Insect Systematics and Diversity, 2026, **10**(1), ixag004
DOI: https://doi.org/10.1093/isd/ixag004
Research

ENTOMOLOGICAL SOCIETY OF AMERICA
SHARING INSECT SCIENCE GLOBALLY

OXFORD

Molecular Phylogenetics, Phylogenomics, and Phylogeography

A novel set of ultraconserved element probes for lacewings and allies (Insecta: Neuropterida)

Alice C. Assmar*[,1], Ziad Khouri[2], Renato J.P. Machado[3], and Jessica P. Gillung[1]

[1]Natural Resource Sciences Department, Lyman Entomological Museum, McGill University, Ste-Anne-de-Bellevue, Quebec, Canada
[2]Department of Entomology and Nematology, University of California, Davis, CA, USA
[3]Department of Zoology, Biological Sciences Sector, Universidade Federal do Paraná (UFPR), Curitiba, Paraná, Brazil

*Corresponding author. Natural Resource Sciences Department, Lyman Entomological Museum, McGill University, 21,111 Lakeshore Road, Macdonald Campus, Ste-Anne-de-Bellevue, Quebec H9X 3V9, Canada (Email: alice.assmar@mail.mcgill.ca).

Subject Editor: Rodrigo Monjaraz-Ruedas

Target-enrichment of ultraconserved elements has proven advantageous in recovering large quantities of phylogenetically informative data even from specimens with suboptimal DNA preservation. It is also possible to develop ultraconserved element probe sets with minimal exemplar taxa, enabling the resolution of recalcitrant nodes in the Tree of Life even of rare groups. Unfortunately, probe sets are still lacking for multiple groups of insects, such as the 3 orders in Neuropterida (Insecta: Holometabola): Megaloptera, Neuroptera, and Raphidioptera. Additionally, a well-curated probe set for neuropterids makes it possible for the investigation of their evolution, including reconstructing the ancestral state of characters and examining their life-history transitions. In this study, we developed the first openly available probe set targeting ultraconserved elements of Neuropterida. We in-silico tested different parameters to design an optimized probe set that efficiently captures conserved loci of the 3 orders of Neuropterida. After comparing taxonomically tailored and universal probes, our study suggests that the latter recovers higher number of shared ultraconserved elements for the superorder. Accordingly, we provide an optimized probe set that successfully targets 5,577 loci across all Neuropterida lineages.

Keywords: Megaloptera, Neuroptera, phylogenomics, probe design, Raphidioptera

Introduction

Modern advances in evolutionary biology methods enable us to obtain new insights into how biodiversity responded to major events that happened on Earth in the past. Phylogenies have long been used for understanding evolutionary history and relationships among taxa, and when combined with the advances in DNA sequencing, they proved to be important to several fields of biology and the key to the understanding of deeper evolutionary and biogeographic histories (Yang and Rannala 2012). A phylogenetic tree not only provides a hypothesis for species relationships but also enables the estimation of the origin of traits, divergence times, and diversification rates, in addition to mapping heritable character states (Yang and Rannala 2012, Kapli et al. 2020). However, the expensive cost of traditional Sanger sequencing and the relatively low quantity of phylogenetically informative data recovered with only a few genes limited the application of molecular phylogenetics (Blaimer et al. 2015, 2016, Young and Gillung 2020).

High-throughput sequencing (HTS), also known as next-generation sequencing, dramatically improved DNA sequencing after the mid-2000s and provided a method for rapid and low-cost sequencing of genomic scale data (Kapli et al. 2020, Young and Gillung 2020). The development of HTS enabled researchers to acquire data from almost the complete genome (whole-genome sequencing—WGS), which may reduce sampling (stochastic) errors (Kapli et al. 2020, Young and Gillung 2020). However, the computational power and bioinformatic skills required to analyze this large amount of data have become a challenge (Kapli et al. 2020, Young and Gillung 2020). Additionally, several loci without phylogenetically informative data can be recovered with WGS, increasing the time required for data cleaning and the potential for systematic error (violation of model assumptions) to occur (Blaimer et al. 2016, Young and Gillung 2020). Methods of target enrichment minimize these challenges, allowing for the sequencing of regions of interest throughout the genome (Faircloth et al. 2012, Blaimer et al. 2016).

The target enrichment of ultraconserved elements (UCEs) has been proven successful as a reduced representation sequencing technique (Young and Gillung 2020). UCEs are regions with highly conserved sequences (95% to 100% similarity) across distantly related taxa (Bejerano et al. 2004, Faircloth et al. 2012), which have been used to investigate phylogenetic relationships of both shallow and deep taxonomic levels

Received: 4 August 2025; **Revised:** 28 December 2025; **Accepted:** 11 January 2026

(Faircloth et al. 2012, Blaimer et al. 2015, Baca et al. 2017, Derkarabetian et al. 2019, Gustafson et al. 2020, Buenaventura et al. 2021, Saenz Manchola et al. 2022, Sota et al. 2022, Baca et al. 2025). This method also performs well in recovering data even from specimens with suboptimal DNA preservation (eg old, pinned insects), allowing the use of specimens from groups that are difficult to collect or even extinct or extirpated (Faircloth et al. 2015, Blaimer et al. 2016, Van Dam et al. 2021). Additionally, through museomics, it allows researchers to increase the taxon sampling for regions of the world that are difficult to access by using specimens already deposited in museums (Derkarabetian et al. 2019, Van Dam et al. 2021). Although UCE sequencing is a broadly applied method, many challenges remain, especially regarding how to best design and curate probe sets (short DNA sequences used to target the region of interest) (Blaimer et al. 2015, Faircloth 2017, Gustafson et al. 2019, 2020, Van Dam et al. 2022, Gustafson et al. 2023).

Shared universal UCE probe sets are currently available for several vertebrate groups, and they also perform well for internal lineages (Gustafson et al. 2019). However, for highly diverse and ancient organisms such as insects, universal probe sets may be insufficient to accurately recover data from lineages that are phylogenetically distant from the exemplar taxa (ie the taxa used for designing the probes) due to reduced hybridization efficiency (Gustafson et al. 2019, 2020, 2023). For insects, tailored probe sets work better in recovering larger numbers of loci from closely related taxa, but designing new probes for each model group results in targeting different UCEs, thus reducing inter-dataset compatibility (Gustafson et al. 2019, 2023). Both universal and tailored probe sets have been designed to capture UCEs of different insect groups, including Coleoptera (Faircloth 2017, Gustafson et al. 2019, 2023), Diptera (Faircloth 2017, Buenaventura et al. 2021), Hemiptera (Faircloth 2017), Hymenoptera (Blaimer et al. 2015, Faircloth et al. 2015, Branstetter et al. 2017, Faircloth 2017), Isoptera (Hellemans et al. 2022), Lepidoptera (Faircloth 2017), and Phthiraptera (Zhang et al. 2019).

When designing new probe sets, it is possible to combine them with previously designed probes, thus diversifying the data and increasing the compatibility with previous datasets (Gustafson et al. 2019). Additionally, it is possible to design universal probes based on representative exemplar taxa and later subset them to specific groups within the main taxa, maintaining the universality of the probes (Gustafson et al. 2019, Young and Gillung 2020). In this study, we tested this last approach to design the first optimized probe set for Neuropterida (Insecta: Holometabola), which will allow future studies to further explore their evolutionary relationships.

Neuropterida is composed of 3 relatively small orders: Megaloptera, Raphidioptera, and Neuroptera (Engel et al. 2018). They display a wide variety of ecological strategies, including freshwater vs terrestrial organisms, widespread vs restrict distributions, generalists vs specialists, predation, parasitism, parasitoidy, extracorporeal digestion vs intracorporeal digestion, etc. (Oswald and Machado 2018). The phylogenetic relationships of Neuropterida have been widely discussed for the last few decades (Aspöck 2002, Haring and Aspöck 2004, Wang et al. 2017, Engel et al. 2018, Winterton et al. 2018, Vasilikopoulos et al. 2020, Lai et al. 2025, Zhang et al. 2025) due to their great morphological diversity. Additionally, their evolutionary history is marked by several events of extinction (Engel et al. 2018), which sometimes obscure their phylogenetic relationships.

Using a representative taxon sampling to design UCE probe sets allows us to identify numerous UCEs across the lineages of Neuropterida. This approach, in turn, enables the reconstruction of both deep and shallow evolutionary relationships within this group by providing access to a whole new set of molecular data. Our goals are to: (i) design probe sets for targeting UCEs of species of Neuropterida and examine the utility of these newly designed universal probes in recovering UCE data, (ii) test different optimization parameters for probe set design, such as stringency and choice of base genomes, (iii) test the efficiency of subsetting probes in recovering UCEs of lower taxonomic levels, and (iv) compare the universal probe set with probes designed separately for each order. Finally, we aim to validate probe design by using the newly generated probe set to capture UCEs from available genomic sequences and reconstruct a phylogenetic tree for Neuropterida.

Materials and Methods

Taxon Sampling for Bait Design

Neuropterida are a diverse superorder comprising approximately 6,500 species and 19 families (Winterton et al. 2018). Their approximate range of genome sizes is 300 Mbp to 1.3 Gbp (Broad et al. 2023, Zhao et al. 2025), but there is a 99 Mbp genome sequence of *Conwentzia psociformis* (Curtis) (Neuroptera) in the National Center for Biotechnology Information (NCBI) database. In total, we included 17 genomes across 9 families. We sequenced 3 of these genomes de novo using low-coverage WGS and downloaded all others from NCBI. In Table 1, we provide a list of the species used for bait design. Additionally, we also included 1 genome of Coleoptera, *Tribolium castaneum* (Herbst) (GCA_000002335.2), as the outgroup, both to root the phylogenies and test the efficiency of the new probe set. We verified the genomes' completeness using BUSCO (Simão et al. 2015) and tested whether the most complete genome yielded the greatest number of UCEs (respective scores showed Fig. 1).

DNA Extraction and Sequencing

We performed the DNA extractions for the 3 newly sequenced genomes in the molecular lab of the Canadian National Collection of Insects, Arachnids, and Nematodes (CNC), Agriculture and Agri-Food Canada. We used the Qiagen QIAamp DNA Micro Kit for DNA extractions and followed the manufacturer's protocol (QIAGEN 2014), except for the amount of proteinase K, which we tripled to improve the concentration of DNA extracted. We submerged each sample in 180 µl Buffer ATL, mixed with 60 µl of proteinase K and placed them in a heated incubator for 12 to 24 h. For the binding step, we incorporated 240 µl of Buffer AL and the optional step of adding 1 µl carrier RNA to increase the possibility of capturing fragmented DNA. Subsequently, we added 240 µl ethanol (96% to 100%) and transferred the lysate to a QIAmp MinElute column. We washed the membrane with 500 µl of Buffer AW1, followed by Buffer AW2 and finally eluted with 200 µl pH-regulated water. We then quantified DNA concentrations using the High Sensitivity Assay Kit on a Qubit 4 Fluorometer.

The extracted DNA was later stored in 50 µl Buffer AE and shipped to the McGill Genome Centre facility for sequencing.

Table 1. List of genomes used as representative taxa to design probe sets of Neuropterida

Order	Family	Species	Source	Estimated read depth	GenBank Genome Assembly
Megaloptera	Corydalidae	*Acanthacorydalis orientalis* (McLachlan 1899)	NCBI	258×	GCA_034766995.1
		Neoneuromus ignobilis (Navás 1932)	NCBI	112×	GCA_029203775.1
	Sialidae	*Sialis lutaria* (Linnaeus 1758)	NCBI	60×	GCA_949319165.1
		Sialis fuliginosa (Pictet 1836)	NCBI	65×	GCA_961205875.1
Neuroptera	Chrysopidae	*Chrysopa pallens* (Rambur 1838)	NCBI	204×	GCA_020423425.1
		Chrysoperla carnea (Stephens 1836)[a]	NCBI	40×	GCF_905475395.1
		Chrysoperla zastrowi sillemi (Esben-Petersen 1935)	NCBI	30×	GCA_040669965.1
		Nineta flava (Scopoli 1763)	NCBI	39×	GCA_963920215.1
	Coniopterygidae	*Conwentzia* sp.	Newly sequenced	5×	SAMN51422129
	Myrmeleontidae	*Brachynemurus nebulosus* (Olivier 1811)	NCBI	34×	GCA_035578155.1
	Osmylidae	*Spilosmylus inthanonensis* (New 1991)	Newly sequenced	5×	SAMN51422130
	Sisyridae	*Sisyra nigra* (Retzius 1783)	NCBI	55×	GCA_958496155.1
		Sisyra terminalis (Curtis 1854)	NCBI	82×	GCA_958496175.1
Raphidioptera	Inocelliidae	*Negha inflata* (Hagen 1861)	Newly sequenced	5×	SAMN53489216
	Raphidiidae	*Mongoloraphidia duomilia* (Yang 1998)	NCBI	64×	GCA_039654895.1
		Venustoraphidia nigricollis (Albarda 1891)	NCBI	20×	GCA_034508555.1
		Xanthostigma xanthostigma (Schummel 1832)	NCBI	29×	GCA_963575645.1

[a]Annotated genome.

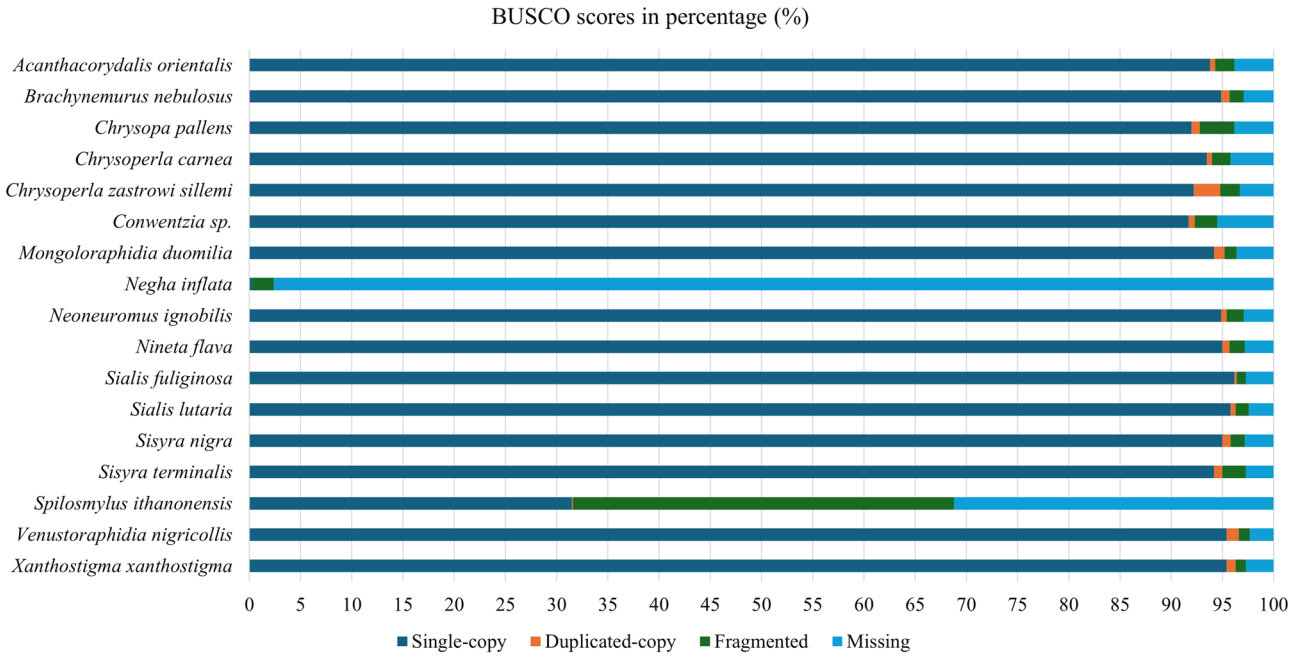

Fig. 1. BUSCO scores of genomes.

Library preparation was performed using a Kapa Hyper Prep Library Preparation kit, with PCR, followed by successive sessions of Ampure clean-up because of the reduced quantity of DNA. Low-coverage WGS was conducted in a NovaSeq6000 S4 v1.5, 150 bp paired-end reads, at 5× coverage.

Data Cleaning and Genome Assembly

We trimmed adapters and other contaminants from raw reads using Trimmomatic v0.39 (Bolger et al. 2014) and assessed the quality of the trimmed reads using FastQC v0.12.0 (Andrews 2023). All paired reads, Anchored Hybrid Enrichment (AHE), and transcriptome sequences were assembled de novo (ie without a reference genome) using SPAdes v4.0.0 (Prjibelski et al. 2020). We tested several SPAdes assembly options before the final run and assessed the quality of the assemblies using QUAST v5.0.2 (Gurevich et al. 2013) to ensure the best method was selected. We assembled the reads using the single cell (-sc) running mode, specifying paired-ended libraries, and for the AHE and transcriptome sequences, we used "isolate." We set the

advanced option "cov-cutoff" to auto (Nurk et al. 2013, Prjibelski et al. 2020). The aforementioned bioinformatic steps were performed on the Digital Research Alliance of Canada computing clusters facilitated by the partnership between McGill University and Calcul Quebec. All bioinformatic steps, from trimming to quality check, are available in the online supplementary data (sm1_scripts).

Probe Set Design

We followed Tutorial IV of PHYLUCE, which is explained in detail by Faircloth (2017), to design multiple probe sets (Faircloth 2016a). For simplicity, we refer to the probes designed for the superorder as "universal" and to the ones designed specifically for each order as "tailored." We modified specific commands to optimize the final probe set, and our full experimental design is detailed in Fig. 2. The PHYLUCE workflow starts by cleaning the genomes using "BioPython" (Cock et al. 2009) and converting them to "2bit" format using the software "faToTwoBit." We used "ART" (Huang et al. 2012) to simulate reads without sequencing error, with a length of 100 bp, covering 2× the genome, insert size of 200 bp, and 150 bp of SD. We prepared the base genomes and aligned the simulated reads on a taxon-by-taxon basis using "Stampy" (Lunter and Goodson 2011), applying the following parameters: --maxbasequal 93, --substitutionrate = 0.05 and insertsize = 400. We removed unmapped reads with "SAMtools" (Danecek et al. 2021) and used "BEDtools" (Quinlan and Hall 2010) to convert "BAM" files into "BEDs," sort the sequence lines, and merge overlapping intervals.

After generating the database containing all alignment intervals that are shared among taxa ("phyluce_probe_get_multi_merge_table"), we included a filter criterion to enforce that only regions present in at least 1 species of all 3 orders were maintained. This filter guarantees that our baits are not over-representing 1 order but are successfully recovering regions shared among the 3 orders (commands available in sm1_scripts > phyluce_scripts > base-genome-test).

We then queried this filtered database with "phyluce_probe_query_multi_merge_table" and identified highly conserved loci (called putative loci "sensu" Gustafson et al. 2019). The stringency is established considering loci found in the base genome plus at least 1 other taxon. In all base genome tests, we decreased stringency and selected the base genome plus 10 additional taxa. By decreasing the number of taxa in this step, the number of recovered loci increases in the in-silico test (Gustafson et al. 2019). We extracted 160 bp-long FASTA sequences using "phyluce_probe_get_genome_sequences_from_bed" and designed the temporary baits using "phyluce_probe_get_tiled_probes," overlapping the middle of the loci, with 3× tiling density and 2 baits per locus. We removed baits with extreme GC content or with more than 25% repeated content. We also removed duplicates using "phyluce_probe_easy_lastz" and "phyluce_probe_remove_duplicate_hits_from_probes_using_lastz" (online supplementary data > sm1_scripts > phyluce_scripts).

We aligned the temporary baits against all exemplar genomes to increase bait representation and more consistently capture the conserved loci across the multiple taxa (Faircloth 2016a, 2017). From this step on, we included the outgroup taxon. After identifying which candidate loci were detected consistently across the multiple taxa, a new stringency filter was applied, now considering all taxa. Selecting only the baits that hit all taxa results in the highest stringency, but it also captures the lowest number of UCEs (Faircloth 2016a, 2017).

Base Genome Test

To optimize the number of putative loci identified, we tested the effect of base genome choice by repeating the previous steps for each sequence whilst considering them as base genomes (Fig. 2). All other parameters remained unchanged to ensure proper comparison among the different base genomes. During candidate loci identification, we set to 16 the number of taxa within which the candidate loci must be found. We compared the length and final alignment matrix, at 70% completeness, generated by the data collected using the baits of each of these base genomes.

Stringency Test and Final Probe Set Design

To establish the ideal stringency for bait design, we also used Tutorial IV of the PHYLUCE pipeline (Faircloth 2016b). More specifically, using the best resulting base genome from previous steps, we repeated the steps in the "Conserved locus validation" section, available at https://phyluce.readthedocs.io/, starting at "Find which loci we detect consistently," and varying only the number of taxa within which the candidate loci must be found. Two stringency groups were tested: the broad, or least stringent, when targeting loci shared among 10 out of the 18 taxa, and strict, or most stringent, when shared among 16 taxa. We evaluated the length of the loci and completeness of the 70% alignment matrix in the in-silico test as measurements of the stringency efficiency.

Final Probe Set Design

After identifying the optimal base genome and stringency, we created the universal probe list (ie the final probe set). As previously mentioned, we used all exemplar genomes, which ensures that the final probe set contains a more representative mix of probes targeting the same locus. We applied the command "phyluce_probe_get_tiled_probe_from_multiple_inputs," which accepts more than 1 sequence as input, employing the same parameters used for the temporary baits. We removed all baits that had identity and coverage equal to or greater than 50. Using CURE (curation of UCEs) pipeline (Freitas et al. 2023), we identified the location of the UCEs captured by the final universal probe set based on the GeneRegion strategy, using the annotated genome of *Chrysoperla carnea* (Stephens) as a reference.

Probe Set Subsetting

To subset the universal probe list generated in the steps above, we selected the desired taxa and used the command "phyluce_probe_get_subsets_of_tiled_probes." We created 1 subset for the superorder by selecting 1 species of each family of the 3 orders, as follows: Megaloptera: *Acanthacorydalis orientalis* (McLachlan) (Corydalidae), *Sialis fuliginosa* (Pictet) (Sialidae); Neuroptera: *Brachynemurus nebulosus* (Olivier) (Myrmeleontidae), *Conwentzia* sp. (Coniopterygidae), *Nineta flava* (Scopoli) (Chrysopidae), *Sisyra nigra* (Retzius) (Sisyridae), *Spilosmylus inthanonensis* (New) (Osmylidae), Raphidioptera: *Negha inflata* (Hagen) (Inoceliidae), and *Mongoloraphidia duomilia* (Yang) (Raphidiidae). For the subset of each order, we included only species of the individual orders as the desired taxa.

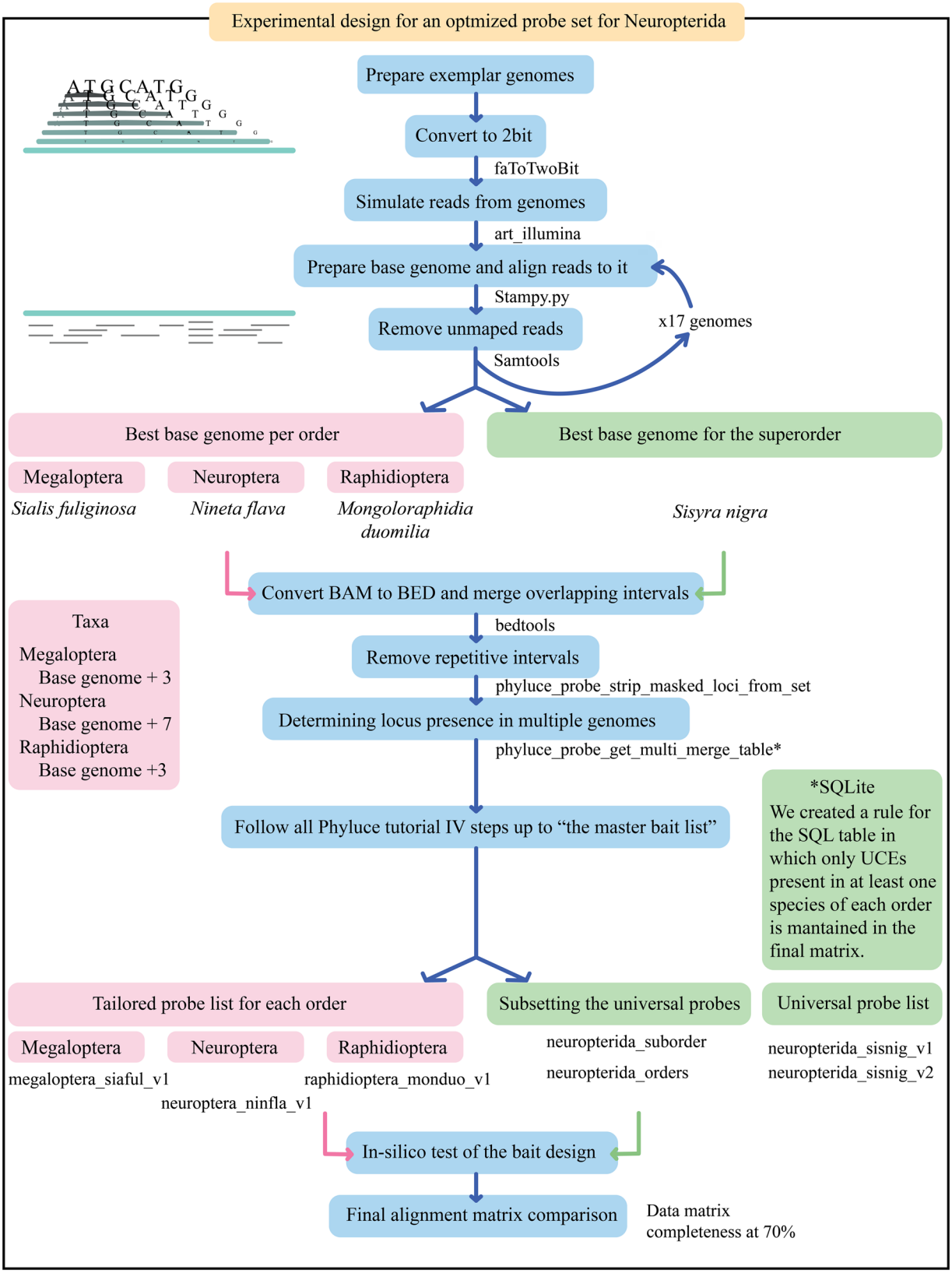

Fig. 2. Experimental design.

For the superorder subset of the combined tailored and universal probes, we included an additional filtering step to reduce the number of probes in the probe set and minimize the costs of synthesizing these probes. Using the SQLite multifasta table, we filtered the UCEs shared across all sequences, except the outgroup, and favored these loci in the final reduced subset.

In-Silico Probe Design Test

Using the same 17 genomes of our taxon sampling, we tested the efficiency of each probe set designed to collect data from genomes that could generate a reliable phylogenetic tree. It starts by aligning the probe set to the sequences using "phyluce_probe_run_multiple_lastzs_sqlite." We extracted the fasta files for each locus, using a flanking length of 600 bp ("phyluce_probe_slice_sequence_from_genomes") and matched the contigs to the probes ("phyluce_assembly_match_contigs_to_probes"). After adding all recovered loci into a monolithic fasta file, we aligned UCEs using MAFFT (Katoh and Standley 2013). We trimmed the alignments using "Gblocks" (Castresana 2000, Talavera and Castresana 2007), cleaned the locus names from alignments, and generated a 70% complete matrix. We computed the statistical measures of the alignments obtained with "phyluce_align_get_align_summary_data," which enables determination of the efficiency across probe sets. Finally, for the phylogenetic reconstruction, we concatenated the alignments, partitioned them by locus and preformed a maximum-likelihood phylogenetic analysis using "IQ-TREE" (Minh et al. 2020). We used the GTR substitution model with a gamma distribution for site rate heterogeneity (with 4 categories), 2,000 ultrafast bootstrap (Hoang et al. 2018) iterations, and a random seed.

Comparison of Universal Probes with Subsets

To create tailored probe sets and compare their efficiency against the universal probe set for the Neuropterida orders, we selected a base genome of each order based on the number of putative loci from the base genome test. However, as exemplar taxa, we selected only species of the same order to comprise the sampling. We sampled 4 species of Megaloptera, 8 of Neuroptera, and 4 of Raphidioptera (Table 1). For all 3 probe sets designed, we included *T. castaneum* as the outgroup. For the development of the tailored probes, it was not necessary to modify the database containing all alignment intervals that are shared among taxa ("phyluce_probe_get_multi_merge_table").

The assigned names of the UCEs detected across probe sets designed using different base genomes are not compatible even if they target the same region. To overcome this issue, we determined which conserved loci were shared among the different probe sets for final comparison. We used the tailored probes to capture UCEs only from species in the order for which they were designed. We followed the bash and R-script "comparing_monolithic_UCE_fastas" pipeline (Alexander 2018, Gustafson et al. 2019) to identify the shared conserved loci. These scripts also further identify problematic loci across the dataset. After defining the matches, we applied a Python script (created by Z. Khouri, supplementary_material > sm2_python_codes) to create a dictionary of matches and combine the exploded fasta files (separated files organized by locus) targeted from separate probe sets. It resulted in a combined dataset with homologous loci captured by different tailored probe sets.

Finally, we conducted 3 separate experiments in-silico: (i) the dataset created using the universal probe set vs the combined datasets obtained from the tailored probes using only genomes; (ii) the universal dataset vs the combined datasets using different genetic datasets (AHE, transcriptomes, and genomes); and (iii) the subset for each order vs the tailored probe set designed for each order. For the first experiment, we compared the alignment summary for the 17 exemplar taxa and 1 outgroup. However, for the second and third experiments, to further investigate the efficiency of the designed probes in targeting conserved loci, we acquired additional sequences available in the NCBI. The dataset for Megaloptera comprised 9 taxa, Neuroptera contained 81 taxa, and Raphidioptera included 12 taxa (see supplementary data > sm3_taxaset). In total, we included 102 Neuropterida sequences and 1 outgroup (*T. castaneum*).

We aligned the loci with MAFFT and followed the PHYLUCE pipeline to trim and concatenate fasta files. We then tested the efficiency of subsetting the universal and tailored probe sets in downstream analyses by comparing locus length and the final 70% complete alignment matrix.

Neuropterida Phylogeny Based on UCEs

For final phylogenetic reconstruction, we used the universal probe set to target UCEs from genomes, transcriptomes, and AHE sequences of species of Neuropterida available on NCBI. In addition, we included the following 9 sequences of Neuroptera to better represent all extant families in the order: *Zachobiella pallida* (Banks), *Aeropteryx monstrosa* (Riek), *Drepanacra binocula* (Newman), *Eumantispa harmandi* (Navás), *Micromus variegatus* (Fabricius), *Plega dactylota* (Rehn), *Rhachiberotha sheilae* (Aspöck and Mansell), *Ululodes arizonensis* (Banks), and *Ululodes bicolor* (Banks). In total, our analysis included sequence data from 112 taxa: 9 Megaloptera (4 genomes, 4 AHEs, and 1 transcriptome), 90 Neuroptera (9 genomes, 61 AHEs, and 20 transcriptomes), 12 Raphidioptera (4 genomes, 4 AHE, and 4 transcriptomes), and 1 genome of the outgroup species. The complete list of taxa and their accession numbers is provided in the online supplementary data > sm3_taxaset. We utilized the same steps we used for the in-silico test to estimate the Neuropterida phylogeny, with the exception of using AMAS (Borowiec 2016) to concatenate the alignments and obtain a partition file by UCE locus. Finally, we reconstructed a maximum-likelihood phylogenetic tree using "IQ-TREE," with the command "-m TEST" activated, which identifies the best-fitting substitution model for each partition with ModelFinder (Kalyaanamoorthy et al. 2017). We set up 2 runs, 2,000 ultrafast bootstrap repetitions, and a random seed (see online supplementary data > sm1_scripts > phylogenetic_reconstruction).

Results

Base Genome

After carrying out the PHYLUCE workflow for each of the 17 different base genomes, *N. flava* performed best as the base genome for the superorder during putative loci recovery (12,410 UCEs and 20,549 probe count) (Table 2). *Sisyra terminalis* and *S. nigra* were the second and third best, capturing 12,302 and 12,239 putative UCEs, respectively. Interestingly, these 3 genomes were not among the ones with the highest BUSCO scores. The genomes of *Venustoraphidia nigricollis* (Albarda), *Sialis fuliginosa* (Pictet), *Sialis lutaria* (Linnaeus), and *Xanthostigma xanthostigma* (Schummel) had 96.6%, 96.4%, 96.3%, and 96.3% BUSCO scores, respectively, which were the highest scores among sampled taxa (Fig. 1). Despite that, they were not among the top 3 in capturing conserved loci. In fact, *S. lutaria* was among the least efficient base genomes. The baits designed using the genomes of *M. duomilia*

Table 2. Results of base genome tests based on the *phyluce_probe_get_tiled_probes* and *phyluce_probe_get_tiled_probe_from_multiple_inputs* conserved locus count

Base-genomes	Putative loci count (str = base + 10)	Candidate loci count (str=16)	70% C.A.M.
Nineta flava (Scopoli 1763)	12,410	5,098	2,444
Sisyra terminalis (Curtis 1854)	12,302	5,669	2,452
Sisyra nigra (Retzius 1783)	12,239	5,739	2,480
Mongoloraphidia duomilia (Yang 1998)	12,081	5,904	2,469
Venustoraphidia nigricollis (Albarda 1891)	11,928	5,741	2,447
Chrysoperla carnea (Stephens 1836)	11,520	5,684	2,320
Xanthostigma xanthostigma (Schummel 1832)	11,435	5,462	2,363
Chrysoperla zastrowi sillemi (Esben-Petersen 1935)	10,592	4,066	2,178
Sialis fuliginosa (Pictet 1836)	10,208	4,671	2,061
Chrysopa pallens (Rambur 1838)	10,038	4,050	2,089
Acanthacorydalis orientalis (McLachlan 1899)	9,247	4,534	2,160
Spilosmylus inthanonensis (New 1991)	5,700	2,052	1,250
Brachynemurus nebulosus (Olivier 1811)	5,175	2,135	1,202
Conwentzia sp.	4,593	1,514	923
Sialis lutaria (Linnaeus 1758)	4,371	1,660	1,019
Neoneuromus ignobilis (Navás 1932)	3,855	1,571	1,052
Negha inflata (Hagen 1861)	3,217	859	671

The table is sorted by descending the number of putative locus count. str, stringency, C.A.M, complete alignment matrix.

and *V. nigricollis* as base genomes also resulted in a high number of putative loci (12,081 and 11,928, respectively).

For the candidate loci search, the probe set developed using *M. duomilia* as base genome performed better and captured the highest number of conserved loci (5,904), followed by *S. terminalis* (5,767). *Sisyra nigra* and *N. flava* were the third and fourth best base genomes, with 5,739 and 5,098 UCE candidate loci identified, respectively. However, during the final in-silico test of the bait design, the 70% alignment matrix generated by the data collected using *S. nigra* as base genome resulted in a higher number of alignments (2,480). The matrix assembled with data collected using probe sets for *M. duomilia*, *S. terminalis*, and *N. flava* resulted in 2,469, 2,452, and 2,444 alignments, respectively. Alignment length and minimum and maximum locus lengths were similar across the data collected using the different probe sets. Overall, our results indicate that the genome of *S. nigra* performed better as base genome in downstream analyses and was, hence, selected to design the universal probe set for Neuropterida.

Broad vs Stringent Probe

When testing for probe set stringency using *S. nigra* as the base genome, we compared it to loci shared by 10 (neuropterida-sisnig-v2) and 16 taxa (neuropterida-sisnig-v1). The total number of loci recovered with the stringency of 10 was 7,698 with a probe count of 232,151. For the stringency of 16, we identified 5,289 conserved loci and 172,861 probes without duplicates (Table 3).

Table 3. List of designed probe sets, with the number of identified UCEs and total probe count for each probe set

Probe sets designed	Conserved locus count	Probe count
neuropterida_sisnig_v1 str 16	5,289	172,861
neuropterida_sisnig_v2 str 10	7,698	232,151
subset_superorder	5,289	84,015
subset_megaloptera	5,280	39,644
subset_neuroptera	5,289	88,286
subset_raphidioptera	5,288	35,580
specific_megaloptera	4,380	42,333
specific_neuroptera	5,357	96,394
specific_raphidioptera	4,687	46,514
combined_universal_tailored	5,577	180,449
combined_subset_superorder	2,038	33,311
combined_subset_megaloptera	5,486	41,222
combined_subset_neuroptera	5,503	91.809
combined_subset_raphidioptera	5,470	37,026

str, stringency.

As expected, the in-silico tests of the 2 probe sets (less stringent vs strict) required a high computational time to identify, slice, and match the conserved loci to the probes, with the less stringent probe set being more computationally demanding compared to the strict one (385 vs 245 min) (64 Gb RAM, 8 dedicated cores, and M1 Max processor). Additionally, it also required more time to align the sequences for the final analysis (110 vs 75 min). Despite targeting 3,000 more loci, the broad probe set did not recover a significantly higher number of alignments in the 70% complete matrix. In total, the final 70% matrix had 2,960 alignments when using the broad probe set and 2,480 using the stringent set, with a difference of less than 500 loci (Table 4).

Regarding the composition of the complete alignment matrix, the dataset generated using the broad probe set included 34,636,696 characters, with 5.1% of missing data. The alignment length was 3,083,567 bp, with a mean of 591 bp. For the strict probe set, alignment length was 2,261,661 bp, with a mean of 590 bp (Table 4). The total character count included 28,190,168, with 5% of missing data.

The topologies recovered using the data collected from the 17 taxa show similar results for both datasets (datasets using the 10-taxon, neuropterida-sisnig-v2 and 16-taxon, neuropterida-sisnig-v1). They consistently recovered similar relationships among the taxa, with high branch support (Fig. 3). Overall, considering the computational time required for each analysis and the results of the in-silico test, the probe set with higher stringency proved to be more advantageous for targeting UCEs of Neuropterida. Figure 4 shows the number of UCEs captured by this probe set based on their location in the genes.

Probe Set Designed for Each Order

The base genomes for each order were *N. flava* (Neuroptera), *S. fuliginosa* (Megaloptera), and *M. duomilia* (Raphidioptera). The probe set designed for Neuroptera yielded 5,357 conserved loci shared by 8 taxa (neuroptera_ninfla_v1.fasta), while the Megaloptera and Raphidioptera probe sets identified 4,380 shared by 4 taxa (megaloptera_siaful_v1.fasta) and 4,687 shared by 5 taxa (raphidioptera_monduo_v1.fasta), respectively (Table 3).

Table 4. PHYLUCE summary of the alignment matrix of 17 taxa obtained using UCEs captured by 6 different probe sets

Probe sets	Alignment length	Mean	Min. length (bp)	Max. length (bp)	Total characters	Missing data characters	70% C.A.M.
neuropterida_sisnig_v1 str 16	2,261,661	590.67	89	1,470	28,190,168	1,425,720	2,480
neuropterida_sisnig_v2 str 10	3,083,567	591.74	78	1,743	34,636,696	1,773,033	2,960
subset_superorder	2,256,305	586.97	113	1,417	27,444,824	1,360,476	2,499
combined_orders	492,793	612.93	138	1,350	5,542,665	390,663	436
combined_universal_tailored	2,323,958	593.60	89	1,581	28,806,649	1,462,412	2,518
combined_subset_superorder	946,629	608.76	113	1,581	12,930,292	688,272	1,164

bp, base pairs; C.A.M, complete alignment matrix; str, stringency.

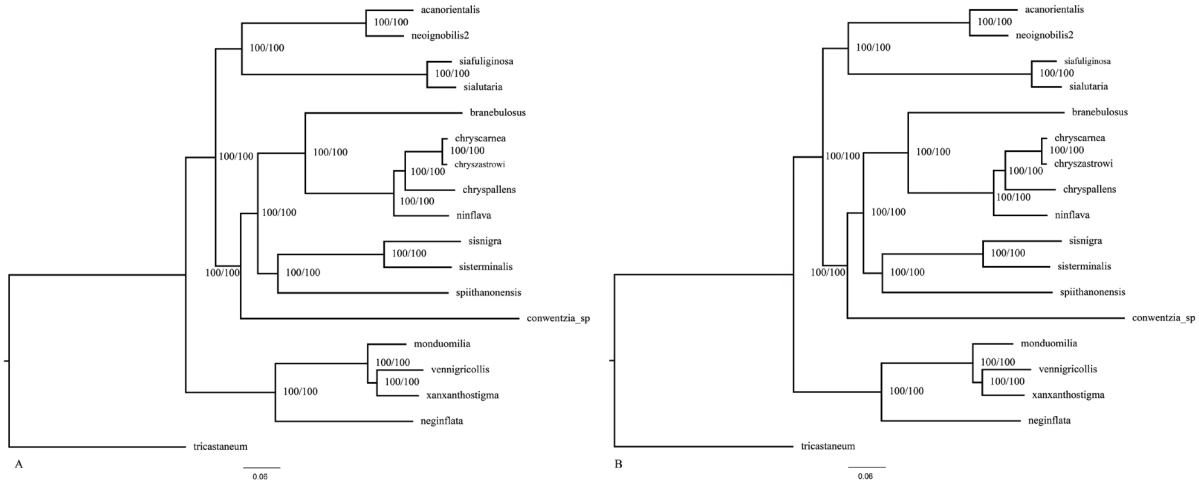

Fig. 3. Maximum likelihood phylogeny reconstructed using data captured using the different stringent probe sets. Tree reconstructed using probe set with shared loci among A) 10 taxa and B) 16 taxa.

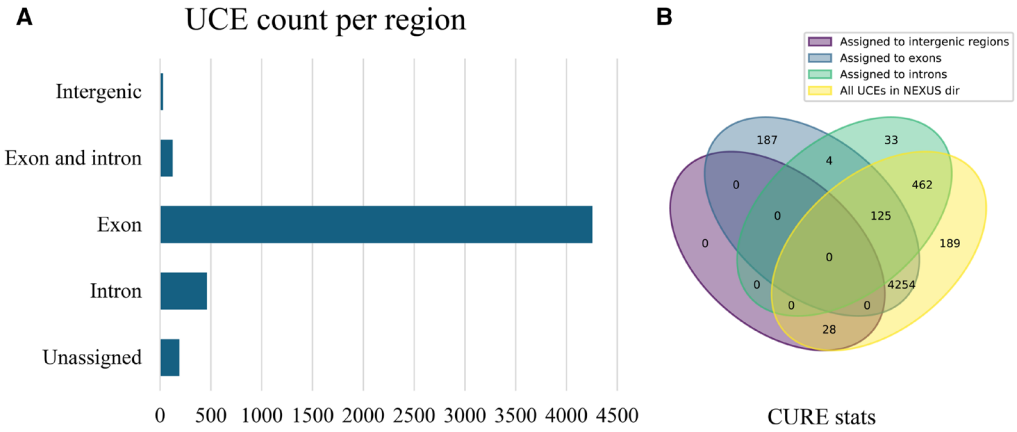

Fig. 4. Curation of UCEs captured from *Sisyra nigra* genome with the universal probe set, using GeneRegion strategy of CURE. A) The bar graph shows the number of UCEs located in each region. B) Same information but in the output file from CURE.

Table 5. PHYLUCE summary of the alignment matrix for the superorder's dataset, generated using UCEs captured by the combined tailored probes, combined tailored and universal probes, and universal probes designed during in-silico test

Probe sets	Alignment length	Mean	Min. length (bp)	Max. length (bp)	Total characters	Missing data characters	70% C.A.M.
neuropterida_sisnig_v1 str 16	1,541,513	323.3	46	1,240	44,860,604	2,260,717	104
combined_orders	298,946	373.22	9	1,241	8,745,345	595,413	29
combined_universal_tailored	1,587,834	326.31	46	1,381	45,775,167	2,330,202	103

The total number of ingroup was 102 taxa, plus 1 outgroup. bp, base pairs, C.A.M, complete alignment matrix, str, stringency.

Table 6. PHYLUCE summary of the alignment matrix for each order-specific dataset, generated using UCEs captured by the tailored probes, subsets of the combined tailored and universal probes, and the subsets of the universal probe set designed for the focal groups during in-silico test

Probe sets	Alignment length	Mean	Min. length (bp)	Max. length (bp)	Total characters	Missing data characters	70% C.A.M
subset_megaloptera	2,651,423	902.77	189	1,394	13,925,141	1,059,383	396
subset_neuroptera	1,584,909	345.67	57	1,364	33,436,136	1,810,495	99
subset_raphidioptera	2,302,192	643.25	57	1,380	13,994,167	1,084,895	535
tailored_megaloptera	2,407,335	853.97	180	1,383	14,351,372	954,469	1,054
tailored_neuroptera	1,791,408	368.68	70	1,352	38,440,876	2,144,979	79
tailored_raphidioptera	2,105,913	559.34	103	1,385	16,165,372	1,391,610	1,650
combined_subset_megaloptera	2,786,640	905.05	189	1,394	14,658,907	1,115,082	420
combined_subset_neuroptera	1,640,502	347.71	57	1,364	35,029,411	1,899,720	103
combined_subset_raphidioptera	2,374,779	641.49	57	1,380	14,551,341	1,129,703	633

The total number of ingroup taxa differed for each order-specific experiments (81 for Neuroptera, 9 for Megaloptera, and 12 for Raphidioptera), plus 1 outgroup. bp, base pairs, C.A.M, complete alignment matrix.

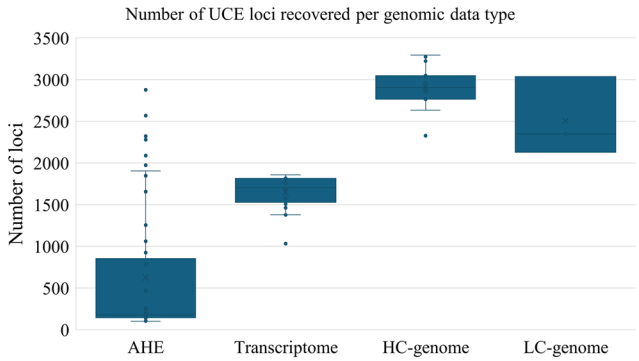

Fig. 5. Boxplot graph showing the number of UCEs captured from different genetic sequences using the newly designed universal probe set, neuropterida-sisnig-v1. AHE, Anchored hybrid enrichment; UCE, ultraconserved elements; HC, high coverage; LC, low coverage.

Sub-Sets vs tailored Probe Sets

Due to the broad and well-representative taxon sampling we used, the final probe set designed for Neuropterida (neuropterida_sisnig_v1) contained a large probe count (172,861). The subset of the universal probe list contained 84,015 probes targeting 5,289 conserved loci (subset_superorder) (Table 3). The final number of alignments obtained in the 70% matrix was slightly higher when using the subset than the universal probe list (2,499 vs 2,480) (Table 4).

The search for shared loci among the probe sets designed individually for Megaloptera, Neuroptera, and Raphidioptera identified 805 nonproblematic UCEs (nonparalogs or divergent sequences). For simplicity, we named this dataset containing 805 UCEs as "combined_orders." When comparing the combined_orders with the universal probe list (neuropterida_sisnig_v1), the alignment matrix obtained using neuropterida_sisnig_v1 included a higher number of loci. For the first experiment, in which we used only 17 genomes, the final alignment matrix using the combined_orders dataset included 436 alignments, while using neuropterida_sisnig_v1, we recovered 2,480 alignments (Table 4). For the second experiment, targeting UCEs from different genetic data type, the combined_orders dataset performed even more poorly than in the previous experiment. The final 70% complete alignment matrix contained 29 alignments when using the combined_orders and 104 when using neuropterida_sisnig_v1 (Table 5).

For the third experiment, the subset of the universal probes for each order was not as efficient as the probe set designed specifically for each order when comparing the number of alignments obtained (Table 6). For Megaloptera, by using the tailored probe set, we obtained more than double the number of alignments in the 70% matrix (1,054) when compared to the subset (396). The tailored and subset probes for Neuroptera overall yielded similar amounts of data but, with subsetting, we acquired a slightly larger number of alignments than using the tailored probe set (99 vs 79). Lastly, for Raphidioptera, the final 70% complete alignment matrix comprised 1,650 loci when using the tailored probe set, as opposed to 585 alignments when using the subset.

For the final probe set, we identified unique UCEs present in the tailored probes and absent in the universal probe set using the bash and R-script "comparing_monolithic_UCE_fastas" pipeline (Alexander 2018, Gustafson et al. 2019). We included a total of 288 unique UCEs in the universal probe set to increase its efficiency when capturing loci from the individual orders (Table 2). Using the combined probe set ("neuropterida_combined_universal_tailored.fasta"), the number of aligned loci increased in the final 70% complete alignment matrix, when compared to the universal probe set, for all experiments (Tables 4 and 6), except the larger in-silico test (Table 5). The results for the latter experiment were equivalent when using the universal and the combined probes (104 vs 103, respectively).

Phylogenetic Results

The final alignment matrix included 28,036 characters, 64 partitions, and 112 terminal taxa. We provide the number of UCEs recovered from AHE, genomes, and transcriptomes in Fig. 5.

The phylogenetic analysis using UCEs recovered a robust topology for the major lineages of Neuropterida, with Raphidioptera as the sister group to Megaloptera + Neuroptera (Fig. 6). All 3 orders were recovered as monophyletic with high branch support (98% or more). Regarding the internal relationships of Raphidioptera, the order was divided into 2 highly supported clades, 1 composed of species of Inocelliidae and the other of species of Raphidiidae. We also recovered the 2 megalopteran families, Sialidae and Corydalidae, as monophyletic. The inferred relationships within Neuroptera indicate that most families are monophyletic, with Coniopterygidae (Coniopterygoidea) as the sister group to all other lacewings. The Osmyloidea clade, comprising Sisyridae + (Osmylidae +

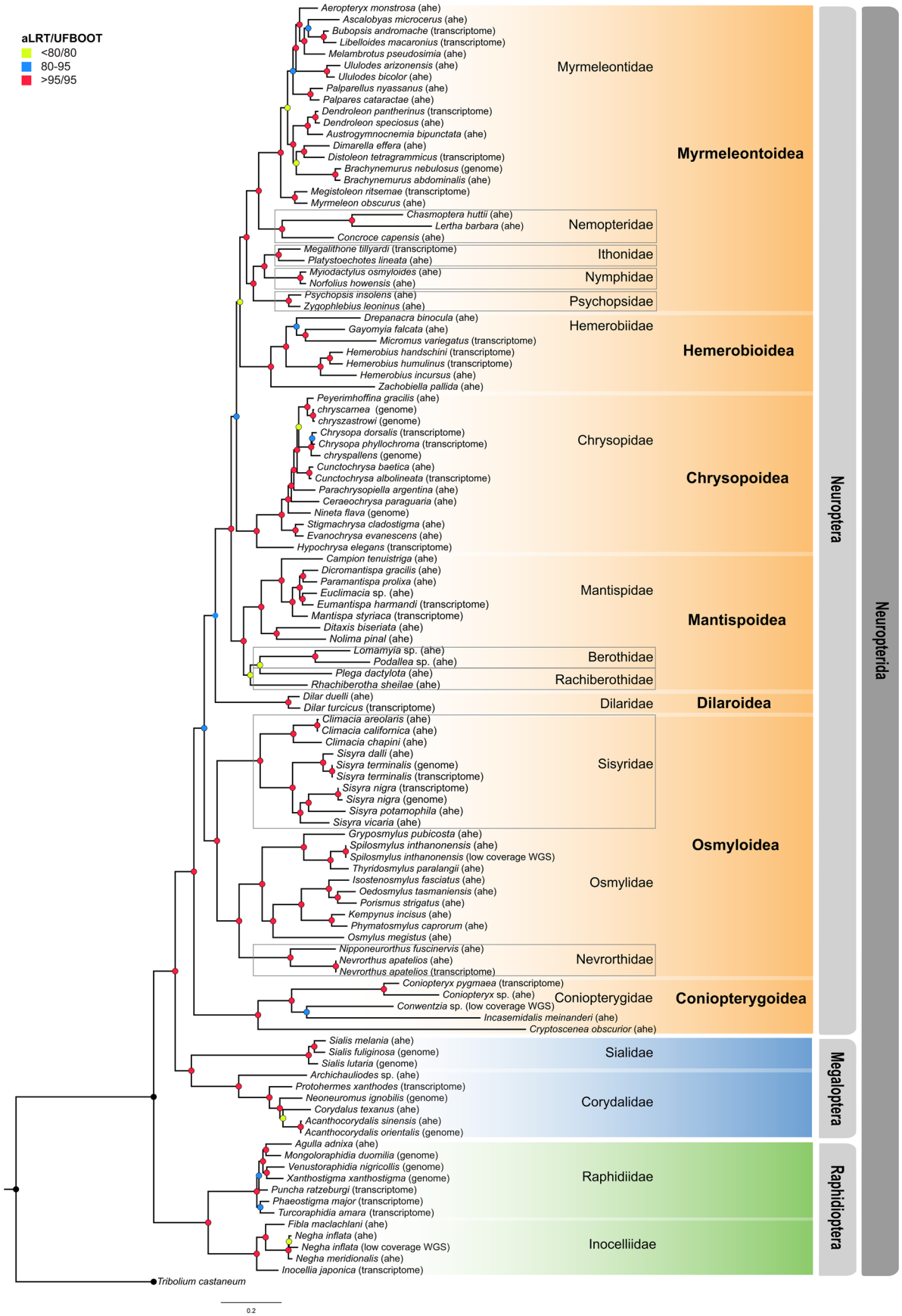

Fig. 6. Phylogenetic relationships among Neuropterida taxa from maximum likelihood analysis of UCEs captured with the neuropterida-sisnig-v1 probe set. The node support is represented by color-coded circles. The statistical analyses for branch support were aLRT (approximate likelihood ratio test) and UFBOOT (ultrafast bootstrap approximation). Color-inclusive design (Stevens et al. 2025).

Nevrorthidae), was recovered with high support. The clade consisting of Osmyloidea and the remaining Neuroptera (without Coniopterygidae) was moderately supported (~90%).

The subsequent branching lineage, Dilaridae, was recovered as the sister group to the remaining Neuroptera, and the family is placed in its own superfamily, Dilaroidea. The Mantispoidea clade was well supported, comprising 3 families, the monophyletic Mantispidae as the sister to Rachiberothidae + Berothidae (Ardila-Camacho et al. 2021). Although Rachiberothidae was recovered as paraphyletic due to *P. dactylota* being placed nearest to Berothidae, the support for these relationships was low. Chrysopidae is the only family within Chrysopoidea and its monophyly is generally well supported. The sister group relationship between Hemerobioidea and Myrmeleontoidea was recovered with extremely low branch support, but Hemerobioidea (comprising only the family Hemerobiidae) was well-supported. Five families were recovered within Myrmeleontoidea, and 2 well defined clades, 1 composed of Psychopsidae + (Nymphidae + Ithonidae), and the other containing Nemopteridae and Myrmeleontidae as sister groups.

Discussion

Our study reinforces previous findings of Gustafson et al. (2019, 2023) that the genome selected as the base significantly alters probe set performance. The most optimal base genome found in our experiment (genome of *S. nigra*) resulted in ~1,800 more alignments than the base genome recovering the fewest number of conserved loci (Table 2). Gustafson et al. (2019) demonstrated that the putative loci count was not a reliable parameter to identify the best base genome, and they also identified the trend that the taxa with the fewest number of putative loci also resulted in fewer recovered loci. Similarly, putative loci counts were not a good indicator of the best base genome detected in our study. Additionally, the base genomes that performed poorly in this first step were likely to continue recovering fewer loci. Our results suggest that selecting a few genomes that performed better during the putative loci identification for the subsequent steps of the workflow will save time and ensure that the best base genome is identified.

Overall, traditional metrics for assessing genome quality, such as coverage and completeness, proved to be of little importance when selecting the base genome. The genome of *V. nigricollis* had the highest BUSCO scores (96.6%), while *A. orientalis* had the highest coverage (258×) and neither genome performed better than *S. nigra* as the best base genome. These results are concordant with those of Gustafson et al. (2019, 2023), who also demonstrated that genome metrics have no correlation with base genome performance. Ideally, well-assembled and annotated genomes are recommended as base genomes, as well as sequences within the focus group (Faircloth 2017). Unfortunately, the only annotated genome available for Neuropterida is *C. carnea* (Stephens), which was used as the base genome in Zhang et al. (2025) to design the UCE probe set they used ("Neuropterida-5Kv1," loci = 5,341). However, this sequence did not perform well as the base genome when compared to others in our study. This reinforces the need for additional annotated genomes for Neuropterida.

When testing how permissive we should be during the stringency test, we considered the results obtained by Gustafson et al. (2019) when setting the number of taxa required in the putative loci step. However, selecting the base genome plus 1 taxon, for example, would result in an exceptionally large number of probes at the end due to the elevated number of exemplar taxa included in this study. By choosing the base genome plus 10 taxa, the resulting probe sets recovered considerably high numbers of UCEs, even with a strict probe set. The broader probe resulted in a higher number of probes and recovered loci.

For bioinformatic analyses and in-silico tests, a larger number of probes in the final universal probes would not be prohibitive. However, synthesizing this high number of probes for in-vitro enrichment might be expensive (Derkarabetian et al. 2025), which makes smaller probe sets or the subsetting universal probes with only the desired taxa potentially desirable from a cost-effectiveness standpoint. In our tests, the smaller strict probe set performed as efficiently as the broad probe set, and the resulting tree was congruent across different levels of stringency. As an alternative, we also included here a filtered subset of the tailored and universal combined probes, targeting 2,038 UCEs and with a 33,311 probe count, which is more financially accessible.

Regarding the experiment testing the efficiency of the universal probe set in comparison with combining data obtained from different tailored probe sets for Neuropterida (Table 5), we observed that the universal probes outperformed the tailored ones, in contrast with the results obtained by Gustafson et al. (2020). Tailored probe sets target different homologous loci, and although some regions may be equivalent, the resulting datasets might not completely overlap. That is, a combined matrix composed of UCEs recovered from several tailored probes will likely have a reduced number of shared conserved loci because most of the regions will not correspond. Additionally, Gustafson et al. (2020, 2023) demonstrated that combining generalized probe sets with tailored probes improved the phylogenetic utility of probes. Our results reinforce this finding considering that our combined tailored and universal probe set recovered slightly more alignments in the final step. Moreover, Gustafson et al. (2023) highlighted that the tailored-only probes did not perform well when applied to phylogenetically distant taxa.

When comparing the tailored probe set and the subset of the universal probe list for each order, we obtained similar results for Neuroptera in both analyses, while for Raphidioptera and Megaloptera, the tailored probe set was more efficient in downstream analyses. Even though we only targeted loci present in all 3 orders during bait design for the universal probe set, we recovered a surprisingly low number of aligned loci across species of Raphidioptera and Megaloptera using the subsets (Table 6). We believe that the base genome choice was the underlying cause of this issue. During bait design, PHYLUCE requires a similarity of at least 95% between the base genome and exemplar taxa (Branstetter et al. 2017), and probably the tailored probes of Raphidioptera and Megaloptera identified more loci unique to these clades.

Even though the 2 datasets differ in the amount of data, the resulting topology is virtually the same under analyses using the same parameters. Additionally, we captured UCEs from transcriptomes and targeted enriched sequences, which are expected to yield fewer data than genomes. The subsets of the universal probe list for each order recovered an average of 2,930 UCEs from genomes, while the tailored probe sets recovered an average of 3,093 loci (log-files of in silico test > phyluce_assembly_get_fasta_from_match_counts). Therefore, synthesizing the subset probe for target enrichment of UCEs

from WGS offers higher universality of the dataset and a considerably high number of conserved loci.

Finally, the complex task of estimating the phylogenetic relationships of Neuropterida seems to progress in the light of every new discovery. Recent phylogenomic studies have investigated different regions of the genome of species of Neuropterida to uncover their evolutionary history, including mitogenomes (Wang et al. 2017), AHE (Winterton et al. 2018), transcriptomes (Vasilikopoulos et al. 2020), USCOs (Lai et al. 2025), and most recently, UCEs (Zhang et al. 2025). Our resulting topology reinforces previous results for the relationships among the major lineages of this superorder (Misof et al. 2014, Wang et al. 2017, Engel et al. 2018, Winterton et al. 2018, Vasilikopoulos et al. 2020, Lai et al. 2025). Additionally, our results are concordant with previous hypotheses of Coniopterygidae as the sister group to the remaining Neuroptera, and Sisyridae, Osmylidae, and Nevrorthidae comprising the monophyletic Osmyloidea (Winterton et al. 2010, 2018, Wang et al. 2017, Engel et al. 2018, Vasilikopoulos et al. 2020, Lai et al. 2025, Zhang et al. 2025). However, the placement of some families varied when compared to previous phylogenetic hypotheses.

The placement of Mantispoidea as sister to all families of Neuroptera, except for Coniopterygoidea, Osmyloidea, and Dilaroidea, was also indicated by Wang et al. (2017), Engel et al. (2018), and Lai et al. (2025). There are conflicting hypotheses for the internal relationship of Mantispoidea in recent studies, including Rachiberothidae as sister group of Berothidae + Matispidae (Winterton et al. 2018), or Mantispidae closely related to Rachiberothidae (Wang et al. 2017, Engel et al. 2018). Here, we recovered Matispidae as the sister group to Berothidae and Rachiberothidae. Our results placing Hemerobioidea as the sister group to Myrmeleontoidea are similar to those of Vasilikopoulos et al. (2020), but within Myrmeleontoidea, we recovered a clade containing Psychopsidae + (Nymphidae + Ithonidae), similar to Winterton et al. (2018) and Machado et al. (2019). Both traditional Myrmeleontidae and Ascalaphidae were recovered as paraphyletic with respect to each other, with similar findings obtained by Winterton et al. (2018) and Machado et al. (2019).

Conclusion

Our study aimed to improve locus recovery while designing a universal probe set for targeting UCEs of Neuropterida (Neuroptera, Megaloptera, Raphidioptera). Given the increasing number of tailored probe sets being designed for various focus groups, creating a representative and well-curated universal probe set for a particular lineage allows for an integrative dataset of homologous loci. Our final, combined probe set targets 5,557 conserved loci and is composed of 180,449 probes (neuropterida_combined_universal_tailored). Although it was developed to cover 3 different orders, the combined universal probe set, and its subsets, worked well for targeting UCEs across different taxonomic levels. Additionally, combining future tailored probes with this probe set will maintain the compatibility of the datasets and increase the number of phylogenetically informative UCEs being enriched. Here, we provided 5 subsets of the master probe list: (i) 2 subsets for the superorder, (ii) subset for Megaloptera, (iii) subset for Neuroptera, and (iv) subset for Raphidioptera. For future studies requiring in-vitro enrichment, we highly recommend using the filtered subset for the superorder (combined-subset -neuropterida-v1) as it represents the highest cost efficiency, whilst the remaining probe sets were found to be more financially costly.

Based on the data acquired using the newly developed probe set, we recovered Neuropterida as monophyletic, with Raphidioptera as the sister group to Neuroptera + Megaloptera. The internal relationships of these taxa were also congruent with previous phylogenetic hypotheses. Although many aspects of the phylogeny of Neuropterida remain controversial, which might be associated with the rapid diversification of the group during the Mesozoic, recent phylogenomic studies have elucidated multiple recalcitrant nodes within this lineage. We expect that our newly developed probe set will enable access to novel and comprehensive datasets for Neuropterida, as well as foster further investigation of the evolutionary history of this enigmatic relictual group.

Specimen Collection Statement

Insect Systematics and Diversity supports compliance with the Nagoya Protocol. The authors attest that all legal and regulatory requirements, including export and import collection permits, have been followed for the collection of specimens from source populations at any international, national, regional, or other geographic level for all relevant field specimens collected as part of this study.

Acknowledgements

We sincerely thank Dr Shaun Winterton for the loan of specimens for DNA extraction. Thank you to the CNC and Dr Sophie Cardinal, in particular, for making their molecular facility available to the authors. Special gratitude to Scott Kelso for his assistance with the DNA purification steps. We thank the FARM computing cluster at University of California in Davis. We also thank Stéphanie Boucher, Curator of the Lyman Entomological Museum, for her support, insight, and zealous management of our research space.

Author contributions

Alice C. Assmar (Conceptualization [lead], Data curation [lead], Formal analysis [equal], Funding acquisition [equal], Project administration [lead], Writing—original draft [lead], Writing—review & editing [equal]), Ziad Khouri (Conceptualization [supporting], Data curation [supporting], Formal analysis [equal], Validation [equal], Writing—review & editing [equal]), Renato J.P. Machado (Resources [equal], Supervision [equal], Writing—review & editing [equal]), and Jessica P. Gillung (Conceptualization [supporting], Funding acquisition [equal], Resources [equal], Supervision [equal], Writing—review & editing [equal])

Supplementary material

Supplementary material is available at *Insect Systematics and Diversity* online.

Funding

ACA is grateful for her doctoral scholarship (B2X 2023-2024, doi: https://doi.org/10.69777/330130) from the Fonds de

Recherche du Quebec—Nature et Technologies (FRQNT). JPG acknowledges funding from NSERC-DG (NSERC RGPIN-2021-02656). RJPM is grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the grants: Proc. 402785/2021-5 and Proc. 441787/2023-1.

Conflicts of interest

None declared.

Data availability

Published low-coverage WGS data are available at NCBI (BioProject PRJNA1329311). Data, probes, log files, and supplementary materials are publicly available from FigShare (https://doi.org/10.6084/m9.figshare.30954212). Accession numbers for the genomes used in this study are presented in Supplementary Table SM3.

References

Alexander A. 2018. comparing_monolithic_UCE_fastas v0.2 [Software]. https://github.com/laninsky/comparing_monolithic_UCE_fastas

Andrews S. 2023. FastQC: a quality control tool for high throughput sequence data [Software]. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Ardila-Camacho A, Martins CC, Aspöck U, et al. 2021. Comparative morphology of extant raptorial Mantispoidea (Neuroptera: Mantispidae, Rhachiberothidae) suggests a non-monophyletic Mantispidae and a single origin of the raptorial condition within the superfamily. *Zootaxa* 4992:1–89. https://doi.org/10.11646/zootaxa.4992.1.1

Aspöck U. 2002. Phylogeny of the Neuropterida (Insecta: Holometabola). *Zool. Scr.* 31:51–55. https://doi.org/10.1046/j.0300-3256.2001.00087.x

Baca SM, Alexander A, Gustafson GT, et al. 2017. Ultraconserved elements show utility in phylogenetic inference of Adephaga (Coleoptera) and suggest paraphyly of "Hydradephaga." *Syst. Entomol.* 42:786–795. https://doi.org/10.1111/syen.12244

Baca SM, Gustafson GT, DeRaad DA, et al. 2025. A shallow-scale phylogenomics approach reveals repeated patterns of diversification among sympatric lineages of cryptic Neotropical aquatic beetles (Coleoptera: Noteridae). *Syst. Entomol.* 50:82–101. https://doi.org/10.1111/syen.12643

Bejerano G, Pheasant M, Makunin I, et al. 2004. Ultraconserved elements in the human genome. *Science* 304:1321–1325. https://doi.org/10.1126/science.1098119

Blaimer BB, Brady SG, Schultz TR, et al. 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. *BMC Evol. Biol.* 15:271. https://doi.org/10.1186/s12862-015-0552-5

Blaimer BB, Lloyd MW, Guillory WX, et al. 2016. Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLoS One*. 11:e0161531. https://doi.org/10.1371/journal.pone.0161531

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

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

Branstetter MG, Longino JT, Ward PS, et al. 2017. Enriching the ant tree of life: enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. *Methods Ecol. Evol.* 8:768–776. https://doi.org/10.1111/2041-210X.12742

Broad GR, Crowley LM, McCulloch J; Darwin Tree of Life Consortium. 2023. The genome sequence of the Black Spongefly, *Sisyra nigra* (Retzius, 1783). *Wellcome Open Res*. 8:511. https://doi.org/10.12688/wellcomeopenres.20295.1

Buenaventura E, Lloyd MW, Perilla López JM, et al. 2021. Protein-encoding ultraconserved elements provide a new phylogenomic perspective of Oestroidea flies (Diptera: Calyptratae). *Syst. Entomol.* 46:5–27. https://doi.org/10.1111/syen.12443

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17:540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Cock PJA, Antao T, Chang JT, et al. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25:1422–1423. https://doi.org/10.1093/bioinformatics/btp163

Danecek P, Bonfield JK, Liddle J, et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience*. 10:giab008. https://doi.org/10.1093/gigascience/giab008

Derkarabetian S, Benavides LR, Giribet G. 2019. Sequence capture phylogenomics of historical ethanol-preserved museum specimens: unlocking the rest of the vault. *Mol. Ecol. Resour.* 19:1531–1544. https://doi.org/10.1111/1755-0998.13072

Derkarabetian S, Lord A, Angier K, et al. 2025. Covering all bases: a universal metazoan UCE probe set to democratize phylogenomics. *Genome Biol. Evol.* 17:1–11. https://doi.org/10.1093/gbe/evaf193

Engel MS, Winterton SL, Breitkreuz LCV. 2018. Phylogeny and evolution of neuropterida: where have wings of lace taken us? *Annu. Rev. Entomol.* 63:531–551. https://doi.org/10.1146/annurev-ento-020117-043127

Faircloth BC, McCormack JE, Crawford NG, et al. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* 61:717–726. https://doi.org/10.1093/sysbio/sys004

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

Faircloth BC. 2016b. PHYLUCE: software for UCE (and general) phylogenomics. Phyluce Softw. UCE Gen. Phylogenomics. https://phyluce.readthedocs.io/en/latest/

Faircloth BC. 2017. Identifying conserved genomic elements and designing universal bait sets to enrich them. *Methods Ecol. Evol.* 8:1103–1112. https://doi.org/10.1111/2041-210X.12754

Faircloth BC, Branstetter MG, White ND, et al. 2015. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Mol. Ecol. Resour.* 15:489–501. https://doi.org/10.1111/1755-0998.12328

Freitas FV, Branstetter MG, Franceschini-Santos VH, et al. 2023. UCE phylogenomics, biogeography, and classification of long-horned bees (Hymenoptera: Apidae: Eucerini), with insights on using specimens with extremely degraded DNA. *Insect Syst. Divers.* 7:3. https://doi.org/10.1093/isd/ixad012

Gurevich A, Saveliev V, Vyahhi N, et al. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Gustafson GT, Alexander A, Sproul JS, et al. 2019. Ultraconserved element (UCE) probe set design: base genome and initial design parameters critical for optimization. *Ecol. Evol.* 9:6933–6948. https://doi.org/10.1002/ece3.5260

Gustafson GT, Baca SM, Alexander A, et al. 2020. Phylogenomic analysis of the beetle suborder Adephaga with comparison of tailored and generalized ultraconserved element probe performance. *Syst. Entomol.* 45:552–570. https://doi.org/10.1111/syen.12413

Gustafson GT, Glynn RD, Short AEZ, et al. 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 Syst. Divers.* 7:4. https://doi.org/10.1093/isd/ixad014

Haring E, Aspöck U. 2004. Phylogeny of the Neuropterida: a first molecular approach: phylogeny of the Neuropterida. *Syst. Entomol.* 29:415–430. https://doi.org/10.1111/j.0307-6970.2004.00263.x

Hellemans S, Wang M, Hasegawa N, et al. 2022. Using ultraconserved elements to reconstruct the termite tree of life. *Mol. Phylogenet. Evol.* 173:107520. https://doi.org/10.1016/j.ympev.2022.107520

Hoang DT, Chernomor O, Von Haeseler A, et al. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35:518–522. https://doi.org/10.1093/molbev/msx281

Huang W, Li L, Myers JR, et al. 2012. ART: a next-generation sequencing read simulator. *Bioinformatics* 28:593–594. https://doi.org/10.1093/bioinformatics/btr708

Kalyaanamoorthy S, Minh BQ, Wong TKF, et al. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods*. 14:587–589. https://doi.org/10.1038/nmeth.4285

Kapli P, Yang Z, Telford MJ. 2020. Phylogenetic tree building in the genomic age. *Nat. Rev. Genet*. 21:428–444. https://doi.org/10.1038/s41576-020-0233-0

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780. https://doi.org/10.1093/molbev/mst010

Lai Y, Du S, Li H, et al. 2025. Lacewing-specific universal single copy orthologs designed towards resolution of backbone phylogeny of Neuropterida. *Syst. Entomol*. 50:309–324. https://doi.org/10.1111/syen.12657

Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res*. 21:936–939. https://doi.org/10.1101/gr.111120.110

Machado RJP, Gillung JP, Winterton SL, et al. 2019. Owlflies are derived antlions: anchored phylogenomics supports a new phylogeny and classification of Myrmeleontidae (Neuroptera). *Syst. Entomol*. 44:418–450. https://doi.org/10.1111/syen.12334

Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol*. 37:1530–1534. https://doi.org/10.1093/molbev/msaa015

Misof B, Liu S, Meusemann K, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763–767. https://doi.org/10.1126/science.1257570

Nurk S, Bankevich A, Antipov D, et al. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads. In: Deng M, Jiang R, Sun F, editors. *Research in computational molecular biology*. Vol. 7821. Springer Berlin Heidelberg. (Hutchison D, Kanade T, Kittler J, et al., editors. Lecture notes in computer science). p. 158–170. https://doi.org/10.1007/978-3-642-37195-0_13

Oswald JD, Machado RJP. 2018. Biodiversity of the Neuropterida (Insecta: Neuroptera, Megaloptera, and Raphidioptera). In: Foottit RG, Adler PH, editors. *Insect biodiversity*. John Wiley & Sons, Ltd. p. 627–672. https://doi.org/10.1002/9781118945582.ch21

Prjibelski A, Antipov D, Meleshko D, et al. 2020. Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinformatics*. 70:e102. https://doi.org/10.1002/cpbi.102

QIAGEN. 2014. QIAamp DNA Micro Handbook. qiagen.com. https://www.qiagen.com/us/resources/resourcedetail?id=085e6418-1ec0-45f2-89eb-62705f86f963&lang=en

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. https://doi.org/10.1093/bioinformatics/btq033

Saenz Manchola OF, Samacá Sáenz E, Virrueta Herrera S, et al. 2022. Mining ultraconserved elements from transcriptome and genome data to explore the phylogenomics of the free-living lice suborder Psocomorpha (Insecta: Psocodea). *Insect Syst. Divers*. 6:1. https://doi.org/10.1093/isd/ixac010

Simão FA, Waterhouse RM, Ioannidis P, et al. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Sota T, Takami Y, Ikeda H, et al. 2022. Global dispersal and diversification in ground beetles of the subfamily Carabinae. *Mol. Phylogenet. Evol*. 167:107355. https://doi.org/10.1016/j.ympev.2021.107355

Stevens HP, Winegar CV, Oakley AF, et al. 2025. Estimating the prevalence of images in the biology literature that are problematic for people with a color-vision deficiency. Colorblind Image Tester. https://bioapps.byu.edu/colorblind_image_tester/.

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

Van Dam AR, Covas Orizondo JO, Lam AW, et al. 2022. Metagenomic clustering reveals microbial contamination as an essential consideration in ultraconserved element design for phylogenomics with insect museum specimens. *Ecol. Evol* 12:e8625. https://doi.org/10.1002/ece3.8625

Van Dam MH, Henderson JB, Esposito L, et al. 2021. Genomic characterization and curation of UCEs improves species tree reconstruction. *Syst. Biol*. 70:307–321. https://doi.org/10.1093/sysbio/syaa063

Vasilikopoulos A, Misof B, Meusemann K, et al. 2020. An integrative phylogenomic approach to elucidate the evolutionary history and divergence times of Neuropterida (Insecta: Holometabola). *BMC Evol. Biol*. 20:64. https://doi.org/10.1186/s12862-020-01631-6

Wang Y, Liu X, Garzón-Orduña IJ, et al. 2017. Mitochondrial phylogenomics illuminates the evolutionary history of Neuropterida. *Cladistics* 33:617–636. https://doi.org/10.1111/cla.12186

Winterton SL, Hardy NB, Wiegmann BM. 2010. On wings of lace: phylogeny and Bayesian divergence time estimates of Neuropterida (Insecta) based on morphological and molecular data. *Syst. Entomol*. 35:349–378. https://doi.org/10.1111/j.1365-3113.2010.00521.x

Winterton SL, Lemmon AR, Gillung JP, et al. 2018. Evolution of lacewings and allied orders using anchored phylogenomics (Neuroptera, Megaloptera, Raphidioptera). *Syst. Entomol*. 43:330–354. https://doi.org/10.1111/syen.12278

Yang Z, Rannala B. 2012. Molecular phylogenetics: principles and practice. *Nat. Rev. Genet*. 13:303–314. https://doi.org/10.1038/nrg3186

Young AD, Gillung JP. 2020. Phylogenomics—principles, opportunities and pitfalls of big-data phylogenetics. *Syst. Entomol*. 45:225–247. https://doi.org/10.1111/syen.12406

Zhang F, Ding Y, Zhu CD, et al. 2019. Phylogenomics from low-coverage whole-genome sequencing. *Methods Ecol. Evol*. 10:507–517. https://doi.org/10.1111/2041-210X.13145

Zhang R, Wang L, Tian S, et al. 2025. Inconsistent performance of multi-type genomic data in phylogenomics of neuropteridan insects, with solutions toward conflicting results. *Syst. Entomol. Early* 50:855–875. https://doi.org/10.1111/syen.12684.

Zhao Y, Zhan Q, Wang Y, et al. 2025. Chromosome-level genome of the brown lacewing *Micromus angulatus* (Stephens, 1836) (Neuroptera: Hemerobiidae). *Sci. Data*. 12:394. https://doi.org/10.1038/s41597-025-04739-w