Cardenas, C. R., Gustafson, G. T., & Toussaint, E. F. A. (2025). Global integration of phylogenomic data and fine-scale partitioning strategies refine the evolutionary tree of Adephaga beetles. *Bulletin of the Society of Systematic Biologists*.
https://doi.org/10.18061/bssb.v4i1.10345

Investigations

Global Integration of Phylogenomic Data and Fine-Scale Partitioning Strategies Refine the Evolutionary Tree of Adephaga Beetles

Cody Raul Cardenas[1], Grey T. Gustafson[2], Emmanuel F. A. Toussaint[1]

[1] Natural History Museum of Geneva, [2] Northern Arizona University

Keywords: Beetle evolution, Carboniferous insects, Coleoptera genomics, convergent evolution, Geadephaga ultraconserved element probe set, gene partitioning

https://doi.org/10.18061/bssb.v4i1.10345

Bulletin of the Society of Systematic Biologists

Abstract

Over the past decade, genomic-scale data has revolutionized insect phylogenetics by allowing the generation of increasingly comprehensive genomic and taxonomic datasets. However, because different approaches have been used, it is often difficult to understand to what extent these data can be integrated to reconstruct evolutionary trees. In this study, we focus on the beetle suborder Adephaga to explore whether genomic data produced in the past decade can be combined to reconstruct the largest phylogenomic tree of this clade to date. To that end, we collected publicly-available transcriptomes, genomes, and target sequence capture data of Adephaga beetles to generate a global dataset. Taking advantage of a newly developed bioinformatic pipeline, we demonstrate the overall compatibility of data types, especially of ultraconserved elements and exon-capture data. We also examined the impact of factors such as the treatment of off-target flanking genomic regions, data trimming regimes and partitioning, as well as varying levels of taxonomic and genomic sampling on phylogenomic inference. Using a matrix of 2,471 loci, we inferred the most comprehensive fossil-based evolutionary tree of Adephaga beetles. Our results confirm the independent colonization of aquatic ecosystems by two lineages. We also reconstruct Hygrobiidae as sister to Amphizoidae and a paraphyletic Aspidytidae, supporting the evolutionary convergence of prothoracic glands in both Hygrobiidae and Dytiscidae. Our results suggest an origin of Adephaga in the Carboniferous, with subsequent diversification of major lineages in the mid-Permian. Future efforts should focus on expanding the taxonomic sampling in Geadephaga, this clade of terrestrial beetles being the most diverse lineage in Adephaga and paradoxically one of the least sampled. To that end, we introduce a new ultraconserved element probe set tailored for Geadephaga beetles that will help generate compatible genomic data to further refine the Adephaga tree of life.

Introduction

Advances in high-throughput sequencing of genome-wide molecular data greatly enhance our ability to resolve the tree of life. This is particularly true for insects (Ribeiro & Espíndola, 2024), the most diverse clade of animals on Earth, for which phylogenomic trees have flourished in the past decade (e.g., Borowiec et al., 2025; Frandsen et al., 2024; Johnson et al., 2018; Kawahara et al., 2019; D. Liu et al., 2024; McKenna et al., 2019; Misof et al., 2014; Peters et al., 2017; Simon et al., 2019; Soghigian et al., 2023; Song et al., 2020; Wutke et al., 2024). Commonly used genomic data generation methods include whole-genome sequencing (Ng & Kirkness, 2010), transcriptomics (McGettigan, 2013; Wang et al., 2009), ultraconserved elements (UCEs; Faircloth et al., 2012), exon capture (ExC; Lemmon et al., 2012), restriction-site associated DNA sequencing methods (RADseq; Davey & Blaxter, 2010; Peterson et al., 2012) and mitogenomics (Curole & Kocher, 1999; Osigus et al., 2013). Capture methods such as ExC, UCEs, or hybrid approaches such as HyRAD-X (Schmid et al., 2017), have the benefit of being able to utilize historically preserved museum specimens in a cost-effective manner (Mayer et al., 2021; Pauli et al., 2024; Van Dam et al., 2017). On the other hand, genomes and transcriptomes, in addition to their high cost, generally require high-quality fresh genomic ma-

terial (but see Gauthier, Cardenas, et al., 2025). While the compatibility of ExC data with genomes and transcriptomes is straightforward, recent research has shown that UCEs as well could be combined, to some extent, with such data. For instance, many hexapod UCEs are composed of protein-coding sequences, allowing researchers to combine them with genomic and transcriptomic data, therefore increasing taxon sampling available for phylogenomic inferences (Baca et al., 2021; Bossert et al., 2019; Gustafson et al., 2023; Van Dam et al., 2021). The ability to integrate multiple sources of genomic data allows reconstructing more comprehensive phylogenomic trees for species-rich clades across the tree of life. With a better understanding of the genomic nature of available data produced in the past decade, it becomes feasible to combine several types of data into larger supermatrices. However, because the sequencing of genomic data is often limited for major insect lineages, few clades offer the chance to empirically explore such approaches (but see Baca et al., 2021; Bossert et al., 2019; Henríquez-Piskulich et al., 2024).

One such example is the beetle suborder Adephaga, comprising *ca.* 50,000 species of mostly predatory beetles inhabiting terrestrial and aquatic ecosystems worldwide. This lineage counts some of the most striking life histories, evolutionary features, and morphologies among beetles (Bergsten & Miller, 2007; Darlington, 1943; Dettner, 1985; Eisner & Aneshansley, 1999; Moore et al., 2022; H. G. Spangler, 1988; Tucker, 1969). Several recently described, relictual aquatic beetle families are also part of Adephaga (e.g., Aspidytidae and Meruidae, Ribera et al., 2002; P. J. Spangler & Steiner, 2005). Because of these characteristics, the systematics of Adephaga have received increased attention using both morphological and molecular approaches (Baca, Alexander, et al., 2017; Baca et al., 2021; Beutel, 1993, 1995, 1998; Beutel et al., 2006, 2013; Beutel, Ribera, et al., 2020; Beutel & Haas, 1996; Beutel & Roughley, 1988; Burmeister, 1976; Gustafson et al., 2020; Hammond, 1979; Ribera et al., 2002; Shull et al., 2001; Toussaint et al., 2016; Vasilikopoulos et al., 2019, 2021).

Traditionally, Adephaga beetles have been divided into two major groups, the Geadephaga and Hydradephaga (Crowson, 1960), corresponding to four extant superfamilies. Terrestrial Geadephaga beetles (= Caraboidea) are the most species-rich group in Adephaga with over 40,000 described species divided into three families: the Carabidae (ground beetles, *ca.* 37,000 species), Cicindelidae (tiger beetles, *ca.* 2,900 species), and Trachypachidae (false ground beetles, six species) (Lorenz, 2021; Wiesner, 2020). While family relationships within Geadephaga have historically been disputed (López-López & Vogler, 2017; Maddison et al., 1999, 2009; Shull et al., 2001), the monophyly of this lineage is currently well-supported (McKenna et al., 2019; Beutel et al., 2020; Baca et al., 2021; Gustafson et al., 2020; Vasilikopoulos et al., 2021). The Hydradephaga are aquatic or semi-aquatic beetles including the Amphizoidae (trout stream beetles, five species), Aspidytidae (water cliff beetles, two species), Dytiscidae (diving beetles, *ca.* 4,800 species), Haliplidae (crawling water beetles, *ca.* 240 species), Hygrobiidae (squeak beetles, six species, but see Nilsson (2006) for a discussion of family name priority and use of Paelobiidae instead of Hygrobiidae), Meruidae (waterfall beetles, 1 species), Noteridae (burrowing water beetles, *ca.* 280 species), and Gyrinidae (whirligig beetles, *ca.* 900 species) (Baca, Toussaint, et al., 2017; Miller & Bergsten, 2012; Nilsson, 2006; Nilsson & Hájek, 2024; Short, 2018). Similarly to Geadephaga, there has also been uncertainty about family relationships in the group; but the paraphyly of Hydradephaga is generally agreed upon, with Gyrinidae (= Gyrinoidea) placed as sister to the rest of the Adephaga, in which Geadephaga are sister to Haliplidae (= Haliploidea) and Dytiscoidea (Baca, Alexander, et al., 2017; Baca et al., 2021; Beutel, Ribera, et al., 2020; Gustafson et al., 2020; López-López & Vogler, 2017; McKenna et al., 2015, 2019; Vasilikopoulos et al., 2021; S.-Q. Zhang et al., 2018). The most recent phylogenomic treatments focusing on Adephaga by Baca et al. (2021) and Vasilikopoulos et al. (2021) resulted in largely consistent relationships although the two studies relied on alternative target-capture approaches to generate genomic data, namely UCEs and ExC. These two studies largely expanded our understanding of Adephaga phylogenomics and concurred in the establishment of a robust backbone for the suborder, with a monophyletic Geadephaga and a paraphyletic Hydradephaga (Baca et al., 2021; Vasilikopoulos et al., 2021). Future work should therefore focus on significantly increasing taxon sampling for all families to test the more fine-scale classification of Adephaga and to better understand the evolutionary history of the various lineages composing this species-rich lineages of beetles.

In that regard, one of the key issues emerging from the modern history of Adephaga phylogenomics, is the variety of data used to infer evolutionary trees. Because existing genomic datasets theoretically differ in nature, it is unclear to what extent these genomic datasets can be integrated into larger-scale genomic matrices for Adephaga and how future research should be undertaken to facilitate the inference of increasingly comprehensive phylogenies. To date, no investigation into the viability of combining the different types of capture data used in Adephaga phylogenomics has been performed. To fill this gap, the aim of this study is to primarily determine if ExC, genomes, transcriptomes, and UCE Adephaga genomic data can be integrated into a large supermatrix. Assuming such a large-scale integration of genomic data is achievable, it would (*1*) result in the largest Adephaga evolutionary tree ever assembled, benefiting from a decade of genomic work led by multiple research groups across the globe, (*2*) demonstrate that any type of future approach among the ones used recently to generate data can be integrated into this global effort, thereby augmenting the taxon sampling of an ever-growing Adephaga tree of life, and (*3*) allow for a further refinement of Adephaga evolutionary history by generating a much denser phylogeny in which additional and more accurately placed fossils could be used to estimate divergence times of Adephaga, pending a critical reassessment of the extant fossil record.

Methods

Taxon sampling

To test the feasibility of generating a large-scale genomic matrix of Adephaga, we combined several, publicly-available, genomic-level datasets (Baca, Alexander, et al., 2017; Baca, Toussaint, et al., 2017; Barclay, Geiser, et al., 2023; Barclay, Natural History Museum Genome Acquisition Lab, et al., 2023; L. Crowley et al., 2023; L. M. Crowley et al., 2021, 2023, 2024; Gough et al., 2020; Gustafson et al., 2020; McKenna et al., 2019; Misof et al., 2014; Peters et al., 2014; Pflug et al., 2020; Sivell et al., 2023; Sota et al., 2022; Van Belleghem et al., 2012; Vasilikopoulos et al., 2019, 2021; Weng et al., 2021). Those datasets were retrieved from the National Center for Biotechnology Information (NCBI) database https://www.ncbi.nlm.nih.gov/, DNA Read Archive of the DNA Database of Japan (DDBJ) https://www.ddbj.nig.ac.jp/, the Dryad Digital Repository: https://datadryad.org, and Zenodo zenodo.org (Table 1 A). Sequence data was composed of exon capture (ExC), genomic assemblies (GEN), RNAseq transcriptomes (TRA) and ultraconserved elements (UCE; Table 1 B). A majority of these data (AllTaxa, Ntax = 317) represents the beetle suborder Adephaga (Ntax = 311), but the other beetle suborders Archostemata (Ntax = 3), Myxophaga (Ntax = 2) and Polyphaga (Ntax =1) were also included as outgroups. A subset of 220 taxa was also generated (SubTaxa) in subsequent analyses to harmonize the taxonomic coverage at the genus-level. In this subset the genus-level representatives with the highest number of identified loci were kept, those representing different data types (e.g., ExC, Genomic, Transcriptomic, and UCE), and those belonging to major species-groups of a given genus based on the literature (e.g., *Carabus* based on Sota et al. (2022); *Calosoma* based on Toussaint & Gillett (2018), Toussaint et al. (2021) and Sota et al. (2022), *Pamborus*: based on Sota et al. (2022)). Additionally, based on the sequence data assemblies (see *Preliminary locus recovery*) we confirmed taxon identity using MitoFinder (Allio et al., 2020). We specifically aimed at recovering fragments of the mitochondrial cytochrome c oxidase subunit I (CO1) commonly found as off-target sequencing (aka bycatch) in capture methods (see supplemental materials). A list of taxa, accession numbers, references, and, in some cases, corrected species identifications used, are available in the supplementary materials.

Table 1. Number of Adephaga samples used in each dataset.

A) Family	AllTaxa	SubTaxa
Amphizoidae	3	3
Aspidytidae	3	3
Carabidae	170	79
Cicindelidae	23	22
Dytiscidae	69	69
Gyrinidae	11	11
Haliplidae	11	9
Hygrobiidae	4	4
Meruidae	1	1
Noteridae	13	11
Trachypachidae	3	2
Total	311	214
B) Data type	**AllTaxa**	**SubTaxa**
ExC	95	90
GEN	15	15
TRA	39	39
UCE	168	76
Total	317	220

(A) The column AllTaxa includes all Adephaga samples used for designing and testing the integrated *in silico* probe sets. The column SubTaxa indicates the number of samples used to reduce taxonomic biases of some groups in subsequent analyses. The total does not include outgroups. (B) The total number of samples per data type for the AllTaxa and SubTaxa datasets. The total includes outgroups.

Integrated probe set

We combined all available Adephaga oligonucleotide probes (i.e., baits) used to generate both ExC and UCE data *in silico*. These probes are designed to hybridize with and isolate genomic regions of interest in high-throughput sequencing libraries (Mayer, 2016; C. Zhang & Mirarab, 2022). After hybridization, libraries are enriched and sequenced; and these reads are assembled and mapped against the probe set used to generate them. The available probe sets that were used in this study are the following: Adephaga ExC (49,786 probes targeting 923 protein-coding exons from 651 genes; Vasilikopoulos et al., 2021), UCE Adephaga 2.9kv1 (38,948 probes targeting 2,941 loci; Gustafson et al., 2019) and the UCE Coleoptera 1.1Kv1 (13,674 probes targeting 1,172 loci; Faircloth, 2017); hereafter referred to as ExC, UCE-Ade, and UCE-Col respectively. By design, the UCE-Ade probe set contains 300 probes from the UCE-Col (Gustafson et al., 2019). These probes targeting *ca.* 300 redundant loci were removed from the UCE-Col in our workflow given that they primarily target Adephaga beetles and are included in the UCE-Ade probes. For the downstream workflow we considered all ExC probes targeting an exon the equivalent of probes targeting a locus.

We mapped all probe sets against four genomes *Nebria brevicollis* (Carabidae; GCA_944738965.1; L. Crowley et al., 2023), *Nebria salina* (Carabidae; GCA_944039245.1; Sivell et al., 2023), *Ophonus ardosiacus* (Carabidae, GCA_943142095.1; L. M. Crowley et al., 2023), and *Pterostichus madidus* (Carabidae; GCA_911728475A.2; L. M. Crowley et al., 2021), in order to identify which genome recovered the most loci. Of the four genomes used, the one of *Pterostichus madidus* recovered the most loci (see supplemental materials and Table S2-S3). Mapping was performed using a combination of *Burrows-Wheeler Aligner MEM* (*BWA MEM* v0.7.17; Li & Durbin, 2009), *SAMtools* (v1.6; Li et al., 2009), and *bedtools* (v2.31.0; https://bedtools.readthedocs.io), to identify and remove overlapping probes (Figure 1A). Additional identification and removal of overlapping probes was performed using *BLAST* (v2.14; Altschul, 1997),

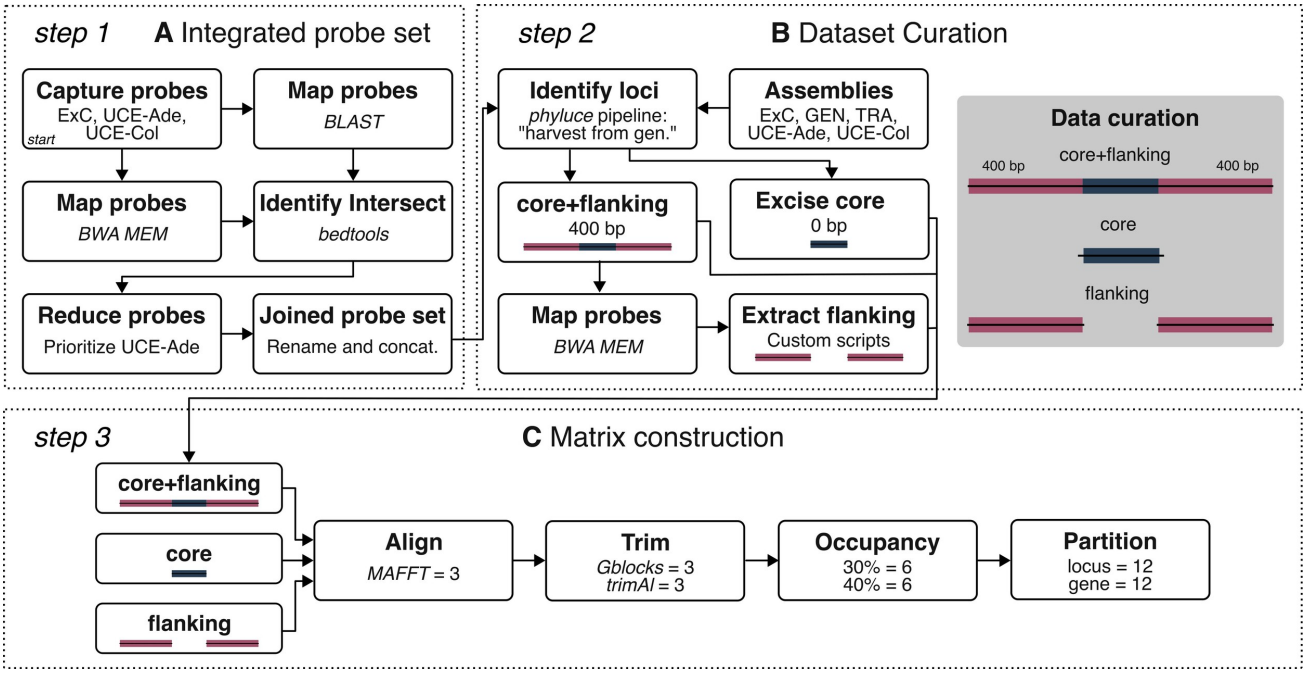

Figure 1. Flow chart of the bioinformatic workflow used in this study.

A flow chart of the bioinformatic workflow used for the AllTaxa dataset. A) The pipeline to create the joined probe set described in section *Integrated probe set*. B) The pipeline to generate the curation datasets using the joined probe set and assemblies as described in the *Dataset curation* section. A visual explanation of the curated data is also presented. C) The pipeline to generate datasets for phylogenomic inference with the AllTaxa dataset. Each step indicates the progressive generation of data matrices as described in *Matrix construction*.

with a percentage identity of 90% and no gaps. Before merging loci, we identified which probes from ExC and UCE-Col overlapped with the UCE-Ade probes, because the latter targets the most loci. This step ensures that multiple probes do not map to more than one contig (See Figure 2). This is because we take advantage of *phyluce*'s (v 1.7.1; Faircloth, 2016) duplicate locus and probe identification, as well as paralog discovery, during locus recovery from sequence data ("Finding UCE step," *phyluce_assembly_match_contigs_to_probes*). This makes the bioinformatic workflow readily accessible and simplifies our downstream workflows.

Up to this step of the bioinformatic pipeline, the different probe sets were kept separate. However, it was necessary to adjust the probe IDs to comply with the expected format in *phyluce*. To maintain separation of probes in the joined probe set, we renamed the remaining probes from each set, keeping the UCE-Ade probes original identifiers (e.g., uce-1, uce-2, ... etc.), while UCE-Col identifiers started at uce-1000000 and ExC at uce-2000000, so that all loci had their own unique ID. All remaining and renamed probes were concatenated, to create a custom probe fasta file for locus recovery called joined probe set (see Supplemental Materials). To quantify the effectiveness of this approach, we identified the number of loci recovered for the AllTaxa dataset with the ExC, UCE-Ade, UCE-Col, and joined probe sets. All custom scripts used for these steps can be found at https://github.com/crcardenas/Adephaga_UCE.

Preliminary locus recovery

We first processed and assembled all data described in the section *Taxon sampling*. UCE sequence data from Sota et al. (2022) was assembled using the default *phyluce* pipeline, while all other short-read sequence data had technical reads trimmed and sequence-quality checked using *fastP* (v0.19.5; Chen et al., 2018), with default parameters. The resulting trimmed reads were assembled using *SPAdes* (v3.14.1; Prjibelski et al., 2020), following the same defaults settings as in *phyluce*. All assembled non-UCE sequence data, including those generated with ExC, were processed using the *phyluce* pipeline "*Harvesting UCE Loci From Genomes*" by matching the UCE-Ade probes to assemblies with a minimum coverage of 50% and identity of 80% (*–coverage 50 –identity 80*; Baca et al., 2021; Bossert et al., 2019). Identified loci were then extracted with 1600 base-pair (bp) flanking regions for genomic loci, and 800 bp flanking regions for ExC and transcriptome loci, following the treatment of genomic and transcriptomic data found in Bossert et al. (2019) and Baca et al. (2021) (e.g., *–flank 800*).

As suggested by the *phyluce* workflow, the resulting datasets were aligned in *phyluce* and "internally trimmed" using the default *phyluce Gblocks* wrapper (v 0.91b; Castresana, 2000), given the age of most Adephaga family-level splits (> 50 million years ago, Ma; Baca et al., 2021). A non-exhaustive examination of alignments was performed in *Geneious* (v 2023.2; http://www.geneious.com/), revealing many lengthy matrices (median maximum length of 7,200 bp, minimum length of 301 bp and maximum length of 26,220 bp), that were longer than the extracted genomic loci (*ca.* 3,300 bp), and excessively gappy (see Figure S1).

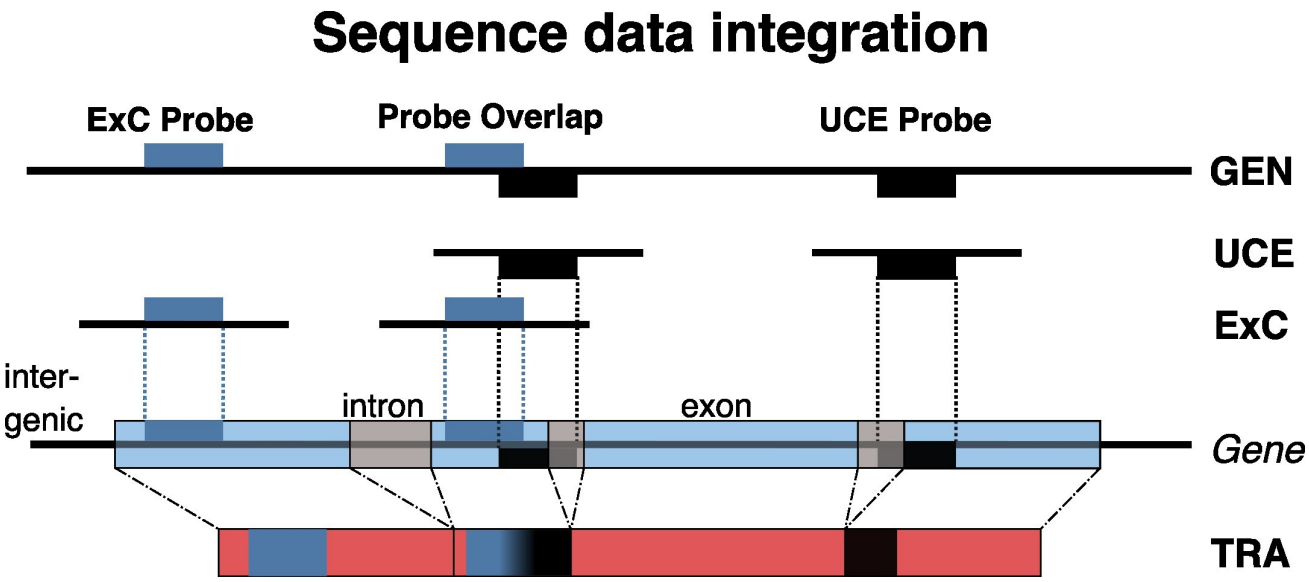

Figure 2. Sequence data integration

A graphic representation of integrating diverse genomic data sources and probe overlap on a theoretical gene; modified from Bossert et al. (2019). All rows are presented as if they are aligned except for the bottom RNAseq example (TRA). The first row simplifies where targeted exon capture (ExC, dark blue rectangles) and ultraconserved element (UCE, black rectangles) probes may target a theoretical genome (GEN) represented by a solid black line. The second and third rows represent the resulting sequence capture data for UCE and ExC data respectively. The fourth row represents a theoretical gene in light blue with three exons and two introns. Here, probes are shown to overlap with each other, and the UCE partially on an intron. Lastly the fifth row represents a TRA sequence in red and where the respective probes map.

As discussed by Bossert et al. (2019) this might be due to a lack of intergenic and intronic data in transcriptomic data (See Figure 1 A). Alternatively, this may be due to the fact that some sequence data, e.g., UCE-Col data (Sota et al., 2022), had extremely long assembled contigs (median length of 2,779 bp and maximum length of 23,670 bp), that appeared to contribute to gappy alignments. The resulting trimming performed with *Gblocks* in *phyluce* returned comparably short alignments (median length of 148 bp and maximum length of 849 bp) that appeared to have suffered a drastic loss of phylogenomic information. Given these initial results, we developed an alternative pipeline to curate the sequence data utilizing methods available in *phyluce* (see below).

Dataset curation

Following the preliminary inspection of resulting alignments pre- and post-trimming (see previous section *Preliminary locus recovery*), we predicted that the core genomic regions targeted by the probes (i.e., excluding the flanking regions), should be both conserved and variable enough for phylogenomic inferences at deep evolutionary scales (see also *Conservedness of probe kits* section). We therefore generated three datasets to test this prediction: core, core+flanking, and flanking. In these datasets, core are the regions of sequence data targeted by the probes, core+flanking are the regions targeted by the probes plus the non-probe flanking regions resulting from off-target sequencing and assembly, lastly flanking are exclusively the non-probe flanking regions resulting from off-target sequencing (see Figure 1 B). We used the same *phyluce* "*Harvesting UCE Loci From Genomes*" pipeline as previously described but reduced the flanking lengths to 400 bp (core+flanking) and 0 bp (core) for all data mapped to the joined probe set (ExC, GEN, TRA, & UCE data). To recover only the flanking regions, the core loci were mapped against the core+flanking, to identify and remove the core region using custom scripts (see supplemental materials).

Locus characterization and partitioning schemes

Characterizing loci targeted by probe sets, allows cogenic loci to be merged for phylogenomic inference, under the assumption that cogenic loci share a similar evolutionary history. Merging cogenic UCEs has been shown to increase bootstrap values and the overall robustness of phylogenomic analyses (Van Dam et al., 2021).

To characterize the genomic features targeted by the UCE-Ade and the joined probe set, we used the genome of *Pterostichus madidus* (genome annotations are available at https://projects.ensembl.org/darwin-tree-of-life/). We annotated intergenic and intronic features not present in the original genome annotation, using *bedtools complement*. The mapping information from the *integrated probe set* section was summarized using *bedtools groupby*, to identify the genomic nature of the targeted loci. Identified cogenic loci were merged in a new partition file using a custom script (see supplemental materials). To account for the potential impact of cogenic loci being identified and merged, we performed dedicated phylogenomic inferences (See *Concatenation-based phylogenomic inference*).

Matrix construction

Each dataset was aligned using *MAFFT* (v7.475; Katoh & Standley, 2013) with the *–auto* option, outside of the *phy-*

luce pipeline (Figure 1 C). We included the comparison of two different trimming methods by generating additional datasets trimmed using *trimAl* (v1.4; Capella-Gutiérrez et al., 2009), and *Gblocks* (Castresana, 2000; Talavera & Castresana, 2007). The parameters used for trimming were *trimAl -automated1* for each alignment, while for *Gblocks* (v 0.91b) parameters were set to the same defaults used by *phyluce* for all alignments (*–b1 0.5 –b2 0.85 –b3 8 –b5 10*). Post alignment and trimming, we discarded loci that contained less than three taxa and alignments shorter than 50 bp. Combining all so far described methods of curation, trimming, and partitioning resulted in 12 variants of the AllTaxa dataset used for phylogenomic inference (see Table 2).

As the amount of available phylogenomic data has grown, concerns about the impact of locus occupancy (i.e., the minimum number/percentage of taxa for a given locus), and taxon sampling on inferences have been raised (Lemmon et al., 2009; Roure et al., 2013; Wiens, 1998). Research has also shown that Including more loci with the drawback of having more missing datat can improve concatenation-based phylogenomic inferences (Huang & Knowles, 2016; Jiang et al., 2014; Streicher et al., 2016; Wiens & Tiu, 2012). To account for these important developments, locus occupancy was taken into account in a set of alternative analyses. We generated subsets with 30% and 40% locus occupancy, thereby augmenting the number of phylogenomic analyses to a total of 24 for the AllTaxa dataset (but see also *Concatenation-based phylogenomic inference* for the SubTaxa dataset).

Concatenation-based phylogenomic inference

Using the previously described 24 matrices, phylogenomic analyses were conducted on the AllTaxa datasets (Ntax = 317; Table 1). The optimal partitioning schemes and corresponding nucleotide substitution models were identified in *IQ-TREE2* (v2.1.2 Minh et al., 2020), using ModelFinder (commands *-MF+MERGE* and *-rfclusterf*; Chernomor et al., 2016; Kalyaanamoorthy et al., 2017), with the best partition selected using the Akaike Information Criterion corrected (AICc). Maximum likelihood (ML) phylogenomic inferences were performed in *IQ-TREE2* with the *-allnni* option. Branch support was estimated using 1000 ultrafast bootstrap replicates (UFBoot; Hoang et al., 2018; Minh et al., 2020), with the *-bnni* option, as well as 1000 SH-like approximate likelihood ratio tests (SH-aLRT; Guindon et al., 2010). Combined values of UFBoot ⩾ 95 and SH-aLRT ⩾ 80 for a given branch were considered robust. Phylogenomic analyses were run on Baobab, one of the high-performance bioinformatic clusters of the University of Geneva.

Following the phylogenomic inferences of these 24 matrices from the AllTaxa datasets, we examined and selected our preferred trimming and locus recovery method (i.e., core with *trimA*l; see *Results*), based on the number of recovered loci and expected topology. We then performed phylogenomic inferences with the SubTaxa dataset to reduce taxonomic bias (Ntax = 220; Table 1). However, we only used the preferred trimming method (i.e. *trimAl*, see *Results*) with the SubTaxa dataset to generate 12 additional matrices: with all curation approaches, 30% and 40% locus occupancy, and gene or locus partitioning (see Table 2). These additional phylogenomic analyses were performed using the same parameters described above.

Coalescent-based species tree estimation

Parallel phylogenomic inferences in a coalescent framework were conducted to account for the potential impact of gene tree discordance in the preferred SubTaxa matrix. We used *IQ-TREE2* to generate locus trees from our partition file (Ntax = 220; Table 1). ModelFinder was implemented without the *+MERGE* option, otherwise the setup for all ML tree reconstructions was the same as described previously. Because we used short alignments (i.e., core probe regions), with moderate degrees of missing data, and selected lower coverage matrices (30% and 40% locus occupancy), we relied on *weighted ASTRAL* (*wASTRAL* v1.15.2.3; C. Zhang & Mirarab, 2022), to take into account branch support as a measure of phylogenetic uncertainty. Local posterior probabilities (LPP) were calculated using default parameters (LPP ⩾ 95 were considered robust).

Estimates of phylogenomic support

We further quantified genealogical and site-specific concordance of the preferred matrix using *IQ-TREE2* (Minh et al., 2020; Mo et al., 2023). With this approach the site and gene concordance factors (sCFs and gCFs respectively), were calculated by quartet analysis providing support for the branch of a tree based on the data (site or genes). A sCF score less than or near 33% indicates that ML methods may favor a different topology, whereas gCFs range from 0% to 100% indicating the proportion of loci or genes that support the topology estimated using ML phylogenomic inference. The locus trees used in *wASTRAL* were used for concordance analyses.

Divergence time estimation

We performed Bayesian relaxed-clock divergence time estimation in *BEAST* (v1.10.4; Suchard et al., 2018), using the preferred ML topology inferred with the SubTaxa (N = 220) core trimAl 30% gene partitioning matrix as a fixed input topology (See *Results*). To reduce the computational burden linked to the use of thousands of genomic-scale sequences for hundreds of lineages, we used the gene-shopping approach developed in *SortaDate* (Smith et al., 2018). We selected 100 loci that best fit the *SortaDate* default criteria and created a new partition and matrix file using *AMAS* (Borowiec, 2016). Using *PartitionFinder2* (v2.1.1; Lanfear et al., 2014, 2016; Stamatakis, 2014), we searched for the optimal partitioning scheme using the options *–min-subset-size* 2000, *–rcluster-max* 1000, and *–rcluster-percent* 20, to minimize overparameterization. The best nucleotide substitution models available in *BEAST* were then searched using *ModelFinder* in *IQ-TREE2*. The identified partitioning scheme and substitution model were used as input in *BEAUTi* v1.10.4 (Suchard et al., 2018). Each partition was

Table 2. Matrix statistics

A) 30% locus occupancy	Locus partitioning	Gene partitioning	Missing data	Alignment length (bp)	Distinct patterns	PIS	IS
AllTaxa core Gblocks	2,931	2,539	58.509 %	337,279	284,534	158,010	147,177
AllTaxa core trimAl	2,941	2,546	61.249 %	423,924	363,598	210,862	171,882
AllTaxa core+flanking Gblocks	2,420	2,232	66.504 %	918,136	848,408	603,803	200,693
AllTaxa core+flanking trimAl	2,420	2,232	66.004 %	706,083	639,817	425,569	184,095
AllTaxa flanking Gblocks	1,167	1,116	58.895 %	59,279	55,997	38,346	10,496
AllTaxa flanking trimAl	2,283	2,108	71.348 %	479,035	464,376	373,015	49,902
SubTaxa core trimAl	2,872	2,471	58.551 %	405,363	320,920	198,822	171,241
SubTaxa core+flanking trimAl	2,325	2,149	65.038 %	808,479	704,679	516,647	197,461
SubTaxa flanking trimAl	1,197	1,148	69.618 %	377,219	358,818	292,774	41,516
B) 40% locus occupancy	**Locus partitioning**	**Gene partitioning**	**Missing data**	**Alignment length (bp)**	**Distinct patterns**	**PIS**	**IS**
AllTaxa core Gblocks	1,741	1,497	54.095 %	225,062	189,424	106,569	99,126
AllTaxa core trimAl	1,741	1,498	56.183 %	265,065	226,075	131,356	110,583
AllTaxa core+flanking Gblocks	1,222	1,151	61.365 %	520,283	481,088	344,465	113,530
AllTaxa core+flanking trimAl	1,221	1,150	61.425 %	407,334	370,095	247,796	104,861
AllTaxa flanking Gblocks	710	687	54.621 %	41,298	38,909	26,357	7,646
AllTaxa flanking trimAl	1,095	1,040	66.804 %	266,271	257,871	204,597	29,481
SubTaxa core trimAl	1,782	1,517	52.943 %	263,670	207,188	130,518	112,703
SubTaxa core+flanking trimAl	1,239	1,161	59.731 %	468,131	409,663	301,754	113,571
SubTaxa flanking trimAl	1,180	1,110	64.799 %	350,865	337,007	280,344	33,827

Matrix statistics for the AllTaxa and SubTaxa datasets. Each sub-table indicates the locus occupancy A) 30% and B) 40%. Each row represents a dataset starting with AllTaxa (N = 317 taxa) or SubTaxa (N = 220 taxa), followed by the curation and trimming method (e.g., SubTaxa core trimAl). For each matrix the following is reported from the *phyluce* log file unless stated otherwise: locus/gene partitioning - the number loci defined by the respective partition file, missing data - the degree of missing data in the alignment (using *AMAS*, Borowiec, 2016), alignment length (bp) - the total length in base pairs of the alignment, distinct patterns - the total distinct patterns in the matrix, parsimony informative sites - the number of parsimony informative sites (PIS) in the matrix, and invariant sites (IS) - the number of invariant sites in the partition.

assigned an uncorrelated lognormal relaxed clock, the mean of which was set up with a uniform distribution (starting value of 0.01, upper value of 1.0, and lower value of 1.0e-6). We enforced node calibrations based on the Adephaga fossil record that we re-curated to account for recent fossil descriptions (e.g., A. G. Kirejtshuk & Ansorge, 2023; H. Liu et al., 2023; Rosová et al., 2023, see supplemental materials), new possible placements due to the expanded taxon sampling, and incorrect placements of previously included taxa (e.g., Schmidt et al., 2023). In total, we relied on 28 carefully selected fossils for divergence time estimation analyses in *BEAST* (Table 3).

Fossil calibrations were set up as exponential or lognormal prior distributions in two sets of analyses with 95% of the distribution ranging from the minimum age of the fossil (i.e., age of the geological stratum in which the fossil was

described) up to 50 million years older. For each fossil calibration, we set the real space of the lognormal distribution to 20 million years and the standard deviation to 13.47 million years. For exponential priors, we set a mean value of 13.65 million years and estimated the offset values to include the 50 million year range previously described (see supplementary materials). Five replicates of ten independent *BEAST* analyses per prior distribution type (50 each) were run on Bamboo, one of the high-performance bioinformatic clusters of the University of Geneva. Each run consisted of 30 million generations, with tree and parameter sampling every 2,500 generations. The log files were examined in *Tracer* (v 1.7.2; Rambaut et al., 2018) to assess convergence. After applying a burnin of 5 million generations, the remaining sampled parameters and trees were combined in *LogCombiner* v1.10.4 (Suchard et al., 2018), before being summarized in *TreeAnnotator* v1.10.4 (Suchard et al., 2018), to produce maximum clade credibility trees.

Conservedness of probe kits

The joined probes used in this phylogenomic approach are designed to target conserved orthologous regions of genomes (Faircloth, 2017; Gustafson et al., 2019; Vasilikopoulos et al., 2021). Moreover, Gustafson et. al (2020, 2023) proposed that the UCE-Col probes are more conserved than that of UCE-Ade. Using a linear mixed effect model (LMM) with the R package *lme4* (Bates et al., 2015) we tested how conserved the ExC, UCE-Ade, and UCE-Col probe sets were. To compare the probe sets, we used parsimony informative and invariant sites of the targeted loci as a measure of conservedness. We predict that a conserved probe set will have fewer parsimony informative sites and more invariant sites compared to others given the sequence data available.

We recovered statistics relative to parsimony informative and invariant sites from the output generated by *IQ-TREE2* for the six AllTaxa 30% locus occupancy datasets (core+flanking trimAl, core+flanking Gblocks, core trimAl, etc.; see *phylogenomic inference* section). We identified loci shared by the six datasets and calculated the proportion of parsimony informative and invariant sites per locus from the *IQ-TREE2* log file as variables of interest. We tested linear models for both parsimony informative and invariant sites, and we defined the fixed effect as probe source (ExC, UCE-Ade, and UCE-Col). We included two random effects: trimming within curation (6 categorical effects) and locus. Using locus as a random effect accounts for variation from each dataset that may be present in each locus. Three models were built: one including all random effects, a second including only curation within trimming, and a third accounting only for variation within loci. The best-fit model was selected using the homoscedasticity and the distribution of residuals along with the Akaike Information Criterion (AIC) of each model. We used the R package *lmerTest* (Kuznetsova et al., 2017) to generate significance values and the R package *emmeans* (Lenth, 2017; Searle et al., 1980) to estimate the means of the fixed effects in the best-fit model and compare each mean estimate.

UCE probe subset optimization for Geadephaga

To support future research in Geadephaga (see *Results* and *Discussion*), we used *phyluce* to develop a dedicated subset of UCE-Ade probes for optimized probe synthesis and *in vitro* targeted capture. We removed probes unlikely to successfully recover loci in most Geadephaga, based on the large genetic distance measures of design taxa as reported in Gustafson et al. (2019), but also based on redundancy, wherein a close relative with smaller genetic distance measures was available. We also tested the impact of excluding all non-base genome probes, except for one other distant taxon in both the tailored probes and generalized UCE-Col probes within the full UCE-Ade probe set (see supplemental material). Following each removal, the remaining probes were aligned to clean, assembled contigs generated from mixed data (transcriptomic and UCE) for ten Geadephaga taxa across the three extant recognized families: Carabidae, Cicindelidae and Trachypachidae (Gough et al., 2020; Pflug et al., 2020). Resulting changes in locus recovery relative to the number of probes remaining in the probe set were investigated. More specific details of the new probe kit design can be found in supplemental material.

The final subset of the UCE-Ade probes was optimized for maximum locus recovery, using a minimal number of probes. We tested the probes *in silico* to ensure that this was the case and that the minimal difference in loci targeted between the subset and full UCE-Ade probe set did not significantly affect phylogenomic performance. *In silico* testing involved the generation of 50% and 80% locus occupancy matrices that were subjected to ML analysis using IQ-TREE, with the model set to GTR and topological support assessed with 1000 ultrafast bootstrap replicates.

Results

Integrated probe set

The overarching goal of this study was to combine the wealth of publicly-available data offered by several recently published studies, to augment both locus and taxon sampling across Adephaga. We combined three probe sets to "recapture" sequence data *in silico*. The resulting joined probe set contains 90,130 probes targeting 4,255 loci; with 46,102 ExC probes targeting 860 loci, 34,995 UCE-Ade probes targeting 2,619 loci, and 9,033 UCE-Col probes targeting 776 loci. We removed 3,684 ExC probes targeting 63 loci and 4,641 UCE-Col probes targeting 396 loci because these probes overlapped with UCE-Ade and targeted multiple loci in the *Pterostichus madidus genome*. Locus recovery of the *P. madidus* genome in *phyluce* using the new joined probe set recovered 677 loci targeted by ExC probes, 362 loci targeted by UCE-Col probes, and 2,486 loci targeted by UCE-Ade probes; for a total of 3,525 targeted loci. *Phyluce* subsequently identified and removed 458 targeted loci for matching multiple contigs, and 452 contigs for matching multiple targeted loci, resulting in a final set of 2,994 loci found in the *P. madidus* genome for phylogenetic inference.

Table 3. List of fossils used for divergence time estimation in BEAST.

Fossil	Number	Suborder	Placement	Min. Age	Family	Subfamily	References
†*Coleopsis archaica* A. G. Kirejtshuk et al., 2014	1	Root Coleoptera	Root Min Age	295	†Tshekardocoleidae	-	A. G. Kirejtshuk et al., 2014
†*Hydroscapha jeholensis* Cai et al., 2012	2	Myxophaga	Stem *Hydroscapha*	122.46	Hydroscaphidae	-	Cai et al., 2012; Chatzimanolis et al., 2012; Fikáček et al., 2020; Fraser et al., 2017
†*Kirghizocupes proporeius* Ponomarenko, 1969	3	Archostemata	Stem Cupedidae	221.5	Cupedidae	-	A. Kirejtshuk et al., 2016; Ponomarenko, 1969
†*Apriacma tuberculosa* J. J. Tan et al., 2006	4	Archostemata	Stem *Priacma*	122.46	Cupedidae	Cupedinae	A. Kirejtshuk et al., 2016; J. J. Tan et al., 2006
†*Cretotortor sp.* Nel, 1989	5	Adephaga	Stem Heterogyrinae	174.1	Gyrinidae	Heterogyrinae	Gustafson et al., 2017; Nel, 1989
†*Cretodineutes rotundus* Liang et al., 2020	6	Adephaga	Stem Dineutini	98.17	Gyrinidae	Gyrininae	Liang et al., 2020; Shi et al., 2012
†*Gyrinus aquisextanea* Nel, 1989	7	Adephaga	Crown *Gyrinus*	23.03	Gyrinidae	Gyrininae	Gustafson et al., 2017; Nel, 1989
†*Apermunda asiaticus* Ponomarenko & Volkov, 2013	8	Adephaga	Stem Trachypachidae	252.3	Trachypachidae	†Eodromeinae	A. G. Kirejtshuk & Ansorge, 2023; Ponomarenko & Volkov, 2013
†*Palaeopronyssiformia groehni* Wiesner et al., 2017	9	Adephaga	Crown Cicindelidae	41.3	Cicindelidae	-	Schmidt et al., 2023; Wiesner et al., 2017
†*Kryzhanovskiana olegi* Kataev et al., 2019	10	Adephaga	Stem Paussinae	98.17	Carabidae	Paussinae	Kataev et al., 2019
†*Cretomophron mutilus* Rosová et al., 2023	11	Adephaga	Stem Omophroninae	98.17	Carabidae	Omophroninae	Rosová et al., 2023
†*Burmapseudomorphus planus* Beutel et al., 2020	12	Adephaga	Stem Harpalinae	98.17	Carabidae	Harpalinae	Beutel, Liu, et al., 2020
†*Calosoma agassizi* Nel, 1989	13	Adephaga	Stem *Calosoma*	23.03	Carabidae	Carabinae	Nel, 1988
†*Limodromus hoffeinsorum* Schmidt, 2015	14	Adephaga	Stem Platynini	41.3	Carabidae	Platyninae	Schmidt, 2015
†*Archaeonebria inexspectata* Schmidt & Kavanaugh 2019	15	Adephaga	Crown Nebriini	41.3	Carabidae	Nebriinae	Schmidt et al., 2019
†*Bembidion bukejsi* Schmidt & Michalik, 2017	16	Adephaga	Stem *Bembidion*	41.3	Carabidae	Trechinae	Schmidt & Michalik, 2017

Fossil	Number	Suborder	Placement	Min. Age	Family	Subfamily	References
†*Haliplus cretaceus* Prokin & Ponomarenko, 2013	17	Adephaga	Stem *Haliplus*	122.46	Haliplidae	Haliplinae	Ponomarenko & Prokin, 2015; Prokin & Ponomarenko, 2013
†*Hydroporus carstengroehni* Balke et al., 2010	18	Adephaga	Stem *Hydroporus*	41.3	Dytiscidae	Hydroporinae	Balke et al., 2010; Wolfe et al., 2016
†*Derovatellus rostrata* (Koch & Berendt 1854)	19	Adephaga	Stem Vatellini	41.3	Dytiscidae	Hydroporinae	Klausnitzer, 2003; Michat & Torres, 2011
†*Coptotomus balticus* Hendrich & Balke, 2020	20	Adephaga	Stem *Coptotomus*	41.3	Dytiscidae	Coptotominae	Hendrich & Balke, 2020
†*Japanolaccophilus beatificus* Balke & Hendrich, 2019	21	Adephaga	Stem Laccophilini	41.3	Dytiscidae	Laccophilinae	Balke & Hendrich, 2019; Wolfe et al., 2016
†*Copelatus aphroditae* Miller & Balke 2003	22	Adephaga	Stem *Copelatus*	41.3	Dytiscidae	Copelatinae	Miller & Balke, 2003; Wolfe et al., 2016
†*Colymbetes miocaenicus* Říha, 1974	23	Adephaga	Stem *Colymbetes*	11.6	Dytiscidae	Colymbetinae	Říha, 1974
†*Hydrotrupes prometheus* Gómez & Damgaard, 2014	24	Adephaga	Stem *Hydrotrupes*	41.3	Dytiscidae	Agabinae	Gómez & Damgaard, 2014; Wolfe et al., 2016
†*Ambarticus myanmaricus* Yang et al., 2019	25	Adephaga	Stem Dytiscinae	98.17	Dytiscidae	Dytiscinae	Shi et al., 2012; Yang et al., 2019
†*Acilius florissantensis* Wickham, 1909	26	Adephaga	Stem Aciliini	33.9	Dytiscidae	Dytiscinae	Wickham, 1909; https://collections.peabody.yale.edu/search/Record/YPM-IP-000005
†*Cybister cf. rotundatus* Říha, 1974	27	Adephaga	Stem Cybistrini	15.9	Dytiscidae	Cybistrinae	Fikáček et al., 2008
†*Palaeodytes gutta* Ponomarenko, 1987	28	Adephaga	Stem Dytiscidae	155.7	Dytiscidae	-	Ponomarenko, 1987; Ponomarenko & Prokin, 2015; Prokin et al., 2013

Details of the 28 fossil taxa used as calibrations in this study to infer divergence times of Adephaga beetles. Some fossils used in Baca et al. (2021) were discarded because new paleontological information and/or their re-examination casts some doubt on their placement (e.g. Schmidt et al., 2023). The table lists all fossils used and provides the following information: fossil name, suborder, placement, minimum age, family, and subfamily. Additional information such as geological deposit or prior settings used in *BEAUti*, are included in the supplemental materials.

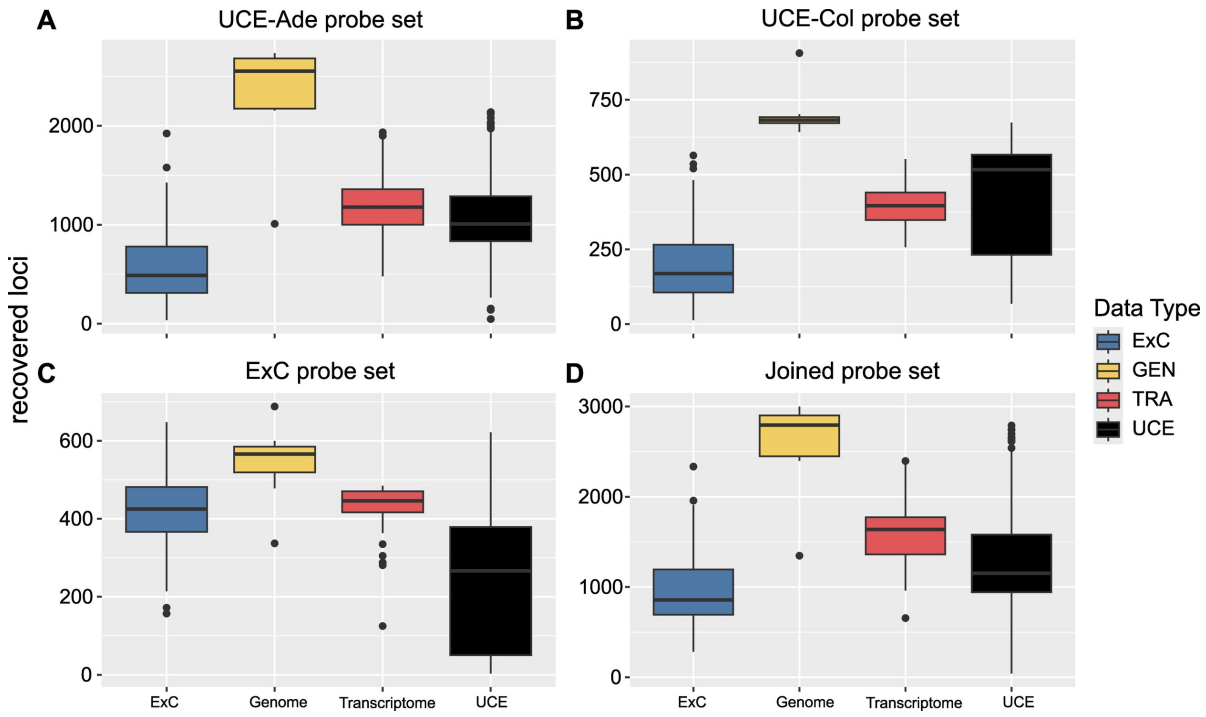

Figure 3. Locus recovery of genomic data for each probe set.

A box and whisker plot of the number of targeted loci found in the available Adephaga sequence data: exon capture (ExC), genomes (GEN), transcriptomes (TRA), and ultraconserved elements (UCE). The statistics are derived from the core+flanking dataset resulting from the standard *phyluce* UCE workflow for phylogenomic data. Each sequence data type is mapped against a) the ultraconserved elements Adephaga UCE probe set (UCE-Ade, 2,941 targeted loci), b) the ultraconserved elements Coleoptera probe set (UCE-Col, 1,172 targeted loci), c) the exon capture probe set (ExC, 923 targeted protein-coding exons), and d) the joined probe set (4,255 targeted loci) generated in this study.

We found that ExC sequence data recovered fewer loci when mapped against just UCE probes. Similarly, UCE sequence data recovered fewer loci when mapped against just the ExC probes (Figure 3 A-C). The ExC sequence data mapped against the UCE-Ade probes recovered an average of *ca.* 60 more loci (*ca.* 120% increase). In some cases, only 12-20% of ExC loci mapped to UCE-Ade, but in others 400% to nearly 545% more loci were recovered; e.g. *Galerita* sp. SRR12339130: 443 ExC to 1,922 UCE-Ade and *Haliplus lathridioides* SRR12339113: 232 ExC to 1263 UCE-Ade (Figure 3 D). Overall, we found a notable average increase in recovered loci using the joined probe set for each data type (Figure 3 D; supplementary materials).

Dataset Curation

To improve phylogenomic inference, we generated a bioinformatic pipeline to curate sequence data, where the targeted core and flanking regions were extracted from targeted loci (Figure 1 B). For each curation method the total locus recovery was *ca.* 99% of the total targeted loci (n = 4,255; Table S3). The core+flanking datasets recovered the most loci and on average the core datasets recovered the most taxa per locus. Because we extracted the core regions from the core+flanking data to generate the flanking dataset, the number of recovered loci and taxa are identical for the core+flanking and flanking datasets (See supplemental).

Locus characterization and partitioning scheme

Mapping the joined probe set to the *Pterostichus madidus* genome using *BWA MEM* and *BLAST* identified 3,378 targeted loci. Of these, 2,600 of loci were identified as genic and the remaining 778 as intergenic genomic features (Figure S3). Approximately 26% of these probes target both intronic-exonic or exonic-intergenic features. In total, 468 genic features recovered 1,122 cogenic targeted loci (Figure S5). Characterization of the UCE-Ade probe set showed that 71% of the targeted genomic features are exonic, including those that target both intronic-exonic and intergenic-exonic features (Figure S4). Importantly, for both the joined and UCE-Ade probe sets a small percentage of probes were found to target loci from all three genomic features: intronic, exonic, and intergenic (UCE-Ade: 1.31% and joined: 1.04%).

Using the characterization of the joined probe set, targeted loci that were identified as cogenic were merged for phylogenomic inference (i.e., gene partitioning datasets). The initial set of gene partitions contained 174 genes with 326 cogenic loci targeted by the ExC probes set, 387 genes with 750 cogenic loci targeted by the UCE-Ade probes and 48 genes with 67 cogenic loci targeted by the UCE-Col probes.

Matrix construction

When aligned, the minimum number of sequences per locus (n = 3), resulted in slightly fewer loci than described in

the *Integrated probe set* section: core 4,212, core+flanking 4,215, and flanking 4,215. The total number of sequences in the alignments were marginally higher for the core curation method (Table S4 A, see supplemental materials), compared to the core+flanking and flanking methods (Table S4 B & C, see supplemental materials). Examination of the same statistics for source probes in our integrated probe set (i.e., the joined probe set), showed that ExC probes recovered more loci compared to the two UCE probe sets (see supplemental materials).

The gappy flanking regions were highly reduced by both trimming approaches. The most extreme reduction in alignment length occurred in the curated datasets that retained flanking data; both *Gblocks* and *trimAl* removed 1.6 to 1.8 kbp from core+flanking and flanking alignments. While the alignments for the core and core+flanking data were on average longer than this minimum length threshold (50 bp; Table S5 A & B), trimming approaches for all curated datasets had some loci reduced below the minimum length and some longer loci still retained. *Gblocks* reduced the length of the flanking data the most, ultimately removing more loci given our minimum alignment length (mean = 55 bp, median 38 bp; Table S5 C, see supplemental materials). The change in alignment lengths also depended on which loci were targeted by the original source probes in our joined probe set. When trimmed, the UCE-Ade and ExC probes reduced the length more so than UCE-Col. While these results were not as extreme as in *Preliminary locus recovery*, we still observed a significant impact on alignment length in our AllTaxa dataset. *trimAl* removed a smaller number of loci from the core+flanking data, and most loci were retained except for the flanking data which was greatly reduced; core: *Gblocks* 4,185, *trimAl* 4,212 (of 4,212 unaligned); core+flanking: *Gblocks* 4,203, t*rimAl* 4,203 (of 4,215 unaligned); flanking: *Gblocks* 2,142, *trimAl* 4,181 (of 4,214 unaligned).

Concatenation phylogenomic inference

The phylogenomic inferences relying on datasets trimmed using *Gblocks* often resulted in a higher number of unrealistic/incoherent topologies compared to those relying on datasets trimmed using *trimAl* (supplemental materials). Comparison of phylogenomic inferences using different curation methods revealed that the core datasets consistently inferred more realistic topologies compared to other curation approaches. While the core+flanking resulted in topologies that were more consistent with core data than those based on flanking, the flanking data always recovered inconsistent and moderately supported topologies (e.g., paraphyletic Gyrinidae or Noteridae + Meruidae as sister to the rest of Adephaga). For most datasets the gene and locus partitioning produced identical family-level topologies, but gene partitioning returned higher branch support in AllTaxa 30% datasets. In some cases when fewer loci were recovered, as in the Gblocks 40% dataset, branch supports were higher in inferences based on locus rather than gene partitioning. The AllTaxa core trimAl datasets (30%, 40%, and both partitioning schemes), included more loci and the respective phylogenomic analyses produced well-resolved topologies with strong branch support. From the previous results (see *Matrix construction* results) and careful screening of phylogenomic relationships recovered in alternative topologies recovered (see supplemental materials), we selected *trimAl* as trimming method for the SubTaxa analyses.

As we observed with the AllTaxa analyses, inferences based on the SubTaxa core datasets returned more robust relationships at different evolutionary time scales than those including flanking regions. The core+flanking and flanking data further revealed that the flanking regions are increasingly difficult to align and likely have a strong impact on phylogenomic inferences at deeper evolutionary relationships. This is most likely because sequence data with an average length of 719.1 bp (234 bp min, 980 bp max) becomes six times as long once aligned and is composed of many gaps (before any trimming, see Figure S1). In contrast, differences between the AllTaxa and SubTaxa core dataset phylogenomic analyses were very limited or null depending on the dataset and partitioning strategy. The results of all SubTaxa phylogenomic analyses (including the ones based on core+flanking and flanking) are summarized at the subfamily level in Figure 4, and all inferred trees and tree statistics can be found in the supplementary materials. We assessed that the SubTaxa core trimAl 30% gene dataset was likely the most robust topology, and it was selected to present the phylogenomic relationships and used as a fixed input for divergence time estimation.

Coalescent-based species tree inference

The higher-level phylogenomic relationships (i.e. family-level relationships) in the *wASTRAL* coalescent analyses of locus trees generated from the SubTaxa core trimAl 30% gene partitioning resulted in low local posterior probabilities along the tree backbone (e.g. LPP < 0.60, see Supplemental Materials). Backbone topological conflict with the ML inferences was significant but with low statistical support. This indicates there is no supported discordance between the ML and *wASTRAL* analyses. Besides, some subfamily relationships were highly supported and shared between *wASTRAL* and *IQ-TREE* analyses (e.g., Hygrobiidae + (Amphizoidae + Aspidytidae), Lancetinae + Coptotominae, Harpalinae + Brachininae, or Trachypachidae + (Cicindelidae + Carabidae), Figure 4).

Estimates of phylogenomic support

The concordance factor analysis of locus trees generated from the SubTaxa core trimAl 30% gene partitions resulted in differing degrees of support at the locus and site levels. The gCF varied across the preferred topology ranging from 0.26 to 13.56, while the sCF had lower variance ranging from 31.93 to 42.75. These results indicate that 6 to 335 of 2471 loci informed the preferred topology, and that approximately one third of the parsimony informative sites supported deep phylogenomic relationships (see also Supplemental Materials for all sCF and gCF values).

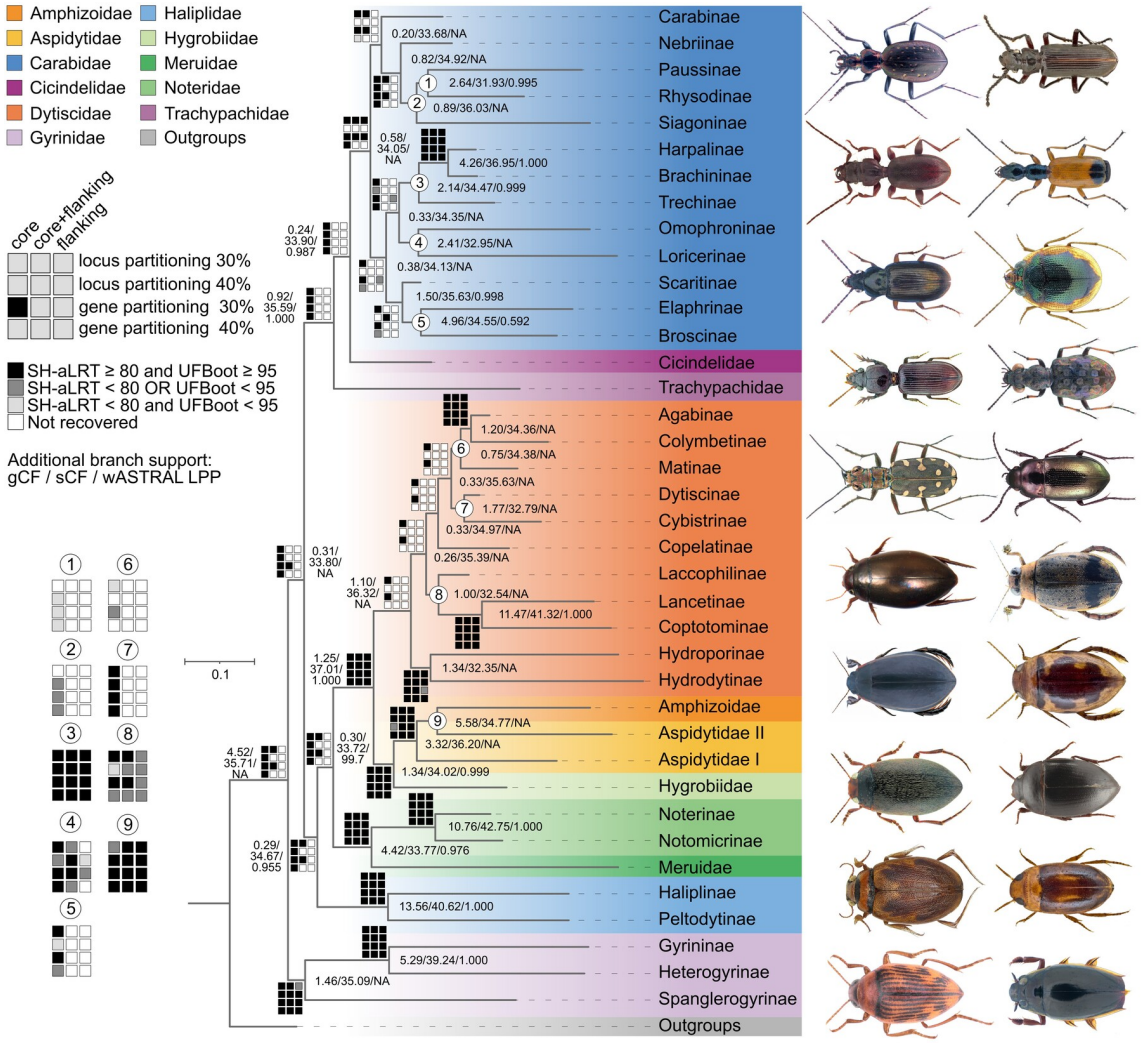

Figure 4. Summary of phylogenomic inferences for Adephaga beetles.

Preferred maximum likelihood (ML) phylogenomic tree of Adephaga resulting from the IQTREE analysis of the SubTaxa core trimAl 30% gene dataset. Only family and subfamily level phylogenomic relationships are presented. Branch support from alternative ML phylogenomic analyses of other datasets are presented in mosaic plots. Branch support values for which space was not available are numbered with their corresponding mosaic plots to the left of the tree. Additional branch information for the preferred topology including the gene and site concordance factors (gCF and sCF respectively), and wASTRAL local posterior probabilities (LPP) are displayed (gCF / sCF / LPP). Photographs of species belonging to major lineages of Adephaga are presented, from left to right and top to bottom: *Carabus depressus* (Carabidae, Carabinae), *Ozaena lemoulti* (Carabidae, Paussinae), *Siagona europea* (Carabidae, Siagoninae), *Odacantha melanura* (Carabidae, Harpalinae), *Pogonus chalceus* (Carabidae, Trechinae), *Omophron brettinghamae* (Carabidae, Omophroninae), *Clivina montei* (Carabidae, Scaritinae), *Elaphrus aureus* (Carabidae, Elaphrinae), *Calomera littoralis* (Cicindelidae), *Trachypachus gibbsii* (Trachypachidae), *Ilybius fenestratus* (Dytiscidae, Agabinae), *Eretes griseus* (Dytiscidae, Dytiscinae), *Cybister lateralimarginalis* (Dytiscidae, Cybistrinae), *Neptosternus kaszabi* (Dytiscidae, Laccophilinae), *Hygrobia hermanni* (Hygrobiidae), *Hydroporus erythrocephalus* (Dytiscidae, Hydroporinae), *Sinaspidytes wrasei* (Aspidytidae), *Neohydrocoptus subvittulus* (Noteridae, Noterinae), *Haliplus confinis* (Haliplidae, Haliplinae), and *Patrus nathani* (Gyrinidae, Gyrininae). Photos were taken by Michael Balke, Udo Schmidt and ukbeetles.co.uk. Specimen images are not to scale.

Divergence time estimation

The independent *BEAST* divergence time analyses recovered consistent estimates. For each run the acceptance ratio of operators was greater than zero, with most > 0.23. The *uniform(nodeHeights(treeModel))* operator acceptance ratio was lower in all runs, with an average of 0.10 for each distribution type. These independent runs had effective sample size (ESS) for most parameters > 200, and all > 100 except for replicate five of the lognormal distribution in which the *Calosoma* fossil calibration posterior had an ESS = 80. Replicate three for exponential and two for the lognormal distributions were selected based on likelihood and ESS values. The estimated ages for the exponential analysis are found in Table 4. Both exponential and lognormal analyses recovered consistent divergence times and credibility intervals (see Supplemental materials). Overall, our median crown estimates tend to push back in time the divergences of major Adephagan lineages (Figure 5). The trees for both analyses are available in the supplementary materials.

Conservedness of probe sets

We find that sequence data generated using UCE probes is generally more conserved than those with ExC regardless of the curation or trimming method used. A total of 1,146 loci were found in common between the six AllTaxa 30% datasets. These loci are composed of 320 ExC, 571 UCE-Ade, and 255 UCE-Col. We find that UCE-Col has the lowest mean parsimony informative sites and most invariant sites (Figure 6 A, B). For all flanking data, some loci can be com-

Table 4. Comparison of divergence times

Crown Group	This study Median age in Ma (95% CI)	McKenna et al. (2019) Median age in Ma (95% CI)	Baca et al. (2021) Median age in Ma (95% CI)
N Fossils	28	18	23
N Adephaga Fossils	24	2	18
Coleoptera	346 (331-364)	327 (297-343)	317 (303-335)
Adephaga	277 (272-283)	230 (197-257)	255 (247-266)
Caraboidea	252 (251-254)	170 (150-198)	227 (221-236)
Gyrinoidea	257 (237-272)	136* (83-179)	242 (230-254)
Haliploidea	176 (162-191)	NA	157 (145-170)
Dytiscoidea	251 (254-266)	NA	220 (212-231)

Comparison of divergence times from the exponential analysis for major clades of Adephaga estimated in this study and a selection of studies that recovered similar backbone relationships. See selected studies for further discussion. The 95% credibility intervals (CI) correspond to the 95% HPD parameter summarized from the BEAST posterior samples in this study, Baca et al. (2021), and from MCMCtree (Dos Reis et al., 2016) in McKenna et al. (2019). (*) Note that Gyrinoidea in McKenna et al. (2019) was only represented by the subfamily Gyrininae and did not include Heterogyrinae or Spanglerogyrinae hence the large deviation with other studies listed in the table.

posed entirely of parsimony informative and invariant sites (Figure 6 A, B). The best-fit model included the random effects of trimming within curation and loci, this model has standard errors that were homogenous with low variance and the overall effect of probe source is strong (Supplemental File 1: SuppFile1_LMM A, B). For selection of the best-fit LMM selection, see supplementary materials. The estimated global means indicates that loci from the ExC probe set contain the highest proportion of parsimony informative sites and UCE-Col the lowest proportion, and the inverse result for the invariant sites (Supplemental File 1: SuppFile1_LMM C, D). Pairwise comparison of estimated means returned standard errors that were homogenous with little variance and significant differences in the estimated means between all probe sets, where UCE-Col has fewer parsimony informative sites and more invariant sites than ExC and UCE-Ade (Supplemental File 1: SuppFile1_LMM E, F).

UCE probe subset optimization for Geadephaga

We found minimal loss in locus recovery following removal of UCE-Ade probes designed for taxa with large genetic distances from the other Geadephaga used in probe design, as well as those with some form of representational redundancy (supplemental materials). Ultimately, we were able to limit the number of UCE-Ade probes to a subset of 11,208 probes (30% of the total probes) targeting 2,925 loci (99% of the total loci), with evidence for minimal reduction in locus recovery across Geadephaga taxa (Table S1).

In silico testing of this final subset of probes relative to the full UCE-Ade probe set revealed a loss of only 68 loci between the 50% complete matrices, and 64 loci between the 80% complete matrices generated (Figure S6 & S7). Phylogenomic analysis showed neither differences in topology nor branch support between trees produced by the 50% matrices (Figure S8 & S9). Trees produced using the 80% complete matrices differed only in the placement of a single clade and had largely similar branch support (Figure S10 & S11). These *in silico* results suggest the two probe sets are likely to perform similarly for *in vitro* targeted-capture and enrichment in Geadephaga taxa, with missing loci and probes unlikely to significantly impact the phylogenomic performance of the UCE-Ade subset relative to the full probe set. We call this new probe subset the "Gea-Sub-2.9Kv1" and make it publicly available under a public domain license from dryad, available in the supplementary materials.

Discussion

Effectiveness of integrating diverse genomic data

The construction of a 2,471-locus supermatrix (512 targeted by ExC probes, 1,282 by UCE-Ade, and 253 by UCE-Col), allowed for the inference of the most comprehensive phylogenomic tree of Adephaga to date (Figures 4 and 5). In total, 1,022 cogenic loci were merged into 424 gene partitions, illustrating the power of locus characterization in genomic data. The overlap between ExC and UCE datasets further supports the genic nature of hexapod UCEs (see Baca et al., 2021; Bossert et al., 2019; Van Dam et al., 2021). By integrating annotated chromosome-level genomes into our pipeline, we identified the genomic features (exonic, intronic, and intergenic) targeted by the joined probe set. This is particularly significant given the incomplete characterization of the UCE-Ade probe set (Baca et al., 2021), due to limited genomic data. Approximately 71% of UCE-Ade probes target exonic regions in the *Pterostichus madidus* genome, aligning with findings by Baca et al. (2021), for UCE-Col probes mapped to the genome of *Tribolium castaneum* (Coleoptera, Polyphaga, Tenebrionidae). Interestingly, *ca.* 1% of loci targeted by the probes originated from all three genomic features: exonic, intronic, and intergenic. This discrepancy may stem from the automated pipelines used for genome annotation. Future improvements to algorithms or transcriptome-based genome annotations are expected to enhance UCE characterization. Regardless, merging cogenic loci targeted by probes based on

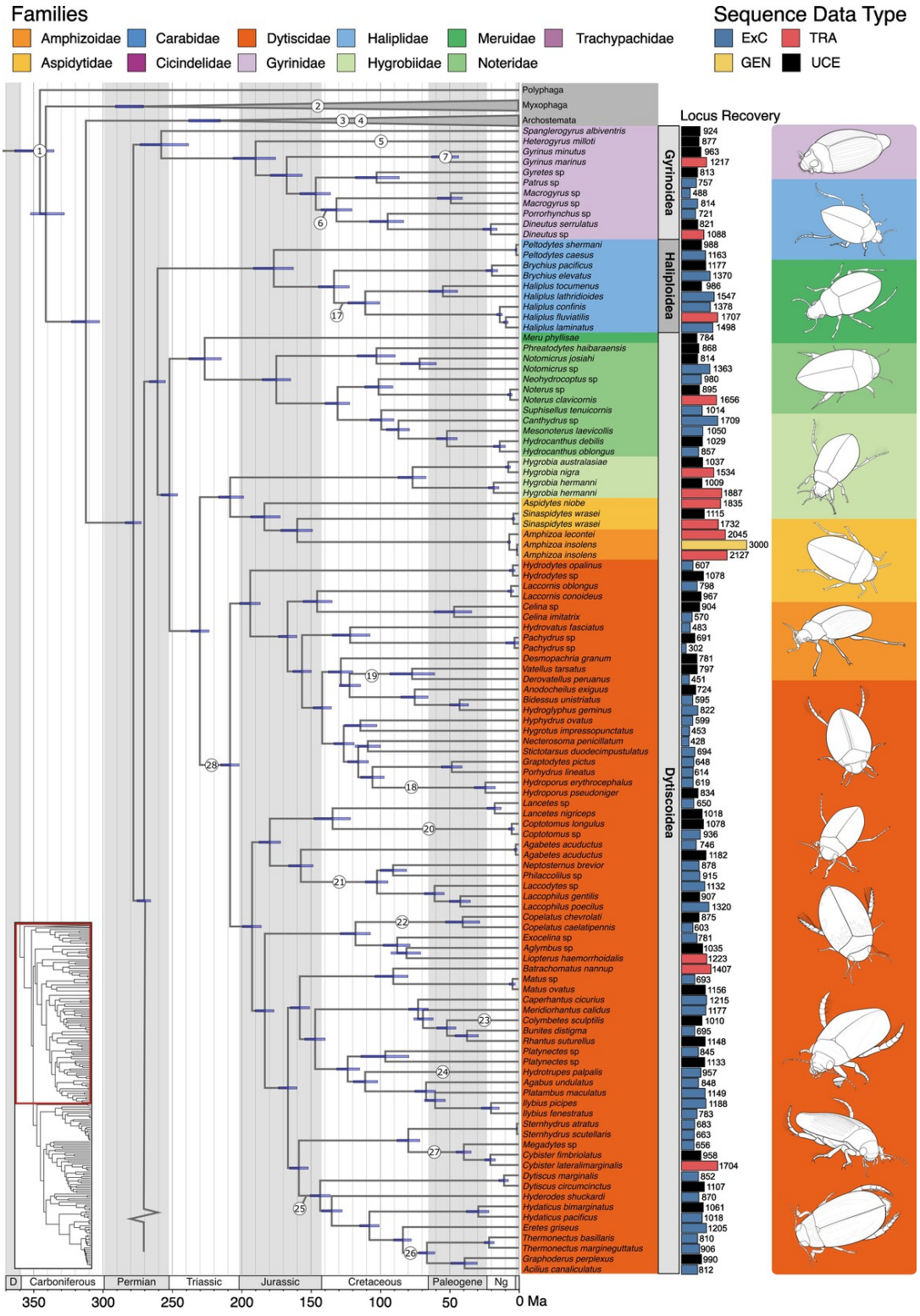

Figure 5 a. Bayesian time-calibrated timetree of Adephaga.

Time-calibrated tree of Adephaga based on 28 fossil calibrations implemented with an exponential prior distribution and a fixed topology derived from the SubTaxa core trimAl 30% gene matrix. The tree was constructed using the *Interactive Tree Of Life* (Letunic & Bork, 2024), FigTree (http://tree.bio.ed.ac.uk/software/figtree/), and merged in *Inkscape* (Inkscape Project, 2022). Outgroups are collapsed. Numbers placed on nodes indicate the placement of fossil calibrations used to estimate divergence times in BEAST. The number of loci recovered per taxon with the joined probe set is presented as a bar plot along the tree tips indicating the original source of the data as exon capture (ExC - blue), transcriptome (red), genome (gold), or ultraconserved elements (UCE - black). Illustrations of Adephaga beetles are presented along the right of the figure, from top to bottom: *Gyrinus* (Gyrinidae, Gyrininae), *Haliplus* (Haliplidae, Haliplinae), *Meru* (Meruidae), *Hydrocanthus* (Noteridae, Noterinae), *Hygrobia* (Hygrobiidae), *Sinaspidytes* (Aspidytidae), Amphizoa (Amphizoidae), *Desmopachria* (Dytiscidae, Hydroporinae), *Hydroporus* (Dytiscidae, Hydroporinae), *Laccophilus* (Dytiscidae, Laccophilinae), *Cybister* (Dytiscidae, Cybistrinae), *Dytiscus* (Dytiscidae, Dytiscinae), and *Eretes* (Dytiscidae, Dytiscinae). Illustrations are not to scale.

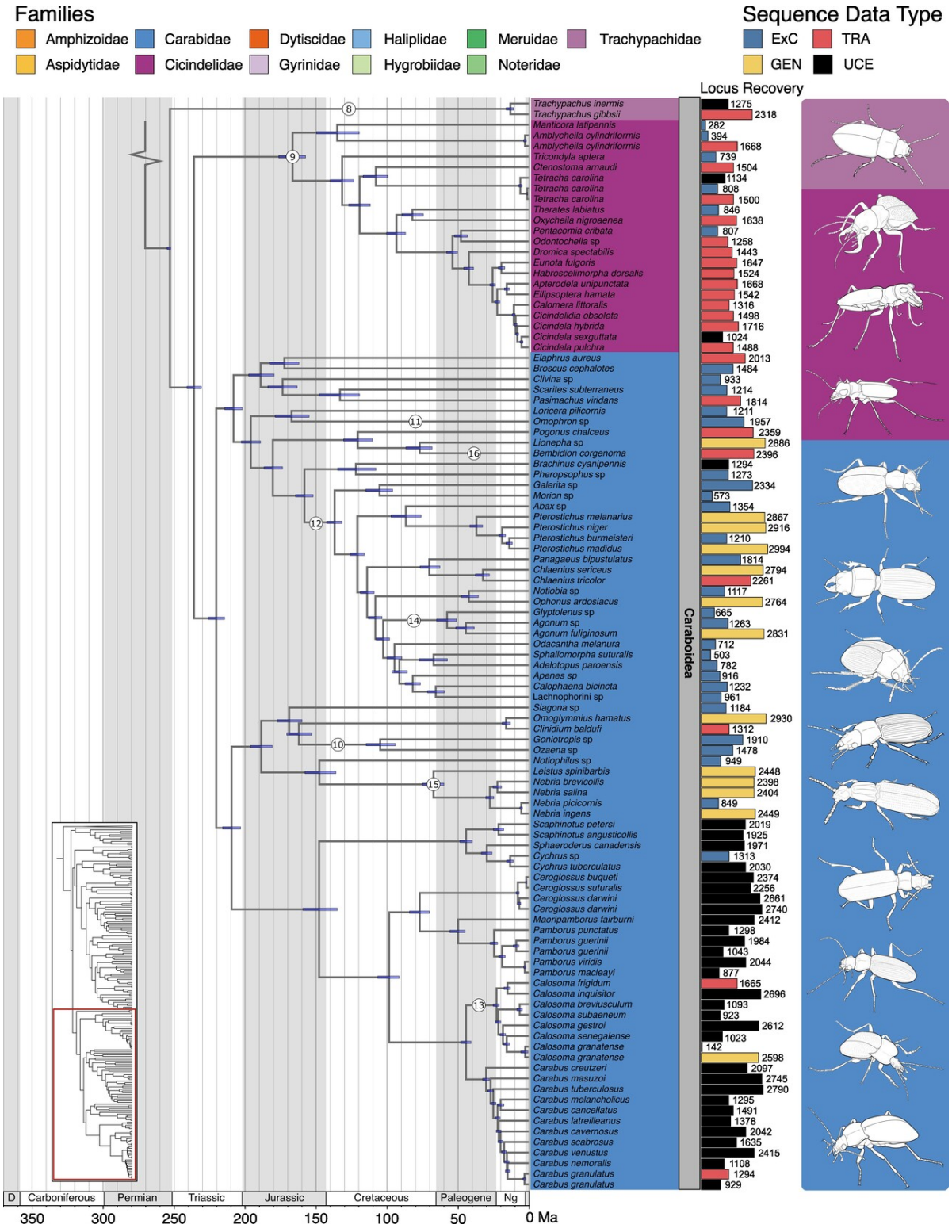

Figure 5 b. Bayesian time-calibrated timetree of Adephaga.

Continuation of the time-calibrated time tree based on 28 fossil calibrations implemented with an exponential prior distribution and a fixed topology derived from the SubTaxa core trimAl 30% gene matrix. Numbers placed on nodes indicate fossil taxa used to calibrate the tree. The number of loci recovered per taxon with the joined probe set is presented as a bar plot to the right of taxa names, color codes indicate the original source of the data as exon capture (ExC - blue), transcriptome (TRA - red), genome (GEN - gold), or ultraconserved elements (UCE - black). Illustrations of Adephaga beetles are presented along the right of the figure, from top to bottom: *Trachypachus* (Trachypachidae), *Manticora* (Cicindelidae), *Therates* (Cicindelidae), *Cicindela* (Cicindelidae), *Elaphrus* (Carabidae, Elaphrinae), *Scarites* (Carabidae, Scaritinae), *Omophron* (Carabidae, Omophroninae), *Pterostichus* (Carabidae, Harpalinae), *Omoglymmius* (Carabidae, Rhysodinae), *Goniotropis* (Carabidae, Paussinae), *Leistus* (Carabidae, Nebriinae), *Scaphinotus* (Carabidae, Carabinae) and *Carabus* (Carabidae, Carabinae). Illustrations are not to scale.

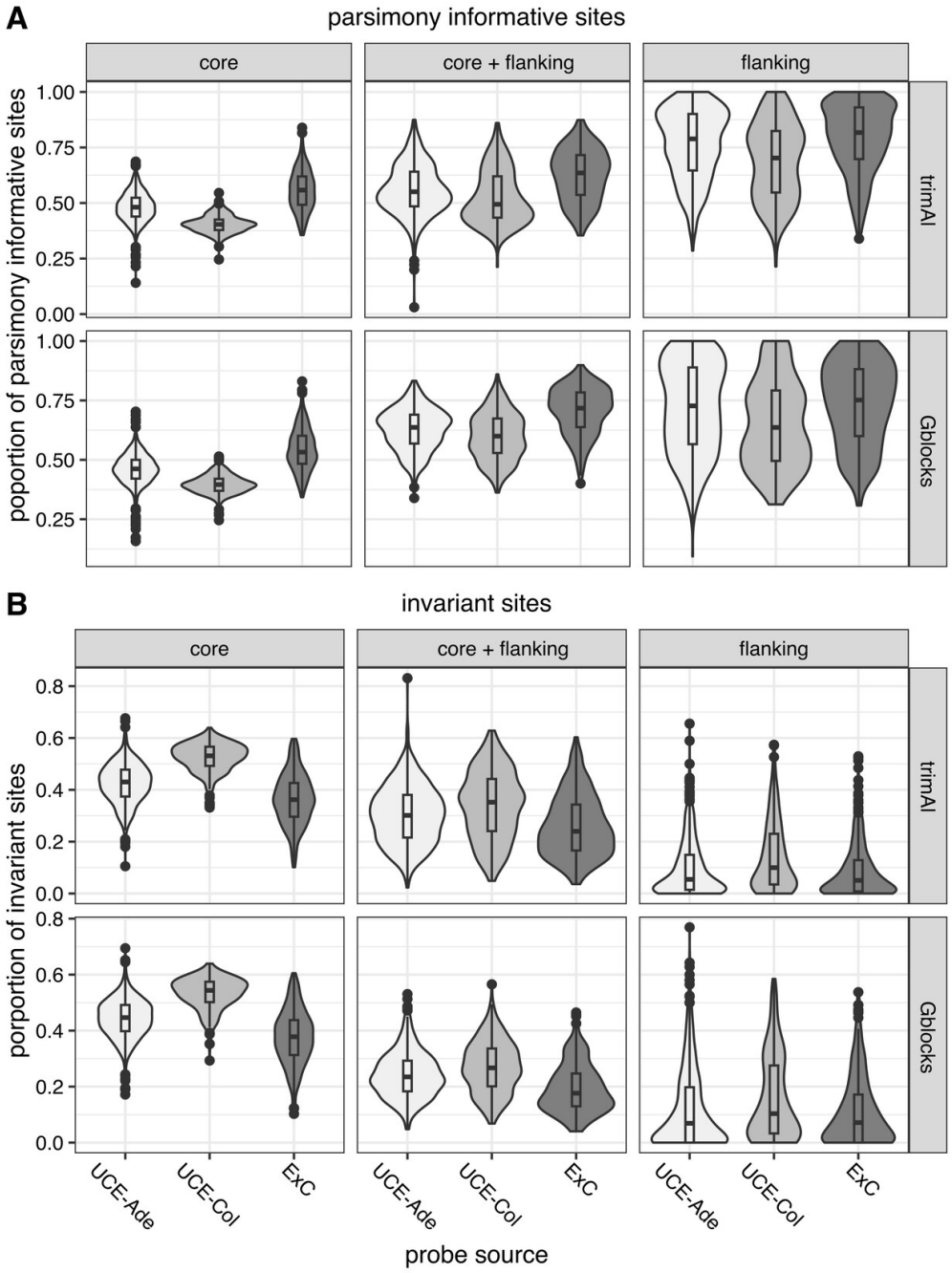

Figure 6. Proportion of parsimony informative and invariant sites.

Violin plots of the distribution of A) parsimony informative sites and B) invariant sites, for AllTaxa 30% datasets. For both A) and B) violin plots, columns indicate the curation method and rows indicate the trimming approach, reflecting the categorical groupings used in the linear mixed models. Each plot contains a violin plot of groupings by UCE-Ade, UCE-Col, and ExC.

characterization, strengthened the robustness of phylogenomic inferences.

The combination of loci targeted by different probe sets within our pipeline improved the compatibility and expanded the number of taxa available for phylogenomic inferences. At deeper evolutionary time scales, fewer conserved loci are available. As a result, increasing the number of targeted loci is necessary to identify informative loci, even at the expense of lower locus occupancy and missing data. Similarly, including more taxa from clades of interest helps ensure their representation despite low locus occupancy. Moreover, integrating probe sets with varying levels of conservedness enhances the robustness of phylogenomic inferences. Most importantly, removing the flanking region and applying a more conserved trimming approach improved matrix construction.

We developed an alternative approach to that of Bossert et al. (2019), eliminating the need for semi-manual curation with liberal trimming approaches like *Gblocks*. Instead, only targeted core probe regions are considered and flanking data is discarded. Generally, multi-sequence alignment trimming is used to reduce non-biological signals introduced by alignment methods. However, while some studies suggest that trimming loci can weaken phylogenomic sig-

nals (G. Tan et al., 2015), others argue that it is essential to improve phylogenomic inferences (Talavera & Castresana, 2007). Our comparison clearly underlines that *Gblocks* removes more informative sites from conserved core regions than *trimAl* (Table 2). This likely stems from *Gblocks* having a fixed trimming parameter applied to all locus alignments, whereas *trimAl* employs a per-locus algorithmic decision-based approach that may be better suited for integrated datasets. Additionally, Talavera & Castresana (2007) suggest that *Gblocks* performance decreases with shorter alignments (1,200-2,400 bp), whereas our preferred curation approach relies on alignments averaging 200 bp before trimming (Table S5). By restricting analyses to the core regions targeted by probes and excluding flanking regions, we consistently recovered stronger phylogenomic signals and better-supported phylogenomic relationships.

By extracting the core region using *phyluce*, we significantly reduced the need to trim persistent gappy regions in capture data. Although some homologous regions remain in the flanking regions, the *Conservedness of probes* results clearly show that trimming flanking data can lead to nearly 100% variant or 0% invariant sites. This suggests that the flanking regions of targeted loci contain more ambiguous homology than previously thought. The presence of gappy regions is partly due to the absence of introns in transcriptomic data, as noted by Bossert et al. (2019) and illustrated in Figure 2. This is further supported by the fact that over 70% of the targeted loci are genic, and nearly 23% of the targeted loci in the joined probe set are exonic-intronic. Additionally, our analyses show that more loci were retained using targeted core regions, and that *trimAl* did not reduce matrix length to the extent of removing loci. Phylogenomic studies focusing on deep-time evolutionary scales or integrating transcriptomes with capture data may benefit from completely excluding flanking regions.

We demonstrate that there is unintended ancestry in the targeted genomic regions shared by the two capture approaches used in Adephaga. This is significant, as selecting or creating probe sets tailored to a specific question is critical in phylogenomic projects (discussed in Gustafson et al., 2023). UCE probes are theoretically designed to target deeply conserved regions of the genome, with the frequency of variable sites increasing further from the core regions (Faircloth et al., 2012; McCormack et al., 2013). In contrast, ExC probes are theoretically designed to target less conserved regions capable of resolving shallower evolutionary relationships with fewer loci (<1000) (Lemmon et al., 2012; McCormack et al., 2013). Here, we further demonstrate that Insecta UCEs are genic in nature; much more than initially anticipated. Moreover, the overlap of UCE and ExC probes suggests that either UCE loci are less conserved than previously thought, and/or that ExC loci are more conserved. Our findings show that the ExC and UCE probe kits exhibit distinct but overlapping levels of conservedness compared to one another (Figure 6). By integrating these probe sets, we retain these varying levels of conservedness into the *in silico* probe set, potentially offering the advantages of both methods: targeting both more and less conserved regions of the genome. While further research into the design and conservedness of these data generation strategies is needed, we have shown that the overlap between these two approaches results in backwards compatible loci (e.g., retrocompatible), where ExC and UCE probes target the same genomic regions and features.

The methods we employed to integrate diverse genomic data for increased taxon sampling, allowed for the inclusion of five additional fossils to estimate the age of the Adephaga tree. The resulting divergence times represent an increase in the age of important crown groups (Table 4). For example, the median age of beetles (order Coleoptera), was pushed back nearly 20–30 million years compared to recent estimates (Baca et al., 2021; McKenna et al., 2019). Such discrepancies are expected, because augmenting the taxon sampling allows the placement of fossils deeper in the tree, thereby applying the fossil associated minimum ages to more derived nodes. The age of Coleoptera was not the focus of this study and other suborders were not adequately sampled to accurately estimate its age (See Table 4). Furthermore, more investigation into the role of maximum ages of prior distributions need to be conducted to contain the ages of Coleoptera as a whole. Regardless, we show that the phylogenomic data available for Adephaga is not yet reflective enough of the true composition of this suborder. For example, Geadephaga represents nearly 90% of Adephaga species diversity, yet this diversity is still underrepresented in currently available data. This lineage (primarily represented by the family Carabidae) is poorly represented in the fossil record, making estimates for this group challenging (see discussion in Baca et al., 2021). Further sampling of missing subfamilies and tribes, aided by the development of the newly introduced UCE probe kit, should improve the incorporation of additional fossil calibrations in the future.

Current hypotheses of Adephaga relationships

The suborder Adephaga has been recognized as a monophylum early on, based on several morphological synapomorphies (Beutel et al., 2008; Beutel & Roughley, 1988; Lawrence et al., 2011), whereas phylogenomic relationships among families have proven more challenging to address until now. The classification and phylogeny of families within Adephaga have quickly and substantially evolved in the past decades. Earlier studies suggested the monophyly of Hydradephaga or the family-level ranking of several lineages (e.g., Paussidae, Rhysodidae), based most frequently on morphological features (Crowson, 1960). However, the progressive introduction of DNA-based phylogenetic inferences has revolutionized our understanding of Adephaga evolutionary history (Baca, Alexander, et al., 2017; Baca et al., 2021; Gustafson et al., 2020; López-López & Vogler, 2017; Maddison et al., 1999, 2009; Ribera et al., 2002; Shull et al., 2001; Toussaint et al., 2016; Vasilikopoulos et al., 2021). In parallel, the combination of Hennigian morphological phylogenetics with increasingly refined morphological character matrices of Adephaga has allowed a convergence to relatively stable family-level relationships (Beutel, 1993; Beutel et al., 2006, 2013; Beutel, Liu, et al., 2020;

Beutel & Roughley, 1988). Despite this progress, several questions remain open with respect to inter- and intra-familial placements across Adephaga. Our supertree approach, based on the cumulative efforts of independent research groups over the past decade, allows to substantially expand the taxon sampling for all Adephaga main lineages and offers new insights into our understanding of the group classification and evolutionary history.

Gyrinoidea

Whirligig beetles (Gyrinidae) are recovered as sister to the rest of Adephaga, with strong support coming from the core region sequence data (Figure 4). This placement is in line with the current treatment of this lineage as a superfamily (Gyrinoidea), with respect to the rest of the Adephaga tree of life (Figures 4 and 5). Our results confirm, once more, the paraphyly of Hydradephaga, with Gyrinidae recovered as sister to both Geadephaga and the remaining aquatic families (i.e., Dytiscoidea; Baca, Alexander, et al., 2017; Baca et al., 2021; Beutel, 1993; Beutel et al., 2006, 2013; Beutel, Liu, et al., 2020; Beutel & Roughley, 1988; Gustafson et al., 2020; McKenna et al., 2019; Vasilikopoulos et al., 2021). The spurious relationship of Gyrinidae + Haliplidae recovered in earlier studies (e.g., Shull et al., 2001; see S.-Q. Zhang et al., 2018), is strongly rejected here in agreement with previous studies based on independent datasets (references herein). Relationships and ages recovered within Gyrinidae (Figure 5) are consistent with the tip-dated phylogeny of the family by Gustafson et al. (2017), and the previous time-calibrated, phylogenomic study of Adephaga by Baca et al. (2021). Within Gyrinidae, we infer Spanglerogyrinae (a single extant species endemic to southern USA) as sister to Gyrininae + Heterogyrinae (a single extant species endemic to Madagascar), while within Gyrininae, Gyrinini is sister to reciprocally monophyletic Dineutini and Orectochilini. The origin of modern whirligig beetles is dated back to the Permian *ca.* 257 Ma (Table 4), the split between Heterogyrinae and Gyrininae dated back to the early Jurassic *ca.* 190 Ma and the origin of modern Gyrininae dated back to the mid-Jurassic *ca.* 167 Ma.

Haliploidea

We infer Haliplidae as sister to all remaining aquatic families (i.e., the Dytiscoidea) which is in line with most recent molecular phylogenomic studies and morphological inferences (Baca et al., 2021; Baca, Toussaint, et al., 2017; Beutel et al., 2013; Beutel, Liu, et al., 2020; Gustafson et al., 2020; McKenna et al., 2019; Toussaint et al., 2016; Vasilikopoulos et al., 2021). The recognition of this evolutionary branch as a superfamily (i.e., Haliploidea) is less obvious than before. Perhaps a more inclusive definition of Dytiscoidea that would include Haliplidae is desirable and would better depict the existence of two aquatic monophylums in Adephaga. The two currently recognized subfamilies Haliplinae and Peltodytinae (van Vondel, 2021) are recovered as sister and reciprocally monophyletic. Genera within Haliplidae were all recovered as monophyletic but a large-scale phylogeny of the family is still lacking despite ongoing work to refine the classification of the family (van Vondel, 2017, 2019, 2021). The phylogenomic relationships inferred in this study are in line with the only recent morphological phylogeny of the group (van Vondel, 2021), indicating that the current classification of Haliplidae is robust. The origin of modern Haliplidae is dated back to the Jurassic *ca.* 176 Ma (Table 4), while the origin of modern Haliplinae is dated back to the early Cretaceous *ca.* 133 Ma.

Dytiscoidea

The monophyly of Dytiscoidea is inferred with maximal support and confirms the placement of this lineage as sister to Haliploidea, both forming the second aquatic monophylum within Adephaga (Figures 4 and 5). Within Dytiscoidea, we recover Meruidae (the unique species *Meru phylissae* dwells in waterfall hygropetric ecosystems of southern Venezuela, P. J. Spangler & Steiner, 2005) as sister to Noteridae, in line with all molecular studies to date (Baca, Alexander, et al., 2017; Baca et al., 2021; Baca, Toussaint, et al., 2017; Balke et al., 2008; Beutel et al., 2006, 2013; Beutel, Liu, et al., 2020; Toussaint et al., 2016). The split between Meruidae and Noteridae is estimated to have occurred during the Triassic *ca.* 226 Ma while the origin of modern Noteridae is dated back to the Jurassic *ca.* 175 Ma in a somewhat synchrony with that of Haliplidae. The phylogenomic relationships within Noteridae are consistent with those of Baca et al. (2017) in recovering Notomicrinae (including Phreatodytini) as sister to Noterinae.

Two major discrepancies in Adephaga phylogenetics have remained largely debated in the past decade. First, the placement of Hygrobiidae has been inconsistently recovered as (1) sister to Aspidytidae and Amphizoidae, (2) sister to Dytiscidae, or (3) sister to Aspidytidae, Amphizoidae and Dytiscidae (reviewed in Gustafson et al., 2021). Second, the monophyly or paraphyly of the water cliff beetle family Aspidytidae have both been proposed in different studies (Baca et al., 2021; Toussaint et al., 2016; Vasilikopoulos et al., 2019, 2021). In this study, we consistently recover Hygrobiidae as sister to Aspidytidae and Amphizoidae, in line with some recent studies (Baca et al., 2021; Gustafson et al., 2020; Toussaint et al., 2016; Vasilikopoulos et al., 2019, 2021), and supporting the idea that the presence of prothoracic glands, a suggested synapomorphy grouping Hygrobiidae and Dytiscidae (Beutel et al., 2006, 2013; Beutel, Liu, et al., 2020), has evolved convergently (Forsyth, 1970). As emphasized in Gustafson et al. (2021), these glands are quite distinct between the two lineages and a closer investigation of morphological structures supports a scenario of evolutionary convergence. We also infer Aspidytidae as paraphyletic as recovered in some studies (e.g., Baca et al., 2021; Toussaint et al., 2016) and in contradiction with others (Vasilikopoulos et al., 2019, 2021). The family consists of two monotypic genera, *Aspidytes* from South Africa and *Sinaspidytes* from China with a similar ecology, adults and larvae being found in hygropetric habitats. These two species are similar in appearance and are classically united by the presence of a strongly shortened pedicellus enclosed in part by the globular distal part of the scapus (Balke et al., 2005; Toussaint et al., 2016). However other charac-

ters proposed to unite the two species are currently considered plesiomorphic (e.g., metacoxal plates). Furthermore, the morphology of Aspidytidae larvae is very similar to that of Amphizoidae, their sister lineage in our phylogenomic inference (Alarie & Bilton, 2005; Toussaint et al., 2016). It is likely that only genome-scale data will finally allow tackling the placement of both *Aspidytes* and *Sinaspidytes* within this clade. We infer a split between Hygrobiidae and the remaining families of this clade in the late Triassic *ca.* 208 Ma, with a subsequent cladogenetic event between *Aspidytes* and *Synaspidytes* + Amphizoidae in the Jurassic *ca.* 175 Ma.

Higher-level relationships within the most diverse aquatic family, Dytiscidae, have been largely unresolved and differed strongly between analyses (see Baca et al., 2021; Gustafson et al., 2020; Vasilikopoulos et al., 2021). The currently recognized subfamilies are supported as being monophyletic based on molecular data (Baca et al., 2021; Gustafson et al., 2020; Miller & Bergsten, 2014; Vasilikopoulos et al., 2021), with most receiving similar support through analysis of larval morphology (Michat et al., 2017). However, the only consistent grouping among subfamilies is the monotypic Hydrodytinae as sister to the diverse Hydroporinae (Baca et al., 2021; Gustafson et al., 2020; Miller & Bergsten, 2014; Vasilikopoulos et al., 2021). Genomic-scale data also groups Agabinae + Colymbetinae and Dytiscinae + Cybistrinae together in a clade, as recovered here (Figure 4; Baca et al., 2021; Gustafson et al., 2020; Vasilikopoulos et al., 2021). An important difference among analyses is the placement of the subfamily Matinae. Matinae was previously placed as sister to all other Dytiscidae (Désamoré et al., 2018; Miller, 2001; Miller & Bergsten, 2014). Here, Matinae is recovered well inside Dytiscidae as part of a clade including Agabinae + Colymbetinae, with low support in analyses of the core region only (Figure 4). A similar placement was recovered in Gustafson et al. (2020) and Baca et al. (2021). Vasilikopoulos et al. (2021) similarly inferred a nested position for Matinae, but as sister to the Hydrodytinae + Hydroporinae clade instead. This same study placed the clade Coptotominae + Lancetinae as the sister lineage to all other Dytiscidae. Our current dataset, and those based largely on UCEs, have placed this clade sister to Laccophilinae, rather than all other lineages. The present study, representing the most comprehensive phylogenomic analysis in terms of taxon sampling and data types, infers a different dytiscid tree from the ExC tree of Vasilikopoulos et al. (2021), comprising two reciprocally monophyletic groups: one consisting of Hydrodytinae + Hydroporinae; and the other containing all other subfamilies (Figure 4). While this pattern was also inferred in Gustafson et al. (2020) using UCE data alone, it was not recovered in Baca et al. (2021). Additionally, taxon sampling for Dytiscidae in the present study and others utilizing genomic-scale data, are limited relative to the known diversity of this group. Therefore, the major branching pattern within Dytiscidae remains unclear and requires future investigation with extensive taxon sampling. We estimate the origin of modern Dytiscidae at the end of the Triassic *ca.* 208 Ma, an age predating previous estimates (e.g., Baca et al., 2021; Désamoré et al., 2018), with potentially important implications for our understanding of diving beetle evolutionary history.

Caraboidea/Geadephaga

We infer Caraboidea (= Geadephaga) as a monophylum, and our phylogenomic tree further clarifies the relationships and temporal origin of many constituent subfamilies and tribes, the placement and evolution of which have been debated for at least a century (reviewed in part in Bousquet, 2012). At the family level, we infer Trachypachidae as sister to Cicindelidae and Carabidae with strong branch support (Figure 4). Despite earlier suggestions that Trachypachidae could be related to Hydradephaga (e.g., Beutel & Roughley, 1988; Hammond, 1979), it was also postulated that this family was likely more closely related to Carabidae (Beutel & Haas, 1996; Kavanaugh, 1986). The monophyly of terrestrial Adephaga families (i.e., Caraboidea/Geadephaga) has since received increasing support from recent morphological and molecular studies, although internal phylogenetic relationships have remained, to some extent, unresolved (e.g., Maddison et al., 2009; Shull et al., 2001). The placement of Trachypachidae in our phylogenomic tree is in line with the most recent studies and is currently undisputed (Baca et al., 2021; Beutel, Liu, et al., 2020; McKenna et al., 2019; Vasilikopoulos et al., 2021). This family consists of six species placed in two extant genera, the South American *Systolosoma* and Holarctic *Trachypachus*, the former not yet placed in a phylogenomic framework although the monophyly of Trachypachidae is supported by both morphology and Sanger sequencing data (Beutel, Liu, et al., 2020; Maddison et al., 2009). We estimate that Trachypachidae diverged from their sister group in the Permian *ca.* 236 Ma, but the origin of the modern crown group and divergence between genera remains unexplored.

The placement of Cicindelidae as sister to Carabidae is consistent with other phylogenomic studies and in line with morphology (Baca et al., 2021; Beutel, Liu, et al., 2020; Gough et al., 2020; McKenna et al., 2019; Vasilikopoulos et al., 2021). Even though the family status of Cicindelidae has been challenged in the past, it is strongly supported by genomic data and the unique morphological characters of this group (Duran & Gough, 2020; Gough et al., 2019; Putchkov & Cassola, 2005). Within Cicindelidae, we infer Manticorini as sister to the rest of the family, in agreement with previous studies (Baca et al., 2021; Galián et al., 2002; Gough et al., 2019, 2020; Vasilikopoulos et al., 2021). Future studies should aim at sampling all recognized genera within this tribe to test the relationships of constituent lineages. We recover the tribe Collyridini as the next lineage branching off the topology, and subsequently *Ctenostoma* (Ctenostomatini) as sister to *Tetracha* (Megacephalini), two unexpected placements in contradiction with earlier studies in which Ctenostomatini are either inferred as sister to Cicindelini and Oxycheilini (Gough et al., 2019, 2020), or as sister to Collyridini (Galián et al., 2002; Vogler & Barraclough, 1998). The placements of Ctenostomatini, Collyridini, Megacephalini and Oxycheilini require additional scrutiny with expanded taxon sampling. Similarly, we infer

Therates (Cicindelini) as sister to Oxycheilini, a result consistent with Vasilikopoulos et al. (2021) but in contradiction with earlier studies based on multilocus data (Galián et al., 2002; Gough et al., 2019; Vogler & Barraclough, 1998). We estimate the origin of modern tiger beetles in the Jurassic *ca.* 166 Ma, largely predating the earliest fossil record for the family (Schmidt et al., 2023). Following an early interest in their molecular phylogeny (Galián et al., 2002; Vogler & Barraclough, 1998; Vogler & Pearson, 1996; Zerm et al., 2007), tiger beetles have subsequently been overlooked for more than a decade, despite their remarkably conserved morphology, diverse ecologies, and striking adaptations (e.g., bat echolocation hearing organs and responses; H. G. Spangler, 1988; Yager et al., 2000). With a limited fossil record and reduced phylogenomic coverage, our understanding of tiger beetle evolution is currently hampered, and future studies should aim at expanding the taxonomic coverage of this family.

Within Carabidae, we recover relatively robust phylogenomic placements of most currently accepted subfamilies across two major clades (Figures 4 and 5). The first clade in the Carabidae phylogeny, is dated back from the Triassic *ca.* 207 Ma, and includes three subclades comprising multiple subfamilies, some of which are the most species-rich in the family (e.g., Trechinae). The first subclade comprises the Broscinae, Elaphrinae and Scaritinae, three subfamilies with no obvious morphological ties. The Elaphrinae and Scaritinae were recovered as sister by Baca et al. (2021) that did not include Broscinae, while Vasilikopoulos et al. (2021) recovered Scaritinae outside of a clade comprising Broscinae and Elaphrinae. This area of the tree requires much deeper sampling to be resolved in the future. Nevertheless, the Scaritinae are inferred as monophyletic with Clivini recovered as sister to Scaritini + Pasimachini, in line with Hogan (2012). We infer an origin of Scaritinae in the Jurassic *ca.* 173 Ma, but understanding the evolutionary history of this subfamily will require a refined placement among Carabidae. In a second subclade, we infer the Loricerinae as sister to the Omophroninae, two species-poor subfamilies represented by a single genus, with surprisingly widespread geographic ranges. This relationship was also recovered in Vasilikopoulos et al. (2021). The robust placement of these subfamilies is key with regards to fossils recently described from mid-Cretaceous Kachin amber in both lineages and used for divergence time estimations (H. Liu et al., 2023; Rosová et al., 2023). The third subclade infers the Trechinae as sister to the Harpalinae and Brachininae, a robust placement recovered in all recent phylogenomic studies (Baca et al., 2021; Kim et al., 2024; Raupach et al., 2022; Vasilikopoulos et al., 2021), as well as to some extent in morphological studies (e.g., Beutel et al., 2008; Beutel, Liu, et al., 2020). Trechinae have been the focus of intense research into their systematics and evolution (e.g., Faille et al., 2021; Maddison et al., 2019; Maddison & Ober, 2011), yet no comprehensive dated evolutionary tree of the subfamily has been inferred to date. Future taxon sampling will be key to refine the very tentative temporal estimates provided here (i.e., divergence time of Bembidiini + Pogonini at *ca.* 120 Ma). The inferred sister-relationship of Brachininae and Harpalinae is well-established in the literature. Although the phylogeny of Brachininae has been largely overlooked in the past, the systematics and evolution of Harpalinae have been the focus of some scrutiny (Ober, 2002; Ober & Heider, 2010; Ober & Maddison, 2008). The internal phylogenomic relationships within Harpalinae are largely in agreement with Vasilikopoulos et al. (2021) from which a majority of the taxa were used. Nevertheless, this very diverse group of ground beetles necessitates a much denser sampling to resolve the phylogenomic relationships among the numerous constituent tribes. The split between Brachininae and Harpalinae is dated back to the Jurassic *ca.* 158 Ma with an origin of modern Harpalinae in the early Cretaceous *ca.* 136 Ma in broad agreement with the earlier fossil-based estimates of Ober & Heider (2010).

The second major clade of Carabidae comprises two subclades. The first subclade is composed of several lineages previously recognized as separate families. For instance, Paussinae ant nest beetles have long been considered a distinct family among Adephaga due to unique morphological adaptations observed in myrmecophilous lineages (Crowson, 1960). Similarly, Rhysodinae wrinkled bark beetles were once considered a family (i.e., Rhysodidae), with uncertain placement among beetles, despite early work suggesting phylogenetic affinities with Carabidae (R. T. Bell & Bell, 1962). We infer Paussinae as sister to Rhysodinae, and within a larger clade that also comprises Siagoninae, although with moderate branch support. This clade was also recovered with moderate support by Vasilikopoulos et al. (2021) based on the analysis of ExC alone, although Paussinae were recovered as sister to Siagoninae. The respective placements of Paussinae and Rhysodinae therefore remain uncertain and additional taxon sampling of constituent tribes would help clarify this matter considering the long branches inferred for each family. The clade comprising Paussinae, Rhysodinae and Siagoninae is in turn recovered as sister to the Nebriinae. Finally, the second subclade is merely composed of the charismatic subfamily Carabinae. The placement of Carabinae within Carabidae has been challenging and various hypotheses have been proposed, usually with low branch support (Maddison et al., 1999, 2009; Ribera et al., 2005; Shull et al., 2001). Our phylogenomic inference is largely in agreement with Baca et al. (2021), although their taxon sampling was significantly less comprehensive, and is in contradiction with Vasilikopoulos et al. (2021) who recovered Carabinae in a phylogenomic grade along with most other subfamilies. The study of genome-scale phylogenomic relationships in Carabidae supports a close relationship between Carabinae and Nebriinae (Gauthier, Blanc, et al., 2025; Gauthier, Cardenas, et al., 2025), as do earlier studies based on morphological grounds (R. T. Bell, 1967; Beutel, 1992; Jeannel, 1940). An exhaustive genomic and taxonomic sampling of all subfamilies including as many tribal representatives as possible might be the only avenue to clearly identify the placement of Carabinae within Carabidae, as well as to disentangle the relationships of other enigmatic subfamilies (e.g., Apotominae, Cicindinae, Nototylinae).

With *ca.* 37,000 described species, ground beetles are the most diverse family of Adephaga and one of the most diverse beetle families overall. Despite an acute interest in their systematics and taxonomy, knowledge of their evolutionary history is trailing far behind, principally because a robust phylogenomic tree of the group is missing. There is a paucity of research in Carabidae phylogenetics using genomic resources (Gauthier, Blanc, et al., 2025; Gauthier, Cardenas, et al., 2025), and existing evolutionary trees based on few loci are not sufficient to explore evolutionary patterns and processes in this family due to unresolved placements and lack of branch support (Maddison et al., 1999, 2009; Ribera et al., 2005). Most future efforts in Adephaga phylogenomics should be directed to expand the depth of coverage at the subfamilial and tribal levels in Geadephaga.

To that end, we introduce a newly optimized UCE probe set named the GeaSub-2.9kv1. The optimization aspect being the probe to locus recovery ratio for the purposes of *in vitro* targeted capture and enrichment. The cost of probe synthesis increases with probe number, therefore this new development will aid phylogenomic investigations within Geadephaga by making probe synthesis more affordable, without impacting phylogenomic performance. Details of the newly introduced probe kit GeaSub-2.9kv1 can be found in supplemental materials on dryad: https://doi.org/10.5061/dryad.w9ghx3fzf.

Evolutionary history of Adephaga

A Permian origin is now well-established for crown Adephaga based on time-calibrated molecular phylogenies (Baca et al., 2021; McKenna et al., 2019; Toussaint et al., 2017) and fossil data, at least in part (Beutel et al., 2024; Boudinot et al., 2023; Ponomarenko & Prokin, 2015; Ponomarenko & Volkov, 2013). Some of these Permian Adephaga fossils have been assigned to extant families such as Gyrinidae (Beutel et al., 2024; Yan et al., 2018) and Trachypachidae (Kirejtshuk & Ansorge, 2023). However, the placement of the former has been contested (Beutel et al., 2019; A. G. Kirejtshuk & Prokin, 2018), and all pre-Cenozoic fossil taxa in the latter belong to the extinct subfamily †Eodrominae, whose exact composition and relationship to Carabidae remain unclear (Kirejtshuk & Ansorge, 2023). The Triassic held the first major diversification event with the origins of the four superfamilies: Gyrinoidea, Caraboidea, Haliploidea, and Dytiscoidea (Baca et al., 2021; Figure 5). There is strong support for the monophyly of these groups and now consensus among phylogenomic studies regarding the early branching pattern of Adephaga with Gyrinoidea sister to the clade (Caraboidea (Haliploidea + Dytiscoidea); Baca et al., 2021; Gustafson et al., 2020; Vasilikopoulos et al., 2021).

Much thought has gone into the likely ancestral ground plan for Adephaga (R. Bell, 1966; Crowson, 1960; Evans, 1982). Given the consensus regarding the deepest splits in the Adephaga tree, and significant differences in the morphology, behavior, and ecology of adult and larval stages of aquatic adephagan lineages (Ribera et al., 2002; see Beutel et al., 2020 for review), numerous independent invasions of diverse aquatic habitats from the terrestrial environment are likely. With the exception of the algophagous Haliplidae and a few carabid taxa (e.g., Rhysodinae wrinkled bark beetles), all adephagans are carnivorous and largely predatory, particularly as larvae. Adult Adephaga also uniformly have their metacoxa fused to the metathorax to some degree, limiting their ability to swing the hind leg forward to at most 5° (Evans, 1977). Crowson (1955; see also Crowson, 1960) proposed that the ancestral Adephaga was likely a subcortical predator. Consistent with this, Evans (1977) argued that the metacoxal fusion would afford improved wedge-pushing ability, as would be required for a predatory beetle to hunt subcortically. Boudinot et al. (2023) also recently proposed that the smooth elytra in Adephaga was a ground-plan adaptation associated with the subcortical movement necessary for predation. Both the metacoxal fusion and smooth elytra would then serve as exaptations within Gyrinoidea and Dytiscoidea for swift swimming capabilities. However, in order to quantitatively infer the likely ancestral character states and timing of the acquisition of particular features, future studies will need to incorporate both fossil and extant crown taxa in a total-evidence framework. This will further allow outstanding questions regarding the monophyly of Aspidytidae in relation to the extinct †Liadytidae (Baca et al., 2021; Toussaint et al., 2016), and the proper placement of pre-Cenozoic fossil taxa to be addressed. By the Jurassic, most modern Adephaga families were present, with continued pulses of diversification occurring in both aquatic and terrestrial lineages throughout the Cretaceous and Cenozoic, culminating in the exceptional diversity of predatory ground and water beetles seen today.

Conclusions

We demonstrate that two targeted capture methods, transcriptomes, and genomes used to study Adephaga over the past two decades are compatible and can be integrated to build comprehensive genomic matrices. Most importantly, the two capture methods ExC and UCE data are indeed congruous. This is due in part to the exonic nature of many arthropod UCE loci. To infer robust evolutionary trees, we develop a bioinformatic pipeline that combines three existing probe sets, modify the workflow of existing softwares, compare the impact of trimming on matrix construction, along with highlighting the effect of data curation and partitioning. In doing so, we demonstrate the challenges faced by target capture approaches, particularly at deep evolutionary time scales. These challenges arise from the sequences flanking the genomic regions targeted by probes (i.e., the core region). These core conserved regions contain sufficient phylogenomic information to confidently ignore flanking regions and their associated challenges. With this integrative approach, we infer the most comprehensive fossil-based evolutionary tree of Adephaga to date. We recover Hygrobiidae as sister to Amphizoidae and a paraphyletic Aspidytidae, further supporting the evolutionary convergence of prothoracic glands in Hygrobiidae and Dytiscidae. Our fossil-calibrated tree supports the origin of stem Ade-

phaga in the Carboniferous and diversification of extant lineages in the mid-Permian. Through this work we find that the most diverse Adephaga family Carabidae remains undersampled, preventing a better understanding of its evolutionary history. To address this challenge, we introduce a new ultraconserved element probe set specifically tailored to Geadephaga beetles that will generate compatible genomic data. By demonstrating that a variety of legacy genomic datasets can be integrated, and by deploying new tools to generate additional genomic data and curate it, it is likely that the Adephaga tree of life will rapidly expand and reveal the underlying mechanisms of diversification in this fascinating group of beetles.

Funding

This study was funded by a Swiss National Science Foundation grant 310030_200491 to Emmanuel Toussaint.

Acknowledgements

We thank Jérémy Gauthier for insightful discussion and advice on bioinformatic analyses. We want to acknowledge Michael Balke, Udo Schmidt, and ukbeetles.co.uk for allowing the use of photographs in the figures of this study. We also thank the handling editor Bryan Carstens as well as reviewers for insightful comments on an earlier version of this study.

Supplementary materials

Supplemental materials include information for the taxa used in analyses, details of the phylogenomic inferences, locus curation and trimming statistics, gene concordance analysis results, dated analysis results, and linear mixed model analysis results. Supplemental files including analyses, figures, tables, and results of all phylogenetic inferences. Data and results of bioinformatic analyses are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.w9ghx3fzf. All scripts used in the bioinformatic pipeline are described and available at https://github.com/crcardenas/Adephaga_UCE.

Author Contributions

CRC, GTG & EFAT conceptualized this study

CRC curated the data, developed the methodology, led the analyses and drafted the original manuscript, tables and figures

GTG & EFAT contributed to the methodology, validated the research, and contributed to the original draft and writing—review and editing

EFAT contributed to the project administration and resources

Submitted: December 11, 2024 EDT. Accepted: March 21, 2025 EDT. Published: May 20, 2025 EDT.

References

Alarie, Y., & Bilton, D. T. (2005). Larval morphology of Aspidytidae (Coleoptera: Adephaga) and its phylogenetic implications. *Annals of the Entomological Society of America*, *98*(4), 417–430. https://doi.org/10.1603/0013-8746(2005)098%5B0417:LMOACA%5D2.0.CO;2

Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B., & Delsuc, F. (2020). MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Molecular Ecology Resources*, *20*(4), 892–905. https://doi.org/10.1111/1755-0998.13160

Altschul, S. F. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Baca, S. M., Alexander, A., Gustafson, G. T., & Short, A. E. Z. (2017). Ultraconserved elements show utility in phylogenetic inference of Adephaga (Coleoptera) and suggest paraphyly of 'Hydradephaga': Phylogeny of Adephaga inferred with UCEs. *Systematic Entomology*, *42*(4), 786–795. https://doi.org/10.1111/syen.12244

Baca, S. M., Gustafson, G. T., Alexander, A. M., Gough, H. M., & Toussaint, E. F. A. (2021). Integrative phylogenomics reveals a Permian origin of Adephaga beetles. *Systematic Entomology*, *46*(4), 968–990. https://doi.org/10.1111/syen.12506

Baca, S. M., Toussaint, E. F. A., Miller, K. B., & Short, A. E. Z. (2017). Molecular phylogeny of the aquatic beetle family Noteridae (Coleoptera: Adephaga) with an emphasis on data partitioning strategies. *Molecular Phylogenetics and Evolution*, *107*, 282–292. https://doi.org/10.1016/j.ympev.2016.10.016

Balke, M., Beigel, A., & Hendrich, L. (2010). *Hydroporus carstengroehni* sp. n. und zwei unbestimmte Hydroporinae aus dem baltischen Bernstein (Dytiscidae: Hydroporinae). *Nachrichtenblatt der Bayerischen Entomologen*, *59*, 2–9.

Balke, M., & Hendrich, L. (2019). †*Japanolaccophilus beatificus* sp. n. from Baltic amber and a key to the Laccophilinae genera of the World (Coleoptera: Laccophilinae). *Zootaxa*, *4567*(1), 176–182. https://doi.org/10.11646/zootaxa.4567.1.10

Balke, M., Ribera, I., & Beutel, R. G. (2005). The systematic position of Aspidytidae, the diversification of Dytiscoidea (Coleoptera, Adephaga) and the phylogenetic signal of third codon positions. *Journal of Zoological Systematics and Evolutionary Research*, *43*(3), 223–242. https://doi.org/10.1111/j.1439-0469.2005.00318.x

Balke, M., Ribera, I., Beutel, R. G., Viloria, 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*(6), 647–650. https://doi.org/10.1111/j.1463-6409.2008.00345.x

Barclay, M. V. L., Geiser, M., Vassiliades, D., Farrell, W. B., Cristóvão, J., Natural History Museum Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. (2023). The genome sequence of a ground beetle, *Pterostichus niger* (Schaller, 1783). *Wellcome Open Research*, *8*, 544. https://doi.org/10.12688/wellcomeopenres.20418.1

Barclay, M. V. L., Natural History Museum Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. (2023). The genome sequence of a ground beetle, *Leistus spinibarbis* (Fabricius, 1775). *Wellcome Open Research*, *8*, 412. https://doi.org/10.12688/wellcomeopenres.19997.1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bell, R. (1966). *Trachypachus* and the Origin of the Hydradephaga (Coleoptera). *The Coleopterists Bulletin*, *20*(4), 107–112. https://doi.org/10.5962/p.372392

Bell, R. T. (1967). Coxal cavities and the classification of the Adephaga (Coleoptera). *Annals of the Entomolological Society of America*, *60*(1), 101–107. https://doi.org/10.1093/aesa/60.1.101

Bell, R. T., & Bell, J. R. (1962). The taxonomic position of the Rhysodidae (Coleoptera). *The Coleopterists Bulletin*, *16*(4), 99–106. https://doi.org/10.5962/p.372317

Bergsten, J., & Miller, K. B. (2007). Phylogeny of Diving Beetles Reveals a Coevolutionary Arms Race between the Sexes. *PLoS ONE*, *2*(6), e522. https://doi.org/10.1371/journal.pone.0000522

Beutel, R. G. (1992). Phylogenetic analysis of thoracic structures of Carabidae (Coleoptera: Adephaga). *Journal of Zoological Systematics and Evolutionary Research*, *30*(1), 53–74. https://doi.org/10.1111/j.1439-0469.1992.tb00390.x

Beutel, R. G. (1993). Phylogenetic analysis of Adephaga (Coleoptera) based on characters of the larval head. *Systematic Entomology*, *18*(2), 127–147. https://doi.org/10.1111/j.1365-3113.1993.tb00658.x

Beutel, R. G. (1995). The Adephaga (Coleoptera): phylogeny and evolutionary history. In *Biology, phylogeny and classification of coleoptera: papers celebrating the 80th birthday of Roy A. Crowson* (pp. 173–217). Muzeum i Instytut Zoologii PAN.

Beutel, R. G. (1998). Trachypachidae and the phylogeny of Adephaga (Coleoptera). In G. E. Ball, A. Casale, A. V. Taglianti, & Museo regionale di scienze naturali (Turin, Italy) (Eds.), *Phylogeny and classification of Caraboidea (Coleoptera: Adephaga): proceedings of a symposium, 28 August 1996, Florence, Italy; XX International Congress of Entomology* (pp. 81–105). Museo regionale di scienze naturali.

Beutel, R. G., Balke, M., & Steiner, W. E., Jr. (2006). The systematic position of Meruidae (Coleoptera, Adephaga) and the phylogeny of the smaller aquatic adephagan beetle families. *Cladistics*, *22*(2), 102–131. https://doi.org/10.1111/j.1096-0031.2006.00092.x

Beutel, R. G., & Haas, A. (1996). Phylogenetic analysis of larval and adult characters of Adephaga (Coleoptera) using cladistic computer programs. *Insect Systematics & Evolution*, *27*(2), 197–205. https://doi.org/10.1163/187631296X00043

Beutel, R. G., Liu, Z., Fikáček, M., Ren, D., Pang, H., & Ślipiński, A. (2020). *Burmapseudomorphus planus* gen. et sp. nov. – a Late Cretaceous stem group member of the specialized Pseudomorphini (Carabidae, Coleoptera) from northern Myanmar. *Cretaceous Research*, *107*, 104274. https://doi.org/10.1016/j.cretres.2019.104274

Beutel, R. G., Ribera, I., & Bininda-Emonds, O. R. P. (2008). A genus-level supertree of Adephaga (Coleoptera). *Organisms Diversity & Evolution*, *7*(4), 255–269. https://doi.org/10.1016/j.ode.2006.05.003

Beutel, R. G., Ribera, I., Fikáček, M., Vasilikopoulos, A., Misof, B., & Balke, M. (2020). The morphological evolution of the Adephaga (Coleoptera). *Systematic Entomology*, *45*(2), 378–395. https://doi.org/10.1111/syen.12403

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*(5), 380–400. https://doi.org/10.1111/j.1439-0469.1988.tb00324.x

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*(2), 147–165. https://doi.org/10.1111/j.1096-0031.2012.00420.x

Beutel, R. G., Xu, C., Jarzembowski, E., Kundrata, R., Boudinot, B. E., McKenna, D. D., & Goczał, J. (2024). The evolutionary history of Coleoptera (Insecta) in the late Palaeozoic and theMesozoic. *Systematic Entomology*, *49*(3), 355–388. https://doi.org/10.1111/syen.12623

Beutel, R. G., Yan, E. V., & Lawrence, J. F. (2019). Phylogenetic methods applied to extinct beetles — the case of †Tunguskagyrus (Gyrinidae or †Triaplidae). *Palaeoentomology*, *2*(4), 372–380. https://doi.org/10.11646/palaeoentomology.2.4.11

Borowiec, M. L. (2016). AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ*, *4*, e1660. https://doi.org/10.7717/peerj.1660

Borowiec, M. L., Zhang, M. Y., Neves, K., Ramalho, M. O., Fisher, B. L., Lucky, A., & Moreau, C. S. (2025). Evaluating UCE data adequacy and integrating uncertainty in a comprehensive phylogeny of ants. *Systematic Biology*, syaf001. https://doi.org/10.1093/sysbio/syaf001

Bossert, S., Murray, E. A., Almeida, E. A. B., Brady, S. G., Blaimer, B. B., & Danforth, B. N. (2019). Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. *Molecular Phylogenetics and Evolution*, *130*, 121–131. https://doi.org/10.1016/j.ympev.2018.10.012

Boudinot, B. E., Yan, E. V., Prokop, J., Luo, X.-Z., & Beutel, R. G. (2023). Permian parallelisms: Reanalysis of †Tshekardocoleidae sheds light on the earliest evolution of the Coleoptera. *Systematic Entomology*, *48*, 69–96. https://doi.org/10.1111/syen.12562

Bousquet, Y. (2012). Catalogue of Geadephaga (Coleoptera, Adephaga) of America, north of Mexico. *ZooKeys*, *245*(245), 1–1722. https://doi.org/10.3897/zookeys.245.3416

Burmeister, E.-G. (1976). Der Ovipositor der Hydradephaga (Coleoptera) und seine phylogenetische Bedeutung unter besonderer Berücksichtigung der Dytiscidae. *Zoomorphologie*, *85*(3), 165–257. https://doi.org/10.1007/BF00993515

Cai, C., Short, A. E. Z., & Huang, D. (2012). The first skiff beetle (Coleoptera: Myxophaga: Hydroscaphidae) from Early Cretaceous Jehol Biota. *Journal of Paleontology*, *86*(1), 116–119. https://doi.org/10.1666/11-050.1

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Castresana, J. (2000). Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. *Molecular Biology and Evolution*, *17*(4), 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Chatzimanolis, S., Grimaldi, D. A., Engel, M. S., & Fraser, N. C. (2012). *Leehermania prorova*, the Earliest Staphyliniform Beetle, from the Late Triassic of Virginia (Coleoptera: Staphylinidae). *American Museum Novitates*, *2012*(3761), 1–28. https://doi.org/10.1206/3761.2

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chernomor, O., Von Haeseler, A., & Minh, B. Q. (2016). Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/sysbio/syw037

Crowley, L. M., Sudworth, J., University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. (2023). The genome sequence of a ground beetle, *Ophonus ardosiacus* (Lutshnik, 1922). *Wellcome Open Research*, *8*, 353. https://doi.org/10.12688/wellcomeopenres.19849.1

Crowley, L. M., Telfer, M. G., Escalona, H. E., University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life Management, Samples and Laboratory team, Wellcome Sanger Institute Scientific Operations: Sequencing Operations, Wellcome Sanger Institute Tree of Life Core Informatics team, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. (2024). The genome sequence of a ground beetle, *Agonum fuliginosum* (Panzer, 1809). *Wellcome Open Research*, *9*, 81. https://doi.org/10.12688/wellcomeopenres.20912.1

Crowley, L. M., University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. (2021). The genome sequence of the black clock beetle, *Pterostichus madidus* (Fabricius, 1775). *Wellcome Open Research*, *6*, 301. https://doi.org/10.12688/wellcomeopenres.17347.1

Crowley, L., University of Oxford and Wytham Woods Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, Garner, B., & Darwin Tree of Life Consortium. (2023). The genome sequence of a ground beetle, *Nebria brevicollis* (Fabricius, 1792). *Wellcome Open Research*, *8*, 20. https://doi.org/10.12688/wellcomeopenres.18749.1

Crowson, R. A. (1955). *The Natural Classification of the Families of Coleoptera*. Nathaniel Lloyd and Co., LTD.

Crowson, R. A. (1960). The Phylogeny of Coleoptera. *Annual Review of Entomology*, *5*, 111–134. https://doi.org/10.1146/annurev.en.05.010160.000551

Curole, J. P., & Kocher, T. D. (1999). Mitogenomics: digging deeper with complete mitochondrial genomes. *Trends in Ecology & Evolution*, *14*(10), 394–398. https://doi.org/10.1016/S0169-5347(99)01660-2

Darlington, P. J. (1943). Carabidae of mountains and islands: Data on the evolution of isolated faunas, and on atrophy of wings. *Ecological Monographs*, *13*(1), 37–61. https://doi.org/10.2307/1943589

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, *9*(5–6), 416–423. https://doi.org/10.1093/bfgp/elq031

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. https://doi.org/10.1111/mec.14492

Dettner, K. (1985). Ecological and Phylogenetic Significance of Defensive Compounds from Pygidial Glands of Hydradephaga (Coleoptera). *Proceedings of the Academy of Natural Sciences of Philadelphia*, *137*(1), 156–171.

Dos Reis, M., Donoghue, P. C. J., & Yang, Z. (2016). Bayesian molecular clock dating of species divergences in the genomics era. *Nature Reviews Genetics*, *17*(2), 71–80. https://doi.org/10.1038/nrg.2015.8

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*(4), 723–729. https://doi.org/10.1111/syen.12440

Eisner, T., & Aneshansley, D. J. (1999). Spray aiming in the bombardier beetle: Photographic evidence. *Proceedings of the National Academy of Sciences USA*, *96*(17), 9705–9709. https://doi.org/10.1073/pnas.96.17.9705

Evans, M. E. G. (1977). Locomotion in the Coleoptera Adephaga, especially Carabidae. *Journal of Zoology*, *181*(2), 189–226. https://doi.org/10.1111/j.1469-7998.1977.tb03237.x

Evans, M. E. G. (1982). Early evolution of the Adephaga—Some locomotor speculations. *The Coleopterists Bulletin*, *36*(4), 597–607.

Faille, A., Balart-García, P., Fresneda, J., Bourdeau, C., & Ribera, I. (2021). A remarkable new genus of Iberian troglobitic Trechodina (Coleoptera: Carabidae: Trechinae: Trechini), with a revisited molecular phylogeny of the subtribe. *Annales de La Société Entomologique de France (N.S.)*, *57*(2), 85–106. https://doi.org/10.1080/00379271.2021.1880339

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*(5), 786–788. https://doi.org/10.1093/bioinformatics/btv646

Faircloth, B. C. (2017). Identifying conserved genomic elements and designing universal bait sets to enrich them. *Methods in Ecology and Evolution*, *8*(9), 1103–1112. https://doi.org/10.1111/2041-210X.12754

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*(5), 717–726. https://doi.org/10.1093/sysbio/sys004

Fikáček, M., Beutel, R. G., Cai, C., Lawrence, J. F., Newton, A. F., Solodovnikov, A., Ślipiński, A., Thayer, M. K., & Yamamoto, S. (2020). Reliable placement of beetle fossils via phylogenetic analyses – Triassic Leehermania as a case study (Staphylinidae or Myxophaga?). *Systematic Entomology*, *45*(1), 175–187. https://doi.org/10.1111/syen.12386

Fikáček, M., Hájek, J., & Prokop, J. (2008). New records of the water beetles (Coleoptera: Dytiscidae, Hydrophilidae) from the central European Oligocene-Miocene deposits, with a confirmation of the generic attribution of *Hydrobiomorpha enspelense* Wedmann 2000. *Annales de La Société Entomologique de France (N,S.)*, *44*(2), 187–199. https://doi.org/10.1080/00379271.2008.10697555

Forsyth, D. J. (1970). The structure of the defence glands of the Cicindelidae, Amphizoidae, and Hygrobiidae (Insecta: Coleoptera). *Journal of Zoology*, *160*(1), 51–69. https://doi.org/10.1111/j.1469-7998.1970.tb02897.x

Frandsen, P. B., Holzenthal, R. W., Espeland, M., Breinholt, J., Thomas Thorpe, J. A., Simon, S., Kawahara, A. Y., Plotkin, D., Hotaling, S., Li, Y., Nelson, C. R., Niehuis, O., Mayer, C., Podsiadlowski, L., Donath, A., Misof, B., Lemmon, M. E., Lemmon, A., Morse, J. C., … Zhou, X. (2024). Phylogenomics recovers multiple origins of portable case making in caddisflies (Insecta: Trichoptera), nature's underwater architects. *Proceedings of the Royal Society B: Biological Sciences*, *291*(2026), 20240514. https://doi.org/10.1098/rspb.2024.0514

Fraser, N. C., Grimaldi, D. A., Axsmith, B. J., Heckert, A. B., Liutkus-Pierce, C., Smith, D., & Dooley, A. C. (2017). The Solite Quarry - a window into life by a late Triassic lake margin. In N. C. Fraser & H.-D. Sues (Eds.), *Terrestrial Conservation Lagerstätten: Windows into the Evolution of Life on Land* (pp. 105–130). Dunedin Academic Press. https://doi.org/10.2307/jj.12638994.7

Galián, J., Hogan, J. E., & Vogler, A. P. (2002). The origin of multiple sex chromosomes in Tiger Beetles. *Molecular Biology and Evolution*, *19*(10), 1792–1796. https://doi.org/10.1093/oxfordjournals.molbev.a004001

Gauthier, J., Blanc, M., & Toussaint, E. F. A. (2025). Chromosome-scale genomes of the flightless Caterpillar Hunter Beetles *Calosoma tepidum* and *Calosoma wilkesii* from British Columbia (Coleoptera: Carabidae). *Genome Biology and Evolution*, *17*(1), evae247. https://doi.org/10.1093/gbe/evae247

Gauthier, J., Cardenas, C. R., Nari, M., Gillett, C. P. D. T., & Toussaint, E. F. A. (2025). Draft genome of the endemic alpine ground beetle *Carabus (Platycarabus) depressus* (Coleoptera: Carabidae) from long-read sequencing of a frozen archived specimen. *G3 Genes Genomes Genetics*, *15*(5), jkaf027. https://doi.org/10.1093/g3journal/jkaf027

Gómez, R. A., & Damgaard, A. L. (2014). A rare diving beetle from Baltic amber: *Hydrotrupes prometheus* new species reveals former widespread distribution of the genus (Coleoptera, Dytiscidae). *Journal of Paleontology*, *88*(4), 814–822. https://doi.org/10.1666/13-017

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*(3), 740–751. https://doi.org/10.1093/biolinnean/blz195

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*(2), 305–321. https://doi.org/10.1111/syen.12324

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*(3), 307–321. https://doi.org/10.1093/sysbio/syq010

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*(12), 6933–6948. https://doi.org/10.1002/ece3.5260

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*(3), 552–570. https://doi.org/10.1111/syen.12413

Gustafson, G. T., Glynn, R. D., Short, A. E. Z., Tarasov, S., & Gunter, N. L. (2023). To design, or not to design? Comparison of beetle ultraconserved element probe set utility based on phylogenetic distance, breadth, and method of probe design. *Insect Systematics and Diversity*, *7*(4), 4. https://doi.org/10.1093/isd/ixad014

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*(3), 473–486. https://doi.org/10.1111/syen.12471

Gustafson, G. T., Prokin, A. A., Bukontaite, R., Bergsten, J., & Miller, K. B. (2017). Tip-dated phylogeny of whirligig beetles reveals ancient lineage surviving on Madagascar. *Scientific Reports*, *7*(1), 8619. https://doi.org/10.1038/s41598-017-08403-1

Hammond, P. M. (1979). Wing-folding Mechanisms of Beetles, with Special Reference to Investigations of Adephagan Phylogeny (Coleoptera). In T. L. Erwin, G. E. Ball, D. R. Whitehead, & A. L. Halpern (Eds.), *Carabid Beetles* (pp. 113–180). Springer Netherlands. https://doi.org/10.1007/978-94-009-9628-1_7

Hendrich, L., & Balke, M. (2020). A Baltic amber species of the diving beetle genus *Coptotomus* Say, 1830 (Coleoptera: Dytiscidae: Coptotominae). *Zootaxa*, *4895*(2), 285–290. https://doi.org/10.11646/zootaxa.4895.2.7

Henríquez-Piskulich, P., Hugall, A. F., & Stuart-Fox, D. (2024). A supermatrix phylogeny of the world's bees (Hymenoptera: Anthophila). *Molecular Phylogenetics and Evolution*, *190*, 107963. https://doi.org/10.1016/j.ympev.2023.107963

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Hogan, J. (2012). *Taxonomy, systematics and biogeography of the Scaritinae (Insecta, Coleoptera, Carabidae)* [Doctoral Dissertation]. Oxford Brookes University.

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, *65*(3), 357–365. https://doi.org/10.1093/sysbio/syu046

Inkscape Project. (2022). *Inkscape*.

Jeannel, R. (1940). Les Calosomes. *Mémoires du Muséum national d'Histoire naturelle*, *13*(1), 1–240.

Jiang, W., Chen, S.-Y., Wang, H., Li, D.-Z., & Wiens, J. J. (2014). Should genes with missing data be excluded from phylogenetic analyses? *Molecular Phylogenetics and Evolution*, *80*, 308–318. https://doi.org/10.1016/j.ympev.2014.08.006

Johnson, K. P., Dietrich, C. H., Friedrich, F., Beutel, R. G., Wipfler, B., Peters, R. S., Allen, J. M., Petersen, M., Donath, A., Walden, K. K. O., Kozlov, A. M., Podsiadlowski, L., Mayer, C., Meusemann, K., Vasilikopoulos, A., Waterhouse, R. M., Cameron, S. L., Weirauch, C., Swanson, D. R., … Yoshizawa, K. (2018). Phylogenomics and the evolution of hemipteroid insects. *Proceedings of the National Academy of Sciences USA*, *115*(50), 12775–12780. https://doi.org/10.1073/pnas.1815820115

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Kataev, B. M., Kirejtshuk, A. G., Manukyan, A. R., & Anokhin, B. A. (2019). *Kryzhanovskiana olegi* gen. et sp. nov., a remarkable eyeless representative of the tribe *Metriini* (Coleoptera: Carabidae: Paussinae) from Upper Cretaceous amber of northern Myanmar. *Cretaceous Research*, *103*, 104168. https://doi.org/10.1016/j.cretres.2019.06.014

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software bersion 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

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*(4), 67–109.

Kawahara, A. Y., Plotkin, D., Espeland, M., Meusemann, K., Toussaint, E. F. A., Donath, A., Gimnich, F., Frandsen, P. B., Zwick, A., Dos Reis, M., Barber, J. R., Peters, R. S., Liu, S., Zhou, X., Mayer, C., Podsiadlowski, L., Storer, C., Yack, J. E., Misof, B., & Breinholt, J. W. (2019). Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proceedings of the National Academy of Sciences USA*, *116*(45), 22657–22663. https://doi.org/10.1073/pnas.1907847116

Kim, D., Kim, G., Shin, C. R., Park, B., Choi, E. H., & Hwang, U. W. (2024). The complete mitochondrial genome of a ground beetle *Synuchus nitidus* (Carabidae: Harpalinae: Sphodrini) from South Korea. *Mitochondrial DNA Part B*, *9*(6), 711–715. https://doi.org/10.1080/23802359.2024.2361689

Kirejtshuk, A. G., & Ansorge, J. (2023). An extraordinarily preserved new genus and species of Trachypachidae (Coleoptera, Adephaga) from the Early Jurassic of Germany and a review of fossil trachypachid genera. *Historical Biology*, *35*(6), 958–975. https://doi.org/10.1080/08912963.2022.2071709

Kirejtshuk, A. G., Poschmann, M., Prokop, J., Garrouste, R., & Nel, A. (2014). Evolution of the elytral venation and structural adaptations in the oldest Palaeozoic beetles (Insecta: Coleoptera: Tshekardocoleidae). *Journal of Systematic Palaeontology*, *12*(5), 575–600. https://doi.org/10.1080/14772019.2013.821530

Kirejtshuk, A. G., & Prokin, A. A. (2018). The position of the Palaeozoic genus *Tunguskagyrus* Yan, Beutel et Lawrence in the family Triaplidae sensu n. (Coleoptera, Archostemata: Schizophoroidea). *Entomological Review*, *98*(7), 872–882. https://doi.org/10.1134/S0013873818070084

Kirejtshuk, A., Nel, A., & Kirejtshuk, P. A. (2016). Taxonomy of the reticulate beetles of the Subfamily Cupedinae (Coleoptera: Archostemata) with a review of their historical development. *Invertebrate Zoology*, *13*, 61–190. https://doi.org/10.15298/invertzool.13.2.01

Klausnitzer, B. (2003). Käferlarven (Insecta: Coleoptera) in Baltischem Bernstein – Möglichkeiten und Grenzen der Bestimmung. *Entomologische Abhandlungen*, *61*(1), 103–108.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., & Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evolutionary Biology*, *14*(1), 82. https://doi.org/10.1186/1471-2148-14-82

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*(3), 772–773. https://doi.org/10.1093/molbev/msw260

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), 1–217. https://doi.org/10.3161/000345411X576725

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*(1), 130–145. https://doi.org/10.1093/sysbio/syp017

Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, *61*(5), 727–744. https://doi.org/10.1093/sysbio/sys049

Lenth, R. V. (2017). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.10.3). https://doi.org/10.32614/CRAN.package.emmeans

Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, *52*(W1), W78–W82. https://doi.org/10.1093/nar/gkae268

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liang, Z., Qi, Z., Chen, J., & Jia, F. (2020). *Cretodineutus rotundus* gen. et sp. nov., the oldest adult whirligig beetle from the Upper Cretaceous of Myanmar (Coleoptera, Gyrinidae, Gyrininae). *Cretaceous Research*, *106*, 104251. https://doi.org/10.1016/j.cretres.2019.104251

Liu, D., Cui, J., Liu, Y., Niu, M., Wang, F., Zhao, Q., Cai, B., Zhang, H., & Wei, J. (2024). Ultraconserved elements from transcriptome and genome data provide insight into the phylogenomics of Sternorrhyncha (Insecta: Hemiptera). *Cladistics*, *40*, cla.12585. https://doi.org/10.1111/cla.12585

Liu, H., Makarov, K. V., Jarzembowski, E. A., Xiao, C., & Luo, C. (2023). *Cretoloricera electra* gen. et sp. nov., the oldest record of Loricerini (Coleoptera: Adephaga: Carabidae: Loricerinae) from mid-Cretaceous Kachin amber. *Cretaceous Research*, *148*, 105540. https://doi.org/10.1016/j.cretres.2023.105540

López-López, A., & Vogler, A. P. (2017). The mitogenome phylogeny of Adephaga (Coleoptera). *Molecular Phylogenetics and Evolution*, *114*, 166–174. https://doi.org/10.1016/j.ympev.2017.06.009

Lorenz, W. (2021). Carabcat database. In *Catalogue of Life Checklist*. http://www.carabidfauna2.net

Maddison, D. R., Baker, M. D., & Ober, K. A. (1999). Phylogeny of carabid beetles as inferred from 18S ribosomal DNA (Coleoptera: Carabidae). *Systematic Entomology*, *24*(2), 103–138. https://doi.org/10.1046/j.1365-3113.1999.00088.x

Maddison, D. R., Kanda, K., Boyd, O. F., Faille, A., Porch, N., Erwin, T. L., & Roig-Juñent, S. (2019). Phylogeny of the beetle supertribe Trechitae (Coleoptera: Carabidae): Unexpected clades, isolated lineages, and morphological convergence. *Molecular Phylogenetics and Evolution*, *132*, 151–176. https://doi.org/10.1016/j.ympev.2018.11.006

Maddison, D. R., Moore, W., Baker, M. D., Ellis, T. M., Ober, K. A., Cannone, J. J., & Gutell, R. R. (2009). (Coleoptera: Carabidae and Trachypachidae). *Zoologica Scripta*, *38*(1), 43–62. https://doi.org/10.1111/j.1463-6409.2008.00359.x

Maddison, D. R., & Ober, K. (2011). Phylogeny of minute carabid beetles and their relatives based upon DNA sequence data (Coleoptera, Carabidae, Trechitae). *ZooKeys*, *147*, 229–260. https://doi.org/10.3897/zookeys.147.1871

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*(3), 649–671. https://doi.org/10.1111/syen.12481

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*(2), 526–538. https://doi.org/10.1016/j.ympev.2011.12.007

McGettigan, P. A. (2013). Transcriptomics in the RNA-seq era. *Current Opinion in Chemical Biology*, *17*(1), 4–11. https://doi.org/10.1016/j.cbpa.2012.12.008

McKenna, D. D., Shin, S., Ahrens, D., Balke, M., Beza-Beza, C., Clarke, D. J., Donath, A., Escalona, H. E., Friedrich, F., Letsch, H., Liu, S., Maddison, D., Mayer, C., Misof, B., Murin, P. J., Niehuis, O., Peters, R. S., Podsiadlowski, L., Pohl, H., ... Beutel, R. G. (2019). The evolution and genomic basis of beetle diversity. *Proceedings of the National Academy of Sciences USA*, *116*(49), 24729–24737. https://doi.org/10.1073/pnas.1909655116

McKenna, D. D., Wild, A. L., Kanda, K., Bellamy, C. L., Beutel, R. G., Caterino, M. S., Farnum, C. W., Hawks, D. C., Ivie, M. A., Jameson, M. L., Leschen, R. A. B., Marvaldi, A. E., Mchugh, J. V., Newton, A. F., Robertson, J. A., Thayer, M. K., Whiting, M. F., Lawrence, J. F., Ślipiński, A., ... Farrell, B. D. (2015). The beetle tree of life reveals that Coleoptera survived end-Permian mass extinction to diversify during the Cretaceous terrestrial revolution. *Systematic Entomology*, *40*(4), 835–880. https://doi.org/10.1111/syen.12132

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. https://doi.org/10.1111/syen.12243

Michat, M. C., & Torres, P. L. M. (2011). Phylogenetic relationships of the tribe Vatellini based on larval morphology, with description of *Derovatellus lentus* (Coleoptera: Dytiscidae: Hydroporinae). *Annals of the Entomological Society of America*, *104*(5), 863–877. https://doi.org/10.1603/AN11054

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*(1), 45–92. https://doi.org/10.1163/187631201X00029

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*(4), 706–746. https://doi.org/10.1111/j.1365-3113.2012.00640.x

Miller, K. B., & Bergsten, J. (2014). The phylogeny and classification of predaceous diving beetles. In D. A. Yee (Ed.), *Ecology, Systematics, and the Natural History of Predaceous Diving Beetles (Coleoptera: Dytiscidae)* (pp. 49–172). Springer. https://doi.org/10.1007/978-94-017-9109-0_3

Minh, B. Q., Hahn, M. W., & Lanfear, R. (2020). New Methods to Calculate Concordance Factors for phylogenomic datasets. *Molecular Biology and Evolution*, *37*(9), 2727–2733. https://doi.org/10.1093/molbev/msaa106

Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., Frandsen, P. B., Ware, J., Flouri, T., Beutel, R. G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A. J., Aspöck, U., Aspöck, H., Bartel, D., ... Zhou, X. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, *346*(6210), 763–767. https://doi.org/10.1126/science.1257570

Mo, Y. K., Lanfear, R., Hahn, M. W., & Minh, B. Q. (2023). Updated site concordance factors minimize effects of homoplasy and taxon sampling. *Bioinformatics*, *39*(1), btac741. https://doi.org/10.1093/bioinformatics/btac741

Moore, W., Scarparo, G., & Di Giulio, A. (2022). Foe to frenemy: Predacious ant nest beetles use multiple strategies to fully integrate into ant nests. *Current Opinion in Insect Science*, *52*, 100921. https://doi.org/10.1016/j.cois.2022.100921

Nel, A. (1988). Les Calosomes fossiles de l'Oligocène du sud-est de la France [Col. Carabidae]. *Bulletin de La Société Entomologique de France*, *93*(9), 257–268. https://doi.org/10.3406/bsef.1988.17562

Nel, A. (1989). Les Gyrinidae Fossiles de France (Coleoptera). *Annales de la Société entomologique de France (N.S.)*, *25*(3), 321–330. https://doi.org/10.1080/21686351.1989.12277593

Ng, P. C., & Kirkness, E. F. (2010). Whole Genome Sequencing. In M. R. Barnes & G. Breen (Eds.), *Genetic Variation* (pp. 215–226). Humana Press. https://doi.org/10.1007/978-1-60327-367-1_12

Nilsson, A. N. (2006). Which name is valid—Hygrobiidae or Paelobiidae? *Latissimus*, *21*, 37–39.

Nilsson, A. N., & Hájek, J. (2024). *A world catalogue of the family Dytiscidae, or the diving beetles (Coleoptera, Adephaga). Version 1.I.2024.*

Ober, K. A. (2002). Phylogenetic relationships of the carabid subfamily Harpalinae (Coleoptera) based on molecular sequence data. *Molecular Phylogenetics and Evolution*, *24*(2), 228–248. https://doi.org/10.1016/S1055-7903(02)00251-8

Ober, K. A., & Heider, T. N. (2010). Phylogenetic diversification patterns and divergence times in ground beetles (Coleoptera: Carabidae: Harpalinae). *BMC Evolutionary Biology*, *10*(1), 262. https://doi.org/10.1186/1471-2148-10-262

Ober, K. A., & Maddison, D. R. (2008). Phylogenetic relationships of tribes within Harpalinae (Coleoptera: Carabidae) as inferred from 28S ribosomal DNA and the wingless gene. *Journal of Insect Science*, *8*(1), 63. https://doi.org/10.1673/031.008.6301

Osigus, H.-J., Eitel, M., Bernt, M., Donath, A., & Schierwater, B. (2013). Mitogenomics at the base of Metazoa. *Molecular Phylogenetics and Evolution*, *69*(2), 339–351. https://doi.org/10.1016/j.ympev.2013.07.016

Pauli, M. T., Gauthier, J., Labédan, M., Blanc, M., Bilat, J., & Toussaint, E. F. A. (2024). Museomics of *Carabus* giant ground beetles shows an Oligocene origin and in situ alpine diversification. *Peer Community Journal*, *4*, e70. https://doi.org/10.24072/pcjournal.445

Peters, R. S., Krogmann, L., Mayer, C., Donath, A., Gunkel, S., Meusemann, K., Kozlov, A., Podsiadlowski, L., Petersen, M., Lanfear, R., Diez, P. A., Heraty, J., Kjer, K. M., Klopfstein, S., Meier, R., Polidori, C., Schmitt, T., Liu, S., Zhou, X., … Niehuis, O. (2017). Evolutionary History of the Hymenoptera. *Current Biology*, *27*(7), 1013–1018. https://doi.org/10.1016/j.cub.2017.01.027

Peters, R. S., Meusemann, K., Petersen, M., Mayer, C., Wilbrandt, J., Ziesmann, T., Donath, A., Kjer, K. M., Aspöck, U., Aspöck, H., Aberer, A., Stamatakis, A., Friedrich, F., Hünefeld, F., Niehuis, O., Beutel, R. G., & Misof, B. (2014). The evolutionary history of holometabolous insects inferred from transcriptome-based phylogeny and comprehensive morphological data. *BMC Evolutionary Biology*, *14*(1), 52. https://doi.org/10.1186/1471-2148-14-52

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pflug, J. M., Holmes, V. R., Burrus, C., Johnston, J. S., & Maddison, D. R. (2020). Measuring genome sizes using read-depth, k-mers, and flow Cytometry: methodological comparisons in Beetles (Coleoptera). *G3 Genes Genomes Genetics*, *10*(9), 3047–3060. https://doi.org/10.1534/g3.120.401028

Ponomarenko, A. G. (1969). Cretaceous insects from Labrador 4. A new family of beetles (Coleoptera: Archostemata). *Psyche: A Journal of Entomology*, *76*(3), 306–310. https://doi.org/10.1155/1969/85918

Ponomarenko, A. G. (1987). New Mesozoic water beetles (Insecta, Coleoptera) from Asia. *Paleontological Journal*, *21*(2), 79–92.

Ponomarenko, A. G., & Prokin, A. A. (2015). Review of paleontological data on the evolution of aquatic beetles (Coleoptera). *Paleontological Journal*, *49*(13), 1383–1412. https://doi.org/10.1134/S0031030115130080

Ponomarenko, A. G., & Volkov, A. N. (2013). *Ademosynoides asiaticus* Martynov, the earliest known member of an extant beetle family (Insecta, Coleoptera, Trachypachidae). *Paleontological Journal*, *47*(6), 601–606. https://doi.org/10.1134/S0031030113060063

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes de novo assembler. *Current Protocols in Bioinformatics*, *70*, e102. https://doi.org/10.1002/cpbi.102

Prokin, A. A., Petrov, P. N., Wang, B., & Ponomarenko, A. G. (2013). New fossil taxa and notes on the Mesozoic evolution of Liadytidae and Dytiscidae (Coleoptera). *Zootaxa*, *3666*(2), 137–159. https://doi.org/10.11646/zootaxa.3666.2.2

Prokin, A. A., & Ponomarenko, A. G. (2013). The first record of crawling water beetles (Coleoptera, Haliplidae) in the Lower Cretaceous of Mongolia. *Paleontological Journal*, *47*(1), 89–93. https://doi.org/10.1134/S0031030113010115

Putchkov, A., & Cassola, F. (2005). Tiger beetles deserve separate family status in suborder Adephaga (Coleoptera, Cicindelidae). *Bulletin de La Société Entomologique de France*, *110*(3), 281–293. https://doi.org/10.3406/bsef.2005.16234

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Raupach, M. J., Deister, F., Villastrigo, A., & Balke, M. (2022). The complete mitochondrial genomes of *Notiophilus quadripunctatus* Dejean, 1826 and *Omophron limbatum* (Fabricius, 1777): New insights into the mitogenome phylogeny of the Carabidae (Insecta, Coleoptera). *Insect Systematics & Evolution*, *53*(3), 242–263. https://doi.org/10.1163/1876312X-bja10027

Ribeiro, T. M., & Espíndola, A. (2024). Integrated phylogenomic approaches in insect systematics. *Current Opinion in Insect Science*, *61*, 101150. https://doi.org/10.1016/j.cois.2023.101150

Ribera, I., Beutel, R. G., Balke, M., & Vogler, A. P. (2002). Discovery of Aspidytidae, a new family of aquatic Coleoptera. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1507), 2351–2356. https://doi.org/10.1098/rspb.2002.2157

Ribera, I., Mateu, J., & Bellés, X. (2005). Phylogenetic relationships of *Dalyat mirabilis* Mateu, 2002, with a revised molecular phylogeny of ground beetles (Coleoptera, Carabidae). *Journal of Zoological Systematics and Evolutionary Research*, *43*(4), 284–296. https://doi.org/10.1111/j.1439-0469.2005.00324.x

Říha, P. (1974). Neue fossile Schwimmkäfer aus dem Tertiär Europas und Westsibiriens (Coleoptera: Dytiscidae). *Acta Entomologica Bohemoslovaca*, *71*, 398–413.

Rosová, K., Prokop, J., Hammel, J. U., & Beutel, R. G. (2023). The earliest evidence of Omophroninae (Coleoptera: Carabidae) from mid-Cretaceous Kachin amber and the description of a larva of a new genus. *Arthropod Systematics & Phylogeny*, *81*, 689–704. https://doi.org/10.3897/asp.81.e101374

Roure, B., Baurain, D., & Philippe, H. (2013). Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Molecular Biology and Evolution*, *30*(1), 197–214. https://doi.org/10.1093/molbev/mss208

Schmid, S., Genevest, R., Gobet, E., Suchan, T., Sperisen, C., Tinner, W., & Alvarez, N. (2017). HyRAD-X, a versatile method combining exome capture and RAD sequencing to extract genomic information from ancient DNA. *Methods in Ecology and Evolution*, *8*(10), 1374–1388. https://doi.org/10.1111/2041-210X.12785

Schmidt, J. (2015). On the Eocene age of *Limodromus* Motschulsky, 1850, with description of *L. hoffeinsorum* sp. n. from Baltic Amber (Coleoptera: Carabidae: Platynini). *Zootaxa*, *3974*(4), 573–581. https://doi.org/10.11646/zootaxa.3974.4.8

Schmidt, J., & Michalik, P. (2017). The ground beetle genus *Bembidion* Latreille in Baltic amber: Review of preserved specimens and first 3D reconstruction of endophallic structures using X-ray microscopy (Coleoptera, Carabidae, Bembidiini). *ZooKeys*, *662*, 101–126. https://doi.org/10.3897/zookeys.662.12124

Schmidt, J., Scholz, S., & Kavanaugh, D. H. (2019). Unexpected findings in the Eocene Baltic amber forests: Ground beetle fossils of the tribe Nebriini (Coleoptera: Carabidae). *Zootaxa*, *4701*(4), 350–370. https://doi.org/10.11646/zootaxa.4701.4.2

Schmidt, J., Scholz, S., Wiesner, J., & Will, K. (2023). MicroCT data provide evidence correcting the previous misidentification of an Eocene amber beetle (Coleoptera, Cicindelidae) as an extant species. *Scientific Reports*, *13*(1), 14743. https://doi.org/10.1038/s41598-023-39158-7

Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. *The American Statistician*, *34*(4), 216–221. https://doi.org/10.1080/00031305.1980.10483031

Shi, G., Grimaldi, D. A., Harlow, G. E., Wang, J., Wang, J., Yang, M., Lei, W., Li, Q., & Li, X. (2012). Age constraint on Burmese amber based on U–Pb dating of zircons. *Cretaceous Research*, *37*, 155–163. https://doi.org/10.1016/j.cretres.2012.03.014

Short, A. E. Z. (2018). Systematics of aquatic beetles (Coleoptera): current state and future directions. *Systematic Entomology*, *43*(1), 1–18. https://doi.org/10.1111/syen.12270

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*(6), 945–969. https://doi.org/10.1080/106351501753462894

Simon, S., Letsch, H., Bank, S., Buckley, T. R., Donath, A., Liu, S., Machida, R., Meusemann, K., Misof, B., Podsiadlowski, L., Zhou, X., Wipfler, B., & Bradler, S. (2019). Old world and new world Phasmatodea: Phylogenomics resolve the evolutionary history of stick and leaf insects. *Frontiers in Ecology and Evolution*, *7*, 345. https://doi.org/10.3389/fevo.2019.00345

Sivell, O., Sivell, D., Natural History Museum Genome Acquisition Lab, Darwin Tree of Life Barcoding collective, Wellcome Sanger Institute Tree of Life programme, Wellcome Sanger Institute Scientific Operations: DNA Pipelines collective, Tree of Life Core Informatics collective, & Darwin Tree of Life Consortium. (2023). The genome sequence of a carabid beetle, *Nebria salina* (Fairmaire & Laboulbène, 1854). *Wellcome Open Research*, *8*, 247. https://doi.org/10.12688/wellcomeopenres.19372.1

Smith, S. A., Brown, J. W., & Walker, J. F. (2018). So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *PLoS ONE*, *13*(5), e0197433. https://doi.org/10.1371/journal.pone.0197433

Soghigian, J., Sither, C., Justi, S. A., Morinaga, G., Cassel, B. K., Vitek, C. J., Livdahl, T., Xia, S., Gloria-Soria, A., Powell, J. R., Zavortink, T., Hardy, C. M., Burkett-Cadena, N. D., Reeves, L. E., Wilkerson, R. C., Dunn, R. R., Yeates, D. K., Sallum, M. A., Byrd, B. D., … Wiegmann, B. M. (2023). Phylogenomics reveals the history of host use in mosquitoes. *Nature Communications*, *14*(1), 6252. https://doi.org/10.1038/s41467-023-41764-y

Song, H., Béthoux, O., Shin, S., Donath, A., Letsch, H., Liu, S., McKenna, D. D., Meng, G., Misof, B., Podsiadlowski, L., Zhou, X., Wipfler, B., & Simon, S. (2020). Phylogenomic analysis sheds light on the evolutionary pathways towards acoustic communication in Orthoptera. *Nature Communications*, *11*(1), 4939. https://doi.org/10.1038/s41467-020-18739-4

Sota, T., Takami, Y., Ikeda, H., Liang, H., Karagyan, G., Scholtz, C., & Hori, M. (2022). Global dispersal and diversification in ground beetles of the subfamily Carabinae. *Molecular Phylogenetics and Evolution*, *167*, 107355. https://doi.org/10.1016/j.ympev.2021.107355

Spangler, H. G. (1988). Hearing in tiger beetles (Cicindelidae). *Physiological Entomology*, *13*(4), 447–452. https://doi.org/10.1111/j.1365-3032.1988.tb01129.x

Spangler, P. J., & Steiner, W. E. (2005). A new aquatic beetle family, Meruidae, from Venezuela (Coleoptera: Adephaga). *Systematic Entomology*, *30*(3), 339–357. https://doi.org/10.1111/j.1365-3113.2005.00288.x

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Streicher, J. W., Schulte, J. A., & Wiens, J. J. (2016). How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in Iguanian lizards. *Systematic Biology*, *65*(1), 128–145. https://doi.org/10.1093/sysbio/syv058

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian Phylogenetic and Phylodynamic Data Integration Using BEAST 1.10. *Virus Evolution*, *4*(1). https://doi.org/10.1093/ve/vey016

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577. https://doi.org/10.1080/10635150701472164

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*(5), 778–791. https://doi.org/10.1093/sysbio/syv033

Tan, J. J., Ren, D., & Shih, C. K. (2006). First record of fossil *Priacma* (Coleoptera: Archostemata: Cupedidae) from the Jehol Biota of western Liaoning, China. *Zootaxa*, *1326*(1), 55–68. https://doi.org/10.11646/zootaxa.1326.1.6

Toussaint, E. F. A., Beutel, R. G., Morinière, J., Jia, F., Xu, S., Michat, M. C., Zhou, X., Bilton, D. T., Ribera, I., Hájek, J., & Balke, M. (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. https://doi.org/10.1111/zoj.12332

Toussaint, E. F. A., Gauthier, J., Bilat, J., Gillett, C. P. D. T., Gough, H. M., Lundkvist, H., Blanc, M., Muñoz-Ramírez, C. P., & Alvarez, N. (2021). HyRAD-X exome capture museomics unravels giant ground beetle evolution. *Genome Biology and Evolution*, *13*(7), evab112. https://doi.org/10.1093/gbe/evab112

Toussaint, E. F. A., & Gillett, C. P. D. T. (2018). Rekindling Jeannel's Gondwanan vision? Phylogenetics and evolution of Carabinae with a focus on Calosoma Caterpillar Hunter Beetles. *Biological Journal of the Linnean Society*, *123*(1), 191–207. https://doi.org/10.1093/biolinnean/blx128

Toussaint, E. F. A., Seidel, M., Arriaga-Varela, E., Hájek, J., Král, D., Sekerka, L., Short, A. E. Z., & Fikáček, M. (2017). The peril of dating beetles. *Systematic Entomology*, *42*, 1–10. https://doi.org/10.1111/syen.12198

Tucker, V. A. (1969). Wave-Making by Whirligig Beetles (Gyrinidae). *Science*, *166*(3907), 897–899. https://doi.org/10.1126/science.166.3907.897

Van Belleghem, S. M., Roelofs, D., Van Houdt, J., & Hendrickx, F. (2012). De novo transcriptome assembly and SNP discovery in the wing polymorphic Salt Marsh Beetle Pogonus chalceus (Coleoptera, Carabidae). *PLoS ONE*, *7*(8), e42605. https://doi.org/10.1371/journal.pone.0042605

Van Dam, M. H., Henderson, J. B., Esposito, L., & Trautwein, M. (2021). Genomic characterization and curation of UCEs improves species tree reconstruction. *Systematic Biology*, *70*(2), 307–321. https://doi.org/10.1093/sysbio/syaa063

Van Dam, M. H., Lam, A. W., Sagata, K., Gewa, B., Laufa, R., Balke, M., Faircloth, B. C., & Riedel, A. (2017). Ultraconserved elements (UCEs) resolve the phylogeny of Australasian smurf-weevils. *PLoS ONE*, *12*(11), e0188044. https://doi.org/10.1371/journal.pone.0188044

van Vondel, B. J. (2017). Haliplidae Aubé, 1836. In I. Löbl & D. Löbl (Eds.), *Catalogue of Plaearctic Coleopteram, Vol. 1. Archostemata– Myxophaga– Adephaga* (Revised, pp. 838–843). Brill.

van Vondel, B. J. (2019). Features of the metacoxal air-storage space as additional characters for reconstructing the phylogeny of Haliplidae (Coleoptera). *Tijdschr Voor Entomol*, *162*(1), 13–32. https://doi.org/10.1163/22119434-20192081

van Vondel, B. J. (2021). Revision of the Nearctic Haliplidae (Coleoptera). *Tijdschrift Voor Entomologie*, *163*(2–3), 101–298. https://doi.org/10.1163/22119434-20202093

Vasilikopoulos, A., Balke, M., Beutel, R. G., Donath, A., Podsiadlowski, L., Pflug, M., Waterhouse, R. M., Meusemann, K., Peters, R. S., Escalona, H. E., Mayer, C., Liu, S., Hendrich, L., Alarie, Y., Bilton, D. T., Jia, F., Zhou, X., Maddison, D. R., Niehuis, O., & Misof, B. (2019). Phylogenomics of the superfamily Dytiscoidea (Coleoptera: Adephaga) with an evaluation of phylogenetic conflict and systematic error. *Molecular Phylogenetics and Evolution*, *135*, 270–285. https://doi.org/10.1016/j.ympev.2019.02.022

Vasilikopoulos, A., Balke, M., Kukowka, S., Pflug, M., Martin, S., Meusemann, K., Hendrich, L., Mayer, C., Maddison, D. R., Niehuis, O., Beutel, R. G., & Misof, B. (2021). Phylogenomic analyses clarify the pattern of evolution of Adephaga (Coleoptera) and highlight phylogenetic artefacts due to model misspecification and excessive data trimming. *Systematic Entomology*, *46*(4), 991–1018. https://doi.org/10.1111/syen.12508

Vogler, A. P., & Barraclough, T. G. (1998). Reconstructing shifts in diversification rate during the radiation of Cicindelidae (Coleoptera). In G. E. Ball, A. Casale, A. V. Taglianti, & Museo regionale di scienze naturali (Turin, Italy) (Eds.), *Phylogeny and classification of Caraboidea (Coleoptera: Adephaga): proceedings of a symposium, 28 August 1996, Florence, Italy; XX International Congress of Entomology* (pp. 251–260). Museo regionale di scienze naturali.

Vogler, A. P., & Pearson, D. L. (1996). A Molecular Phylogeny of the Tiger Beetles (Cicindelidae): Congruence of Mitochondrial and Nuclear rDNA Data Sets. *Molecular Phylogenetics and Evolution*, *6*(3), 321–338. https://doi.org/10.1006/mpev.1996.0083

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57–63. https://doi.org/10.1038/nrg2484

Weng, Y., Francoeur, C. B., Currie, C. R., Kavanaugh, D. H., & Schoville, S. D. (2021). A high-quality carabid genome assembly provides insights into beetle genome evolution and cold adaptation. *Molecular Ecology Resources*, *21*(6), 2145–2165. https://doi.org/10.1111/1755-0998.13409

Wickham, H. F. (1909). New fossil Coleoptera from Florissant. *American Journal of Science*, *28*(164), 126–130. https://doi.org/10.2475/ajs.s4-28.164.126

Wiens, J. J. (1998). Does adding characters with missing data increase or decrease phylogenetic accuracy? *Systematic Biology*, *47*(4), 625–640. https://doi.org/10.1080/106351598260635

Wiens, J. J., & Tiu, J. (2012). Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling. *PLoS ONE*, *7*(8), e42925. https://doi.org/10.1371/journal.pone.0042925

Wiesner, J. (2020). *Checklist of the Tiger Beetles of the World* (2nd ed.).

Wiesner, J., Will, K., & Schmidt, J. (2017). Two new genera and species of tiger beetles from Baltic amber (Coleoptera: Carabidae: Cicindelinae). *Insecta Mundi*, *0577*, 1–14.

Wolfe, A. P., McKellar, R. C., Tappert, R., Sodhi, R. N. S., & Muehlenbachs, K. (2016). Bitterfeld amber is not Baltic amber: Three geochemical tests and further constraints on the botanical affinities of succinite. *Review of Palaeobotany and Palynology*, *225*, 21–32. https://doi.org/10.1016/j.revpalbo.2015.11.002

Wutke, S., Blank, S. M., Boevé, J.-L., Faircloth, B. C., Koch, F., Linnen, C. R., Malm, T., Niu, G., Prous, M., Schiff, N. M., Schmidt, S., Taeger, A., Vilhelmsen, L., Wahlberg, N., Wei, M., & Nyman, T. (2024). Phylogenomics and biogeography of sawflies and woodwasps (Hymenoptera, Symphyta). *Molecular Phylogenetics and Evolution*, *199*, 108144. https://doi.org/10.1016/j.ympev.2024.108144

Yager, D. D., Cook, A. P., Pearson, D. L., & Spangler, H. G. (2000). A comparative study of ultrasound-triggered behaviour in tiger beetles (Cicindelidae). *Journal of Zoology*, *251*(3), 355–368. https://doi.org/10.1111/j.1469-7998.2000.tb01086.x

Yan, E. V., Beutel, R. G., & Lawrence, J. F. (2018). Whirling in the late Permian: ancestral Gyrinidae show early radiation of beetles before Permian-Triassic mass extinction. *BMC Evolutionary Biology*, *18*, 33. https://doi.org/10.1186/s12862-018-1139-8

Yang, Q., Chen, Z.-Y., & Jia, F.-L. (2019). *Ambarticus myanmaricus* gen. et sp. nov., the first diving beetle from mid-Cretaceous amber of northern Myanmar (Coleoptera, Dytiscidae, Dytiscinae). *Cretaceous Research*, *102*, 1–6. https://doi.org/10.1016/j.cretres.2019.05.005

Zerm, M., Wiesner, J., Ledezma, J., Brzoska, D., Drechsel, U., Cicchino, A. C., Rodríguez, J. P., Martinsen, L., Adis, J., & Bachmann, L. (2007). Molecular phylogeny of *Megacephalina* Horn, 1910 Tiger Beetles (Coleoptera: Cicindelidae). *Studies on Neotropical Fauna and Environment*, *42*(3), 211–219. https://doi.org/10.1080/01650520701409235

Zhang, C., & Mirarab, S. (2022). Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution*, *39*(12), msac215. https://doi.org/10.1093/molbev/msac215

Zhang, S.-Q., Che, L.-H., Li, Y., Liang, D., Pang, H., Ślipiński, A., & Zhang, P. (2018). Evolutionary history of Coleoptera revealed by extensive sampling of genes and species. *Nature Communications*, *9*(1), 205. https://doi.org/10.1038/s41467-017-02644-4