis on data partitioning strategies. *Molecular Phylogenetics and Evolution*, **107**, 282–292.
- Balke, M., Ribera, I. & Beutel, R.G. (2003) ASPIDYTIDAE: on the discovery of a new beetle family: detailed morphological analysis, description of a second species, and key to fossil and extant adephagan families (Coleoptera). *Water Beetles of China* (ed. by M.A. Jäch and L. Ji), pp. 53–66. Wien: Zoologisch-Botanische Gesellschaft & Wiener Coleopterologenverein.
- Balke, M., Ribera, I., Beutel, R., Vilorio, A., Garcia, M. & Vogler, A.P. (2008) Systematic placement of the recently discovered beetle family Meruidae (Coleoptera: Dytiscoidea) based on molecular data. *Zoologica Scripta*, **37**, 647–650.
- Ballesteros, J.A. & Sharma, P.P. (2019) A critical appraisal of the placement of Xiphosura (Chelicerata) with account of known sources of phylogenetic error. *Systematic Biology*, **68**, 896–917.
- Bank, S., Sann, M., Mayer, C. *et al.* (2017) Transcriptome and target DNA enrichment sequence data provide new insights into the phylogeny of vespidae wasps (Hymenoptera: Aculeata: Vespidae). *Molecular Phylogenetics and Evolution*, **116**, 213–226.
- Belkaceme, T. (1991) Skelet und Muskulatur des Kopfes und Thorax von *Noterus laevis* Sturm. Ein Beitrag zur Morphologie und Phylogenie der Noteridae (Coleoptera: Adephaga). *Stuttgarter Beiträge zur Naturkunde*, **462**, 1–94.
- Beutel, R.G. (1992a) Larval head structures of *Omoglymmius hamatus* and their implications for the relationships of Rhysodidae (Coleoptera: Adephaga). *Insect Systematics & Evolution*, **23**, 169–184.
- Beutel, R.G. (1992b) Phylogenetic analysis of thoracic structures of Carabidae (Coleoptera: Adephaga). *Journal of Zoological Systematics and Evolutionary Research*, **30**, 53–74.
- Beutel, R.G. (1994) On the systematic position of *Hydrotrupes palpalis* Sharp (Coleoptera: Dytiscidae). *Aquatic Insects*, **16**, 157–164.
- Beutel, R.G. & Roughley, R.E. (1987) On the systematic position of the genus *Notomicrus* Sharp (Hydradephaga, Coleoptera). *Canadian Journal of Zoology*, **65**, 1898–1905.
- Beutel, R.G. & Roughley, R.E. (1988) On the systematic position of the family Gyrinidae (Coleoptera: Adephaga). *Journal of Zoological Systematics and Evolutionary Research*, **26**, 380–400.
- Beutel, R.G. & Roughley, R.E. (1993) Phylogenetic analysis of Gyrinidae based on characters of the larval head (Coleoptera: Adephaga). *Insect Systematics & Evolution*, **24**, 459–468.
- Beutel, R.G. & Ruhnau, S. (1990) Phylogenetic analysis of the genera of Haliplidae (Coleoptera) based on characters of adults. *Aquatic Insects*, **12**, 1–17.
- Beutel, R.G., Balke, M. & Steiner, W.E. (2006) The systematic position of Meruidae (Coleoptera, Adephaga) and the phylogeny of the smaller aquatic adephagan beetle families. *Cladistics*, **22**, 102–131.
- Beutel, R.G., Wang, B., Tan, J.J., Ge, S.Q., Ren, D. & Yang, X.K. (2013) On the phylogeny and evolution of Mesozoic and extant lineages of Adephaga (Coleoptera, Insecta). *Cladistics*, **29**, 147–165.
- Beutel, R.G., Yan, E., Richter, A., Büsse, S., Miller, K.B., Yavorskaya, M. & Wipfler, B. (2017) The head of *Heterogyrus milloti* (Coleoptera: Gyrinidae) and its phylogenetic implications. *Arthropod Systematics and Phylogeny*, **75**, 261–280.
- Beutel, R.G., Pohl, H., Yan, E.V. *et al.* (2019a) The phylogeny of Coleopterida (Hexapoda) – morphological characters and molecular phylogenies. *Systematic Entomology*, **44**, 75–102.
- Beutel, R.G., Yan, E., Yavorskaya, M., Büsse, S., Gorb, S.N. & Wipfler, B. (2019b) On the thoracic anatomy of the Madagascan *Heterogyrus milloti* and the phylogeny of Gyrinidae (Coleoptera). *Systematic Entomology*, **44**, 336–360.
- Beutel, R.G., Ribera, I., Fikáček, M., Vasilikopoulos, A., Misof, B. & Balke, M. (2020) The morphological evolution of the Adephaga (Coleoptera). *Systematic Entomology*, **45**, 378–395.
- Bi, K., Vanderpool, D., Singhal, S., Linderroth, T., Moritz, C. & Good, J.M. (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*, **13**, 403.
- Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
- Bowker, A.H. (1948) A test for symmetry in contingency tables. *Journal of the American Statistical Association*, **43**, 572–574.
- Bradley, R.K., Roberts, A., Smoot, M. *et al.* (2009) Fast statistical alignment. *PLoS Computational Biology*, **5**, e1000392.
- Bragg, J.G., Potter, S., Bi, K. & Moritz, C. (2016) Exon capture phylogenomics: efficacy across scales of divergence. *Molecular Ecology Resources*, **16**, 1059–1068.
- Bryant, D. & Hahn, M.W. (2020) The concatenation question. *Phylogenetics in the genomic era* (ed. by C. Scornavacca, F. Delsuc and N. Galtier), pp. 3.4:1–3.4:23. <https://hal.archives-ouvertes.fr/hal-02535651>.
- Burmeister, E.-G. (1976) Der Ovipositor der Hydradephaga (Coleoptera) und seine phylogenetische Bedeutung unter besonderer Berücksichtigung der Dytiscidae. *Zoomorphologie*, **85**, 165–257.
- Cai, C., Tihelka, E., Pisani, D. & Donoghue, P.C.J. (2020) Data curation and modeling of compositional heterogeneity in insect phylogenomics: a case study of the phylogeny of Dytiscoidea (Coleoptera:

- Adephaga). *Molecular Phylogenetics and Evolution*, **147**, 106782.
- Chernomor, O., Von Haeseler, A. & Minh, B.Q. (2016) Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, **65**, 997–1008.
- Cloutier, A., Sackton, T.B., Grayson, P., Clamp, M., Baker, A.J. & Edwards, S.V. (2019) Whole-genome analyses resolve the phylogeny of flightless birds (Palaeognathae) in the presence of an empirical anomaly zone. *Systematic Biology*, **68**, 937–955.
- Criscuolo, A. & Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, **10**, 210.
- Crotty, S.M., Minh, B.Q., Bean, N.G., Holland, B.R., Tuke, J., Jermiin, L.S. & Von Haeseler, A. (2020) GHOST: recovering historical signal from heterotachously evolved sequence alignments. *Systematic Biology*, **69**, 249–264.
- Crowson, R.A. (1960) The phylogeny of Coleoptera. *Annual Review of Entomology*, **5**, 111–134.
- Delsuc, F., Brinkmann, H. & Philippe, H. (2005) Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*, **6**, 361–375.
- Désamoré, A., Laenen, B., Miller, K.B. & Bergsten, J. (2018) Early burst in body size evolution is uncoupled from species diversification in diving beetles (Dytiscidae). *Molecular Ecology*, **27**, 979–993.
- Dietz, L., Dömel, J.S., Leese, F., Mahon, A.R. & Mayer, C. (2019) Phylogenomics of the longitarsal Colossendeidae: the evolutionary history of an Antarctic sea spider radiation. *Molecular Phylogenetics and Evolution*, **136**, 206–214.
- Dressler, C. & Beutel, R.G. (2010) The morphology and evolution of the adult head of Adephaga (Insecta: Coleoptera). *Arthropod Systematics and Phylogeny*, **68**, 239–287.
- Duran, D.P. & Gough, H.M. (2020) Validation of tiger beetles as distinct family (Coleoptera: Cicindelidae), review and reclassification of tribal relationships. *Systematic Entomology*, **45**, 723–729.
- Evangelista, D., Simon, S., Wilson, M.M. *et al.* (2021) Assessing support for Blaberoidea phylogeny suggests optimal locus quality. *Systematic Entomology*, **46**, 157–171.
- Faircloth, B.C., McCormack, J.E., Crawford, N.G., Harvey, M.G., Brumfield, R.T. & Glenn, T.C. (2012) Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, **61**, 717–726.
- Feuda, R., Dohrmann, M., Pett, W. *et al.* (2017) Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Current Biology*, **27**, 3864–3870.
- Forsyth, D.J. (1970) The structure of the defence glands of the Cicindelidae, Amphizoidae, and Hydrobiidae (Insecta: Coleoptera). *Journal of Zoology*, **160**, 51–69.
- Foster, P. (2004) Modeling compositional heterogeneity. *Systematic Biology*, **53**, 485–495.
- Freitas, F.V., Branstetter, M.G., Griswold, T. & Almeida, E.A.B. (2021) Partitioned gene-tree analyses and gene-based topology testing help resolve incongruence in a phylogenomic study of host-specialist bees (Apidae: Eucerinae). *Molecular Biology and Evolution*, **38**, 1090–1100.
- Gatesy, J. & Springer, M.S. (2014) Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalence conundrum. *Molecular Phylogenetics and Evolution*, **80**, 231–266.
- Gontcharov, A.A., Marin, B. & Melkonian, M. (2004) Are combined analyses better than single gene phylogenies? A case study using SSU rDNA and rbcL sequence comparisons in the Zygnematomyceae (Streptophyta). *Molecular Biology and Evolution*, **21**, 612–624.
- Gough, H.M., Duran, D.P., Kawahara, A.Y. & Toussaint, E.F.A. (2019) A comprehensive molecular phylogeny of tiger beetles (Coleoptera, Carabidae, Cicindelinae). *Systematic Entomology*, **44**, 305–321.
- Gough, H.M., Allen, J.M., Toussaint, E.F.A., Storer, C.G. & Kawahara, A.Y. (2020) Transcriptomics illuminate the phylogenetic backbone of tiger beetles. *Biological Journal of the Linnean Society*, **129**, 740–751.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, **59**, 307–321.
- Gustafson, G.T., Bergsten, J., Ranarilalaitiana, T., Randriamihaja, J.H. & Miller, K.B. (2017a) The morphology and behavior of the endemic Malagasy whirligig beetle *Heterogyryus milloti* Legros, 1953 (Coleoptera: Gyrinidae: Heterogyryinae). *The Coleopterists Bulletin*, **71**, 315–328.
- Gustafson, G.T., Prokin, A.A., Bukontaite, R., Bergsten, J. & Miller, K.B. (2017b) Tip-dated phylogeny of whirligig beetles reveals ancient lineage surviving on Madagascar. *Scientific Reports*, **7**, 8619.
- Gustafson, G.T., Alexander, A., Sproul, J.S., Pflug, J.M., Maddison, D.R. & Short, A.E.Z. (2019) Ultraconserved element (UCE) probe set design: base genome and initial design parameters critical for optimization. *Ecology and Evolution*, **9**, 6933–6948.
- Gustafson, G.T., Baca, S.M., Alexander, A.M. & Short, A.E.Z. (2020) Phylogenomic analysis of the beetle suborder Adephaga with comparison of tailored and generalized ultraconserved element probe performance. *Systematic Entomology*, **45**, 552–570.
- Gustafson, G.T., Miller, K.B., Michat, M.C., Alarie, Y., Baca, S.M., Balke, M. & Short, A.E.Z. (2021) The enduring value of reciprocal illumination in the era of insect phylogenomics: a response to Cai *et al.* (2020). *Systematic Entomology*, **46**, 473–486.
- Hawlitcschek, O., Hendrich, L. & Balke, M. (2012) Molecular phylogeny of the squeak beetles, a family with disjunct Palearctic-Australian range. *Molecular Phylogenetics and Evolution*, **62**, 550–554.
- Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q. & Le, S.V. (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, **35**, 518–522.
- Hosner, P.A., Faircloth, B.C., Glenn, T.C., Braun, E.L. & Kimball, R.T. (2016) Avoiding missing data biases in phylogenomic inference: an empirical study in the landfowl (Aves: Galliformes). *Molecular Biology and Evolution*, **33**, 1110–1125.
- Huerta-Cepas, J., Serra, F. & Bork, P. (2016) ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Molecular Biology and Evolution*, **33**, 1635–1638.
- Hunt, T., Bergsten, J., Levkanicova, Z. *et al.* (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**, 1913–1916.
- Irisarri, I. & Meyer, A. (2016) The identification of the closest living relative(s) of tetrapods: phylogenomic lessons for resolving short ancient internodes. *Systematic Biology*, **65**, 1057–1075.
- Jäch, M.A. & Balke, M. (2008) Global diversity of water beetles (Coleoptera) in freshwater. *Hydrobiologia*, **595**, 419–442.
- Jermin, L.S. & Misof, B. (2020) Measuring historical and compositional signals in phylogenetic data. *bioRxiv*. <https://doi.org/10.1101/2020.01.03.894097>.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**, 587–589.
- Kapli, P. & Telford, M.J. (2020) Topology dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha. *Science Advances*, **6**, eabc5162.

- Kapli, P., Yang, Z. & Telford, M.J. (2020) Phylogenetic tree building in the genomic age. *Nature Reviews Genetics*, **21**, 428–444.
- Karin, B.R., Gamble, T., Jackman, T.R. & Vidal, N. (2020) Optimizing phylogenomics with rapidly evolving long exons: comparison with anchored hybrid enrichment and ultraconserved elements. *Molecular Biology and Evolution*, **37**, 904–922.
- Kavanaugh, D.H. (1986) A systematic review of amphizoid beetles (Amphizoidae: Coleoptera) and their phylogenetic relationships to other Adephaga. *Proceedings of the California Academy of Sciences*, **44**, 67–109.
- Klopfstein, S., Massingham, T. & Goldman, N. (2017) More on the best evolutionary rate for phylogenetic analysis. *Systematic Biology*, **66**, 769–785.
- Kocot, K.M., Struck, T.H., Merkel, J. *et al.* (2017) Phylogenomics of Lophotrochozoa with consideration of systematic error. *Systematic Biology*, **66**, 256–282.
- Kück, P. & Longo, G.C. (2014) FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology*, **11**, 81.
- Kück, P. & Struck, T.H. (2014) BaCoCa - a heuristic software tool for the parallel assessment of sequence biases in hundreds of gene and taxon partitions. *Molecular Phylogenetics and Evolution*, **70**, 94–98.
- 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.
- Kück, P., Meid, S.A., Groß, C., Wägele, J.W. & Misof, B. (2014) AliGROOVE - visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support. *BMC Bioinformatics*, **15**, 294.
- Lanfear, R., Calcott, B., Kainer, D., Mayer, C. & Stamatakis, A. (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, **14**, 82.
- Lartillot, N. & Philippe, H. (2004) A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Molecular Biology and Evolution*, **21**, 1095–1109.
- Lartillot, N., Brinkmann, H. & Philippe, H. (2007) Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. *BMC Evolutionary Biology*, **7**, S4.
- Lartillot, N., Rodrigue, N., Stubbs, D. & Richer, J. (2013) PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Systematic Biology*, **62**, 611–615.
- Laumer, C.E., Fernández, R., Lemer, S. *et al.* (2019) Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proceedings of the Royal Society B: Biological Sciences*, **286**, 20190831.
- Lawrence, J.F. & Newton, A.F. JR (1982) Evolution and classification of beetles. *Annual Review of Ecology and Systematics*, **13**, 261–290.
- 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.
- Lemmon, E.M. & Lemmon, A.R. (2013) High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 99–121.
- Lemmon, A.R., Brown, J.M., Stanger-Hall, K. & Lemmon, E.M. (2009) The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. *Systematic Biology*, **58**, 130–145.
- Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, **61**, 727–744.
- Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
- Li, H., Handsaker, B., Wysoker, A. *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
- Li, Y., Shen, X.-X., Evans, B., Dunn, C.W. & Rokas, A. (2021) Rooting the animal tree of life. *Molecular Biology and Evolution*. <https://doi.org/10.1093/molbev/msab170>.
- Linkem, C.W., Minin, V.N. & Leaché, A.D. (2016) Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). *Systematic Biology*, **65**, 465–477.
- López-López, A. & Vogler, A.P. (2017) The mitogenome phylogeny of Adephaga (Coleoptera). *Molecular Phylogenetics and Evolution*, **114**, 166–174.
- Lorenz, W. (2020) *CarabCat: global database of ground beetles (version Oct 2017)* [WWW document]. Species 2000 ITIS Cat. Life, 2020-12-01. URL <https://www.catalogueoflife.org/> [accessed on 12 November 2020].
- Lozano-Fernandez, J., Tanner, A.R., Giacomelli, M., Carton, R., Vinther, J., Edgecombe, G.D. & Pisani, D. (2019) Increasing species sampling in chelicerate genomic-scale datasets provides support for monophyly of Acari and Arachnida. *Nature Communications*, **10**, 2295.
- Maddison, D.R., Moore, W., Baker, M.D., Ellis, T.M., Ober, K.A., Cannone, J.J. & Gutell, R.R. (2009) Monophyly of terrestrial adephagan beetles as indicated by three nuclear genes (Coleoptera: Carabidae and Trachypachidae). *Zoologica Scripta*, **38**, 43–62.
- Mayer, C., Sann, M., Donath, A. *et al.* (2016) BaitFisher: a software package for multispecies target DNA enrichment probe design. *Molecular Biology and Evolution*, **33**, 1875–1886.
- Mayer, C., Dietz, L., Call, E., Kukowka, S., Martin, S. & Espeland, M. (2021) Adding leaves to the Lepidoptera tree: capturing hundreds of nuclear genes from old museum specimens. *Systematic Entomology*, **46**, 649–671.
- McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**, 526–538.
- McKenna, D.D., Wild, A.L., Kanda, K. *et al.* (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.
- McKenna, D.D., Shin, S., Ahrens, D. *et al.* (2019) The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences*, **116**, 24729–24737.
- Michat, M.C., Alarie, Y. & Miller, K.B. (2017) Higher-level phylogeny of diving beetles (Coleoptera: Dytiscidae) based on larval characters. *Systematic Entomology*, **42**, 734–767.
- Miller, K.B. (2001) On the phylogeny of the Dytiscidae (Insecta: Coleoptera) with emphasis on the morphology of the female reproductive system. *Insect Systematics & Evolution*, **32**, 45–92.
- Miller, K.B. (2009) On the systematics of Noteridae (Coleoptera: Adephaga: Hydradephaga): phylogeny, description of a new tribe, genus and species, and survey of female genital morphology. *Systematics and Biodiversity*, **7**, 191–214.
- Miller, K.B. & Bergsten, J. (2012) Phylogeny and classification of whirligig beetles (Coleoptera: Gyrinidae): relaxed-clock model outperforms parsimony and time-free Bayesian analyses. *Systematic Entomology*, **37**, 706–746.
- Miller, K.B. & Bergsten, J. (2014) The phylogeny and classification of predaceous diving beetles. *Ecology, systematics, and the natural history of predaceous diving beetles (Coleoptera: Dytiscidae)* (ed. by D. Yee), pp. 49–172. Springer, Dordrecht.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. & Lanfear, R. (2020) IQ-TREE 2: new models

- and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, **37**, 1530–1534.
- Mirarab, S., Bayzid, M.S. & Warnow, T. (2016) Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. *Systematic Biology*, **65**, 366–380.
- 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., Rickert, A.M., Buckley, T.R., Fleck, G. & Sauer, K.P. (2001) Phylogenetic signal and its decay in mitochondrial SSU and LSU rRNA gene fragments of Anisoptera. *Molecular Biology and Evolution*, **18**, 27–37.
- Misof, B., Meyer, B., von Reumont, B.M., Kück, P., Misof, K. & Meusemann, K. (2013) Selecting informative subsets of sparse supermatrices increases the chance to find correct trees. *BMC Bioinformatics*, **14**, 348.
- Misof, B., Liu, S., Meusemann, K. *et al.* (2014a) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, **346**, 763–767.
- Misof, B., Meusemann, K., von Reumont, B.M., Kück, P., Prohaska, S.J. & Stadler, P.F. (2014b) *A priori* assessment of data quality in molecular phylogenetics. *Algorithms for Molecular Biology*, **9**, 22.
- Mongiardino Koch, N. & Thompson, J.R. (2021) A total-evidence dated phylogeny of Echinoidea combining phylogenomic and paleontological data. *Systematic Biology*, **70**, 421–439.
- Naser-Khdour, S., Minh, B.Q., Zhang, W., Stone, E.A., Lanfear, R. & Bryant, D. (2019) The prevalence and impact of model violations in phylogenetic analysis. *Genome Biology and Evolution*, **11**, 3341–3352.
- Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274.
- Nilsson, A.N. (1989) On the genus *Agabetes* Crotch (Coleoptera, Dytiscidae), with a new species from Iran. *Annales Entomologici Fennici*, **55**, 35–40.
- Nosenko, T., Schreiber, F., Adamska, M. *et al.* (2013) Deep metazoan phylogeny: when different genes tell different stories. *Molecular Phylogenetics and Evolution*, **67**, 223–233.
- Pease, J.B., Brown, J.W., Walker, J.F., Hinchliff, C.E. & Smith, S.A. (2018) Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, **105**, 385–403.
- Peng, Y., Leung, H.C.M., Yiu, S.M. & Chin, F.Y.L. (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, **28**, 1420–1428.
- Petersen, M., Meusemann, K., Donath, A. *et al.* (2017) Orthograph: a versatile tool for mapping coding nucleotide sequences to clusters of orthologous genes. *BMC Bioinformatics*, **18**, 111.
- Philippe, H., de Vienne, D.M., Ranwez, V., Roure, B., Baurain, D. & Delsuc, F. (2017) Pitfalls in supermatrix phylogenomics. *European Journal of Taxonomy*, **283**, 1–25.
- Phillips, M.J., Delsuc, F. & Penny, D. (2004) Genome-scale phylogeny and the detection of systematic biases. *Molecular Biology and Evolution*, **21**, 1455–1458.
- Portik, D.M. & Wiens, J.J. (2021) Do alignment and trimming methods matter for phylogenomic (UCE) analyses? *Systematic Biology*, **70**, 440–462.
- Quang, L.S., Gascuel, O. & Lartillot, N. (2008) Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics*, **24**, 2317–2323.
- R Core Team (2020) *R: A Language and Environment for Statistical Computing*.
- Ranwez, V., Douzery, E.J.P., Cambon, C., Chantret, N. & Delsuc, F. (2018) MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Molecular Biology and Evolution*, **35**, 2582–2584.
- Ribera, I., Hogan, J.E. & Vogler, A.P. (2002) Phylogeny of hydrade-phagan water beetles inferred from 18S rRNA sequences. *Molecular Phylogenetics and Evolution*, **23**, 43–62.
- Ribera, I., Vogler, A.P. & Balke, M. (2008) Phylogeny and diversification of diving beetles (Coleoptera: Dytiscidae). *Cladistics*, **24**, 563–590.
- Robinson, D.F. & Foulds, L.R. (1981) Comparison of phylogenetic trees. *Mathematical Biosciences*, **53**, 131–147.
- Rodríguez-Ezpeleta, N., Brinkmann, H., Roure, B., Lartillot, N., Lang, B.F. & Philippe, H. (2007) Detecting and overcoming systematic errors in genome-scale phylogenies. *Systematic Biology*, **56**, 389–399.
- Sann, M., Niehuis, O., Peters, R.S. *et al.* (2018) Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. *BMC Evolutionary Biology*, **18**, 71.
- Sayyari, E., Whitfield, J.B. & Mirarab, S. (2017) Fragmentary gene sequences negatively impact gene tree and species tree reconstruction. *Molecular Biology and Evolution*, **34**, 3279–3291.
- Sayyari, E., Whitfield, J.B. & Mirarab, S. (2018) DiscoVista: interpretable visualizations of gene tree discordance. *Molecular Phylogenetics and Evolution*, **122**, 110–115.
- 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.
- Short, A.E.Z. (2018) Systematics of aquatic beetles (Coleoptera): current state and future directions. *Systematic Entomology*, **43**, 1–18.
- Shull, V.L., Vogler, A.P., Baker, M.D., Maddison, D.R. & Hammond, P.M. (2001) Sequence alignment of 18S ribosomal RNA and the basal relationships of adephagan beetles: evidence for monophyly of aquatic families and the placement of Trachypachidae. *Systematic Biology*, **50**, 945–969.
- Simion, P., Belkhir, K., François, C. *et al.* (2018) A software tool ‘CroCo’ detects pervasive cross-species contamination in next generation sequencing data. *BMC Biology*, **16**, 28.
- Simion, P., Delsuc, F. & Philippe, H. (2020) To what extent current limits of phylogenomics can be overcome? *Phylogenetics in the genomic era* (ed. by C. Scornavacca, F. Delsuc and N. Galtier), pp. 2.1:1–2.1:34. <https://hal.archives-ouvertes.fr/hal-02535366>.
- Simmons, M.P. & Kessenich, J. (2020) Divergence and support among slightly suboptimal likelihood gene trees. *Cladistics*, **36**, 322–340.
- Song, H., Sheffield, N.C., Cameron, S.L., Miller, K.B. & Whiting, M.F. (2010) When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. *Systematic Entomology*, **35**, 429–448.
- Song, F., Li, H., Jiang, P. *et al.* (2016) Capturing the phylogeny of Holometabola with mitochondrial genome data and Bayesian site-heterogeneous mixture models. *Genome Biology and Evolution*, **8**, 1411–1426.
- Spangler, P.J. & Steiner, W.E. Jr (2005) A new aquatic beetle family, Meruidae, from Venezuela (Coleoptera: Adephaga). *Systematic Entomology*, **30**, 339–357.
- Steenwyk, J.L., Buida, T.J. III, Li, Y., Shen, X.-X. & Rokas, A. (2020) ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biology*, **18**, e3001007.
- 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.
- Stuart, A. (1955) A test for homogeneity of the marginal distributions in a two-way classification. *Biometrika*, **42**, 412–416.

- Suyama, M., Torrents, D. & Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, 609–612.
- Talavera, G. & Castresana, J. (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, **56**, 564–577.
- Tan, G., Muffato, M., Ledergerber, C., Herrero, J., Goldman, N., Gil, M. & Dessimoz, C. (2015) Current methods for automated filtering of multiple sequence alignments frequently worsen single-gene phylogenetic inference. *Systematic Biology*, **64**, 778–791.
- Tilic, E., Sayyari, E., Stiller, J., Mirarab, S. & Rouse, G.W. (2020) More is needed – thousands of loci are required to elucidate the relationships of the ‘flowers of the sea’ (Sabellida, Annelida). *Molecular Phylogenetics and Evolution*, **151**, 106892.
- Toussaint, E.F.A., Beutel, R.G., Morinière, J. *et al.* (2016) Molecular phylogeny of the highly disjunct cliff water beetles from South Africa and China (Coleoptera: Aspidytidae). *Zoological Journal of the Linnean Society*, **176**, 537–546.
- Toussaint, E.F.A., Seidel, M., Arriaga-Varela, E. *et al.* (2017) The peril of dating beetles. *Systematic Entomology*, **42**, 1–10.
- Vasilikopoulos, A., Balke, M., Beutel, R.G. *et al.* (2019) Phylogenomics of the superfamily Dytiscoidea (Coleoptera: Adephaga) with an evaluation of phylogenetic conflict and systematic error. *Molecular Phylogenetics and Evolution*, **135**, 270–285.
- Vasilikopoulos, A., Gustafson, G.T., Balke, M., Niehuis, O., Beutel, R.G. & Misof, B. (2021) Resolving the phylogenetic position of Hygrobiidae (Coleoptera: Adephaga) requires objective statistical tests and exhaustive phylogenetic methodology: a response to Cai *et al.* (2020). *Molecular Phylogenetics and Evolution*, **162**, 106923.
- van Vondel, B.J. (2019) Features of the metacoxal air-storage space as additional characters for reconstructing the phylogeny of Haliplidae (Coleoptera). *Tijdschrift voor Entomologie*, **162**, 13–32.
- Wang, H.-C., Minh, B.Q., Susko, E. & Roger, A.J. (2018) Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Systematic Biology*, **67**, 216–235.
- Wang, H.-C., Susko, E. & Roger, A.J. (2019) The relative importance of modeling site pattern heterogeneity versus partition-wise heterotachy in phylogenomic inference. *Systematic Biology*, **68**, 1003–1019.
- Whelan, S., Irisarri, I. & Burki, F. (2018) PREQUAL: detecting non-homologous characters in sets of unaligned homologous sequences. *Bioinformatics*, **34**, 3929–3930.
- Wong, T.K.F., Kalyanamoorthy, S., Meusemann, K., Yeates, D.K., Misof, B. & Jermini, L.S. (2020) A minimum reporting standard for multiple sequence alignments. *NAR Genomics and Bioinformatics*, **2**, lqaa024. <https://doi.org/10.1093/nargab/lqaa024>.
- Young, A.D. & Gillung, J.P. (2020) Phylogenomics – principles, opportunities and pitfalls of big-data phylogenetics. *Systematic Entomology*, **45**, 225–247.
- Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. (2018a) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, **19**, 153.
- Zhang, S.-Q., Che, L.-H., Li, Y., Liang, D., Pang, H., Ślipiński, A. & Zhang, P. (2018b) Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature Communications*, **9**, 205.

Accepted 1 July 2021

First published online 26 July 2021