Phylogenomics and Fossilized Birth-Death Dating Reveals Gondawanan Origin and Extensive Post-Cretaceous Diversification of Worldwide Cicadidae.

Mark Stukel[1,2*], Nicholas J. Matzke[3,5], Jordan Douglas[4,5,6], Tatiana Petersen Ruschel[7,8], Stéphane Puissant[9], Ben W. Price[10], Martin Villet[11,12], Alan R. Lemmon[13], Emily Moriarty Lemmon[14], Chris Simon[1]

[1]Ecology & Evolutionary Biology; University of Connecticut; Storrs, CT 06269 USA

[2]Department of Entomology & Nematology; University of California, Davis; Davis, CA 95616 USA

[3]School of Biological Sciences, University of Auckland, Auckland 1010, New Zealand

[4]Department of Physics, University of Auckland, Auckland 1010, New Zealand

[5]Centre for Computational Evolution, University of Auckland, Auckland 1010, New Zealand

[6]Research School of Biology, Australian National University, Canberra 0200, ACT, Australia

[7]Programa de Pós-Graduação em Biologia Animal, Departamento de Zoologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[8]Programa de Pós-Graduação em Entomologia, Departamento de Zoologia, Universidade Federal do Paraná, Curitiba, PR, Brazil

[9]Muséum d'Histoire naturelle de Dijon, 14, rue Jehan de Marville, F-21000 Dijon, France

[10]The Natural History Museum, London, UK

[11]Department of Zoology & Entomology, Rhodes University, Makhanda, South Africa

[12]Research and Exhibitions Department, South African Museum, Iziko Museums of South Africa, Cape Town, South Africa

[13]Department of Scientific Computing, Florida State University, Tallahassee, FL 32306 USA

[14]Department of Biological Science, Florida State University, Tallahassee, FL 32306 USA

*Corresponding author:

Mark Stukel

Department of Entomology & Nematology

University of California, Davis

Davis, CA 95616

USA

Tel: +1 (630) 205-1193

Email: mstukel@ucdavis.edu

Abstract

The Cretaceous-Paleogene (K-Pg) mass extinction resulted in a massive turnover in biodiversity. The oldest fossils of the globally-distributed insect family Cicadidae date to the Paleocene, suggesting that this family diversified after the K-Pg boundary. We analyzed 490 nuclear Anchored Hybrid Enrichment loci as well as mitochondrial genomes for the Cicadidae, sampling all five subfamilies, 85% of tribes, and 25% of genera, using concatenated maximum-likelihood and multi-species coalescent approaches to resolve the phylogenetic relationships of Cicadidae subfamilies. We estimated divergence times of Cicadidae lineages using a fossilized birth-death model augmented into the multispecies coalescent with 44 fossil taxa. We estimated a Cretaceous origin for Cicadidae with four of the five subfamilies diversifying shortly after the K-Pg extinction event. Our fossilized birth-death tip-dating approach improved the precision of age estimates for many cicada clades compared to dates based on node-calibration from previous studies. An ancestral area reconstruction inferred a Gondwanan origin for Cicadidae with multiple lineages dispersing across the planet. Our results augment insight into how the K-Pg mass extinction shaped present-day diversity.

Keywords: Phylogenomics, Bayesian, multi-species coalescent, divergence time estimation, K-Pg, Cicadidae, biogeography

Introduction

Tming and geography of diversification events are important for understanding processes that generate biodiversity. In geological time, mass extinctions played a crucial role in shaping biodiversity through accelerated extinction and subsequent diversification. The most recent mass extinction event, the Cretaceous-Paleogene (K-Pg) transition approximately 66 million years ago, involved around 76% of all species, and helped to shape which evolutionary lineages are most diverse today (Schulte et al. 2010). The mechanism of post-extinction diversification is thought to be adaptive radiation, when lineages surviving mass extinctions face much less ecological competition (Erwin 2001; Jablonski 2001; Hull 2015; Hoyal Cuthill et al. 2020). Understanding mass extinction events and their associated evolutionary radiations provides insight into the historical development of biodiversity and the processes that may shape future biodiversity.

Studies using a handful of genes are frequently sufficient to infer phylogenetic trees for taxa that underwent well-spaced speciation events. However, these genes often do not have enough phylogenetic information to resolve the short branches that are typical of episodic evolutionary radiations (Whitfield and Kjer 2008). Furthermore, the assumption implicit in analyses of concatenated sequences is violated for most datasets: that individual genes all share the same evolutionary history as the species tree (Degnan and Rosenberg 2006; Salichos and Rokas 2013). Concatenated sequences can therefore lead to biased estimates of substitution rates and divergence times (Kubatko et al. 2011; Mendes and Hahn 2016; Ogilvie et al. 2016), and also inconsistently estimate the tree topology (Heled and Drummond 2010; Ogilvie et al. 2017).

This necessitates the use of the multispecies coalescent (MSC) to create species trees from unconcatenated genes (Maddison 1997; Degnan and Rosenberg 2009; Edwards 2009; Liu et al. 2009). Genomic datasets and analytical methods that account for deep coalescence (also called incomplete lineage sorting, or ILS) have shown their utility for phylogenetic problems such as evolutionary radiations (Hime et al. 2020; Owen and Miller 2022; Smith et al. 2023; Stukel et al. 2024). However, the increased data in phylogenomic studies have been problematic for divergence time estimation, because Bayesian molecular dating methods are too computationally intensive to handle hundreds of genes for many taxa. Some strategies for reducing the computational complexity include: fixing the maximum-likelihood tree topology and estimating divergence times using a subset of clock-like genes (Smith et al. 2018); estimating divergence times using relaxed clock models in a maximum-likelihood framework (Smith and O'Meara 2012; Tamura et al. 2012); dividing a phylogeny into a backbone tree and subtrees for separate inference, followed by rejoining subtrees to the backbone tree; or using multiple smaller samples of genes and assessing whether they produce similar species phylogenies.

Study system

Cicadas (Cicadoidea, Auchenorrhyncha, Hemiptera) are an excellent model lineage of insects for understanding evolutionary radiations. Cicadoidea contains two families: Tettigarctidae (hairy cicadas) and Cicadidae (singing cicadas) (Moulds 2005). Cicadidae is distributed globally (except for Antarctica) with over 3,400 species, while Tettigarctidae is restricted to southeastern Australia and Tasmania with two extant species (World Auchenorrhyncha Database, accessed 18 April 2025). Cicadidae comprises five subfamilies:

Cicadinae and Cicadettinae, both with global distributions; Tibicininae, distributed across the Neotropics, Holarctic, and North Africa (Puissant and Gurcel 2023); Tettigomyiinae, in Africa and Madagascar; and the monogeneric Neotropical Derotettiginae (Marshall et al. 2018; Simon et al. 2019; Owen et al. 2022). These subfamilies house 55 tribes and over 500 genera (World Auchenorrhyncha Database, accessed 18 April 2025). Due to their species-specific songs, cicadas have been considered a model system for studying speciation and reproductive isolation (Villet 1995; Marshall and Cooley 2000; Marshall et al. 2011; Hertach et al. 2016). Furthermore, due to their long underground life cycles, short-lived adults, and low dispersal ability (Tobella et al. 2025), they have clear population genetic structure that has been analyzed to answer biogeographic and phylogeographic questions (e.g. Hill et al. 2009, 2021; Marshall et al. 2009, 2012; Seabra et al. 2009; Price et al. 2010, 2019; Hertach et al. 2016; Bator et al. 2021; Nguyen et al. 2022; Stukel et al. 2024).

The fossil record suggests that Cicadidae may have originated or diversified around the time of the K-Pg mass extinction, between the Late Mesozoic and Early Cenozoic. Although all fossils unambiguously assigned to crown Cicadidae are from the Cenozoic (Moulds 2018; Moulds et al. 2022, 2023; Boderau et al. 2025b, 2025a), a recent morphological phylogenetic study has suggested that some Mesozoic fossils traditionally assigned to Tettigarctidae are in fact fossils branching off the stem lineage of Cicadidae (Jiang et al. 2024). At the same time fossils assigned to Tettigarctidae appear as early as the Triassic period (Jouault et al. 2024), are numerous throughout the Mesozoic with much of the diversity found in the Jurassic, but decline in the Cenozoic (Shcherbakov 2009; Moulds 2018). This distribution of fossils raises the

possibility that the K-Pg extinction event was an important factor for diversification of the Cicadidae.

Previous studies have investigated multiple evolutionary radiations at different taxonomic levels within the Cicadidae. These include a recent New Zealand cicada radiation of five genera and around 60 species (Arensburger et al. 2004; Buckley et al. 2006; Buckley and Simon 2007; Marshall et al. 2008, 2011, 2012; Bator et al. 2021). Other radiations at deeper levels include tribe-level radiations, such as of the Tacuini (Hill et al. 2015) with over 200 described species distributed in North America, Europe, Asia, SE Asia, Oceania, Australia, and northern South America, the Platypleurini with over 250 described species distributed across Africa and eastern Asia (Price et al. 2019), and the Cicadettini with almost 600 described species and a global distribution except Madagascar (Marshall et al. 2016). Other cicada radiations include collections of tribes, such as a radiation of over 10 Asian cicada tribes with over 700 described species (Hill et al. 2021), or the entire family (Owen et al. 2022). While many previous studies investigating these cicada radiations at different taxonomic levels using small numbers of genes were able to resolve some relationships among taxa, many areas of the phylogeny remained unresolved, requiring phylogenomic approaches to improve our understanding of the evolutionary history of cicadas. In particular, Owen et al. (2022) dramatically improved the phylogenetic resolution of the entire family, but their study included several important but undersampled lineages that are now addressed in the present study.

The geographic origins of Cicadidae and its major lineages have not been conclusively established. Previous researchers have investigated the historical biogeography of some cicada

lineages, mainly at the tribal level. For example, Marshall et al. (2016) inferred an Australasian origin for the ultra-diverse Cicadettini. Other studies have determined an African origin for Platypleurini (Price et al. 2019), a southern Chinese origin for the tribe Polyneurini (Wang et al. 2024), and an Asian origin for a wide collection of predominantly Asian cicada tribes (Hill et al. 2021). Biogeographic inferences for deeper clades in the cicada phylogeny have been hampered by the lack of time-calibrated trees on which to base the interpretation. Simon et al. (2019) proposed a South American origin for the family Cicadidae, basing this on two factors: 1) Derotettiginae, which they found sister to all other Cicadidae, is restricted to South America, and 2) Tibicininae, which is sister to the remaining three subfamilies, is found in the Neotropics, Holarctic, and North Africa and has many early-diverging lineages restricted to South America.

In this study, we used a phylogenomic dataset for Cicadoidea with expanded taxon sampling compared to Owen et al. (2022), with 60% more taxa comprising one additional subfamily, three additional tribes, and 37 additional genera. The expanded taxa also increased the sampling for South America, Africa, and Madagascar, improving the geographic breadth and evenness for the taxon sampling. Finally, we conducted the first model-based biogeographic analysis for the entire Cicadidae family and its component subfamilies to test the hypothesis of South American origin (Simon et al. 2019).

Methods

Data and scripts for all analyses described below are deposited on Dryad: http://datadryad.org/share/LINK_NOT_FOR_PUBLICATION/DHJLmoEdp3L8QYJeiUO-UiGPCO0INIRmAoNoMxAOA-E.

Taxon sampling

We expanded the data sampling of a previous phylogenomic dataset for Cicadidae (Owen et al. 2022), with focus on subfamilies and tribes that were undersampled, especially Afrotropical and Neotropical taxa. This allowed our expanded dataset to better span the phylogenetic and biogeographical diversity of Cicadidae. Cicada specimens were primarily collected by members of the Simon lab or by specialist collaborators, with the exception of six specimens that were loaned by the Natural History Museum, London and one specimen loaned by the Muséum national d'Histoire naturelle, Paris. Cicada specimens were identified using original literature and reference collections by specialist collaborators.

AHE sequencing and data processing

Cicada specimens were sequenced by Anchored Hybrid Enrichment (AHE; Lemmon et al. 2012) at the Center for Anchored Phylogenomics (www.anchoredphylogeny.com) except for specimen PL623 (*Derotettix mendosensis* Berg), for which sequence reads were obtained from NCBI (SRA: SRS20953264). We prepared Illumina libraries for cicada specimens following Prum et al. (2015) from genomic DNA. Briefly, DNA was fragmented to 125-400 bp with a Covaris ultrasonicator, ligated to Illumina adaptors with 8 bp indexes, and the libraries pooled in groups of ~24 samples. Each library pool was enriched using the Auchenorrhyncha probe set

178 developed by Dietrich et al. (2017) and described in Simon et al. (2019). Enriched reads were
179 sequenced on an Illumina NovaSeq6000 with a PE150bp protocol. All sequence reads were
180 assembled using the data processing pipeline described in Stukel et al. (2024) with a few
181 modifications. To summarize, Illumina adaptors were trimmed from the reads with
182 Trimmomatic, the trimmed reads merged with BBMerge, and the reads from each cicada sample
183 assembled using SPAdes (Bankevich et al. 2012; Bolger et al. 2014; Bushnell et al. 2017). The
184 assemblies were queried for AHE locus sequences using BLAST, the top matching sequences for
185 each locus in the assemblies extracted using AliBaSeq, and the extracted sequences aligned using
186 MAFFT (Camacho et al. 2009; Katoh and Standley 2013; Knyshov et al. 2021). Finally, the
187 alignments were clustered using CDHit (Fu et al. 2012) to remove reduntant sequences and close
188 in-paralogs. UPhO (Ballesteros and Hormiga 2016) was used to identify more divergent paralog
189 groups, which were later split off into new locus alignments following the decision process
190 outlined by Stukel et al. (2024).

191 The modifications to the pipeline involved the six pairs of taxa that Owen et al. (2022)
192 identified as having contamination: *Zammara* cf. *erna* and *Platypleura octoguttata*, *Durangona*
193 *tigrina* and *Hamza ciliaris*, *Quintilia wealei* and *Talcopsaltria olivei*, *Tettigomyia vespiformi*s and
194 *Adusella insignifera*, *Parnisa* sp. and *Kikihia rosea*, and *Lembeja paradoxa* and *H. ciliaris*. The
195 presence of contaminated samples complicated the paralogy identification step of our pipeline.
196 During the UPhO step, the contaminated sequences were allowed as in-paralogs. This enabled
197 multiple sequences from a contaminated taxon to be included in an ortholog cluster. Each
198 contaminated locus identified by Owen et al. (2022) was examined for the presence of multiple

sequences from a taxon in a contaminated taxon pair. For many loci, the contaminant sequence was easy to identify: one taxon in the pair had one recovered sequence and the other taxon had two, one of which was nearly identical to the single sequence of the first taxon in the pair, indicating that it was the contaminant. In these cases we were able to identify and remove the contaminant sequence in each affected locus (instead of removing the taxon pair from the locus entirely; Owen et al. 2022), allowing us to increase our sequence data for these taxa. For contaminated loci that did not have recovered sequences that could be easily identified as contaminants in this way, we followed Owen et al. (2022) and removed the taxon pair from the contaminated locus.

We conducted our final alignment processing and screened the resulting locus alignments for the number of nuclear loci recovered. To summarize the final alignment processing, we used ALiBaSeq and MAFFT to re-extract alignments with 300 bp flanking regions without paralogs or contaminants from the assemblies. Misaligned or misassembled regions in these assemblies were removed using HMMCleaner (Franco et al. 2019) and the ends of the alignments were trimmed until 75% of the taxa were present with a custom script. As several of our assemblies were sequenced from museum specimens, we were concerned about unreliable sequence data from poor-quality DNA. Because taxa yielding abnormally few loci seemed the most likely to be unreliable, we removed taxa with fewer than 100 recovered loci from all alignments (taxa that were removed are not included in the specimen table). Subsequently, we screened the new alignments for contamination using the Owen et al. (2022) detection pipeline and removed putative contaminants from the affected loci.

220

221 *Assembly of mitochondrial genomes from AHE bycatch*

222 In addition to extracting nuclear AHE loci, we assembled mitochondrial genomes from
223 off-target AHE capture data in the SPAdes assemblies (Simon et al. 2019; Haji et al. 2022; Stukel
224 et al. 2024). Mitochondrial contigs in the assemblies were identified using a *Derotettix*
225 *mendosensis* reference genome (Łukasik et al. 2019), reads mapped to these contigs using BWA
226 (Li and Durbin 2009), and the mapped reads reassembled with SPAdes (Bankevich et al. 2012).
227 The reassembled contigs were aligned to the reference with MAFFT, the contigs merged into a
228 single genome sequence in Geneious, and misassembled regions removed (Kearse et al. 2012).
229 MITObim (Hahn et al. 2013) was then used to extend the contigs with unassembled sequence
230 reads, the extended genome sequences aligned in MAFFT, and the resulting alignment edited
231 manually in Geneious to remove misassembled regions and adjust reading frames to match the
232 reference genome. Mitochondrial genomes from the *Kikihia*, *Maoricicada*, *Rhodopsalta*, and
233 *Derotettix* were obtainedfrom NCBI (GenBank accessions MG737807, OR413956, OR459873,
234 OR459874, OR459906, OR413963) rather than re-extracting them from our assemblies.

235

236 *Preliminary phylogenomic analysis*

237 Preliminary phylogenetic analyses were conducted to identify support for various cicada
238 clades for subsequent analyses. A concatenated maximum-likelihood (ML) phylogeny for the
239 nuclear AHE data was constructed with IQ-Tree v. 2.2.2 (Minh et al. 2020). The data were
240 partitioned by gene and he best model and partitioning scheme selected using the -m

TESTMERGE setting in ModelFinder, which uses the PartitionFinder "greedy" algorithm (Chernomor et al. 2016; Kalyaanamoorthy et al. 2017; Lanfear et al. 2017). To save computation time, the relaxed hierarchical clustering algorithm was used to examine only the top 10% partition merging schemes (Lanfear et al. 2014). AHE loci were not partitioned by codon due to their typically short length. Branch support statistics were estimated using 1000 ultrafast bootstrap (UFB) and 1000 Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT) replicates (Guindon et al. 2010; Hoang et al. 2018). Species trees were estimated from the nuclear AHE data using ASTRAL-III and SVDQuartets (Chifman and Kubatko 2014; Zhang et al. 2018). For ASTRAL, we individual gene trees were first estimated in IQ-Tree with model selection using ModelFinder and 1000 SH-aLRT replicates, and then all gene tree branches with 0% SH-aLRT support were removed with *newick-utilities* to reduce the effect of gene tree estimation error before the gene trees were fed into ASTRAL-III. While majority-rule consensus trees have been argued to best represent uncertainty for a set of trees (Berry and Gascuel 1996; Holder et al. 2008), ASTRAL has been shown to have better performance with ML trees rather than with bootstrap consensus trees (Sayyari and Mirarab 2016). Based on the recommendations of Simmons and Gatesy (2021), we chose this SH-aLRT threshold rather than a bootstrap support threshold. The SVDQuartets analysis was run in PAUP* (Wilgenbusch and Swofford 2003) using the concatenated alignment and with 100 multilocus bootstrap replicates to estimate branch supports. Although methods for estimating branch lengths have recently been developed for SVDQuartets (Peng et al. 2022), we estimated the topology only for our SVDQuartets phylogeny.

A ML mitochondrial phylogeny was constructed from the recovered mitochondrial genomes using IQ-Tree. The mitochondrial genomes were partitioned by gene and codon, and the same ModelFinder settings as the concatenated nuclear analysis above were used to merge the partitioning schemes. Branch support statistics were again estimated using 1000 UFB and SH-aLRT replicates.

Bayesian species tree inference and divergence time estimation

Bayesian species tree inference and divergence time estimation were conducted in BEAST v. 2.7.5 (Bouckaert et al. 2019). The species tree topology and divergence times were jointly estimated using clades that were well-supported in all three nuclear analyses (IQ-Tree, ASTRAL, or SVDQuartets) as monophyly constraints for fossil placements. Divergence time estimation was performed in an MSC framework to avoid inflating terminal branch lengths when conducting inference with concatenation (Carruthers et al. 2022). Due to the size of the dataset, a divide and conquer approach was used, in which a reduced taxon set was used to estimate a backbone tree, and taxon sets representing subclades from that backbone were used to estimate subtrees that were grafted onto the backbone tree. To ensure that the subtrees had dates compatible with the backbone tree, we used approximations of the divergence time posterior distributions in the backbone tree as prior distributions for the subtrees. This necessitated using different subsets of loci between the backbone tree and the subtrees to avoid “double-dipping” (i.e. using the same data for determining both prior and likelihood). The backbone taxon set consisted of 20 taxa spanning the deepest splits among and within the five Cicadidae subfamilies

(Table 1). The subtree taxon sets consisted of the taxa belonging to Tibicininae, Cicadettinae, Tettigomyiinae, Tacuini, and Cicadinae without Tacuini. For the backbone and subtree taxon sets, species tree inference was performed using StarBeast3 v. 1.1.8 with a relaxed clock, using a fossilized birth-death (FBD) tree prior from Sampled-Ancestors v. 2.1.1 for time calibration and substitution model averaging from bModelTest v. 1.3.3 (Gavryushkina et al. 2014; Bouckaert and Drummond 2017; Douglas et al. 2021, 2022). Additional details for the different taxon set analyses are explained below.

For the backbone tree, nuclear AHE loci with all 20 taxa were identified and from them the 50 loci with the highest number of parsimony-informative sites as calculated by PhyKIT (Steenwyk et al. 2021) were selected. Because of the selective taxon sampling and deep splits in the backbone taxon set, we were concerned that the assumption of constant rates through time would lead to underestimation of diversification, turnover, and fossil sampling rates and corresponding overestimation of divergence times for deep clades, a phenomenon known as "deep root attraction" (Ronquist et al. 2016; Matschiner 2019; Luo et al. 2023; Zhang et al. 2023). To address this, hree versions of the backbone analysis were compared: an analysis with the standard FBD model with constant rates; a Skyline-FBD model with diversified sampling assumptions and two rate transition points at the Jurassic-Cretaceous and Cretaceous-Paleogene boundaries; and a Skyline-FBD model with diversified sampling but no rate transition points (Stadler et al. 2013; Stadler and Smrckova 2016; Zhang et al. 2016). We included 15 fossil tips for all three versions of the backbone analysis (Table 1, first 15 lines). Eight fossil tips were stem-group taxa identified in a recent morphological phylogenetic analysis (Jiang et al. 2024),

consisting of one stem cicadoid, one stem tettigarctid, and six stem cicadids. Seven further tips were fossils of two Cicadinae, two Cicadettinae, and three Tibicininae, all of which represent the oldest unambiguous crown Cicadidae fossils older than our diversified sampling cutoff of 25 Ma (Moulds 2018). To avoid biases caused by ignoring stratigraphic age uncertainty, we specified a log-normal prior for each fossil tip such that the published fossil age range spanned the central 95% of the prior distribution (Barido-Sottani et al. 2019). Based on the estimate of 220 Ma for the divergence between Cicadoidea and Cercopoidea in Johnson et al. (2018) we used a log-normal prior for the origin with a mean in real space of 54 and an offset of 166 (the age of the oldest fossil in our dataset). We set ρ (extant sampling proportion) to 0.004 based on an estimate of approximately 4,000 extant species of cicadas. We chose uniform priors for d (net diversification rate) and r (turnover) (Table 2). Due to the few number of fossils, the fossil sampling proportion s was expected to be small, so the number of sampled fossils was divided by 4,000 extant species of cicadas to get a rough approximation for the sampling proportion, and a beta distribution with a mean of this value was chosen to be the prior for s.

Table 1: List of fossil calibrations

Because phylogenetic constraints are nested, only the narrowest constraint for each analysis is listed for each fossil. All fossils are necessarily included in constraints for deeper clades that contain the listed constraint.

Fossil	**Age (Ma)**		**Phylogenetic constraint**	**Notes**	**Reference**
	Min	**Max**			
Macrotettigarcta	157	166	Backbone: stem		Jiang et al. 2024

Fossil	**Age (Ma)** **Min**	**Max**	**Phylogenetic constraint**	**Notes**	**Reference**
obesa			Cicadidae		
Sanmai mengi	157	166	Backbone: stem Cicadidae		Jiang et al. 2024
Shuraboprosbole daohugouensis	157	166	Backbone: stem Cicadidae		Jiang et al. 2024
Tianyuprosbole zhengi	157	166	Backbone: stem Tettigarctidae		Jiang et al. 2024
Cretotettigarcta problematica	98.2	99.4	Backbone: stem Cicadidae; in a clade with *Pranwanna*, *Vetuprosbole*, and crown Cicadidae	Placement informed by well-supported published morphological analysis	Jiang et al. 2024
Eunotalia emeryi	98.2	99.4	Backbone: stem Cicadoidea		Jiang et al. 2024
Pranwanna xiai	98.2	99.4	Backbone: stem Cicadidae; sister to crown Cicadidae	Placement informed by well-supported published morphological analysis	Jiang et al. 2024
Vetuprosbole parallelica	98.2	99.4	Backbone: stem Cicadidae; in a clade with *Cretotettigarcta*, *Pranwanna*, and crown Cicadidae	Placement informed by well-supported published morphological analysis	Jiang et al. 2024
Davispia bearcreekensis	56	59.2	Backbone and Tibicininae: Tibicinini		Moulds 2018
Lyristes sp.	30.2	30.5	Backbone: Tacuini; in a clade with *Hadoa* and *Megatibicen*. Tacuini: in a clade with *Auritibicen* and *Lyristes*	Placement in backbone tree is the branch of Tacuini where the genus belongs. Placement in Tacuini tree based on old age of fossil vs. low divergence between extant *Auritibicen* and *Lyristes*	Moulds 2018
Hadoa grandiosa	27.8	33.9	Backbone: Tacuini; in a clade with *Lyristes* and *Megatibicen.*	Placement in backbone tree is the branch of Tacuini where the genus belongs. Placement in Tacuini tree is based on old	Moulds 2018

Fossil	**Age (Ma)** **Min**	**Max**	**Phylogenetic constraint**	**Notes**	**Reference**
			Tacuini: sister to extant *Cacama* and *Hadoa*	age of fossil vs. low divergence between extant *Cacama* and *Hadoa*.	
Lithocicada perita	27.8	33.9	Backbone and Tibicininae: Tibicinini		Moulds 2018
Paracicadetta oligocenica	27.8	33.9	Backbone: in a clade with *Cicadatra* and *Magicicada.* Cicadettinae: Pagiphorini	Placement on the backbone tree is the branch of Cicadettinae where the genus belongs.	Moulds 2018
Platypedia primigenia	27.8	33.9	Backbone and Tibicininae: Platypediini	Placement in the Tibicininae tree is based on old age of fossil vs. low divergence between extant *Neoplatypedia* and *Platypedia*	Moulds 2018
Cicadatra? serresi	26.0	27.8	Backbone and Cicadettinae: Cicadatrini	Placement chosen despite tentative assignment of fossil to *Cicadatra*	Moulds 2018
Tibicina sakalai	17.8	17.9	Tibicininae: in a clade with *Okanagana* and *Tibicina*	Placement is based on old age of fossil vs. low divergence between extant *Okanagana* and *Tibicina*	Moulds 2018
Meimuna protopalifera	16	23	Cicadinae (without Tacuini): Dundubiini		Moulds 2018
Tymocicada gorbunovi	16	23	Tacuini: Tacuini		Moulds 2018
Paleopsalta ungeri	16	20.4	Cicadettinae: Cicadettini		Moulds 2018
Minyscapheus dominicanus	15	20	Cicadettinae: Taphurini		Moulds 2018
Cryptotympana incasa	11.6	16	Tacuini: *Cryptotympana*		Moulds 2018
Tanyocicada lapidescens	11.6	16	Cicadinae (without Tacuini): in a clade	Placement is based on the close relationship of	Moulds 2020

Fossil	**Age (Ma)** **Min**	**Max**	**Phylogenetic constraint**	**Notes**	**Reference**
			with extant Gaeanini and extinct Leptopsaltriini	Leptopsaltriini (the tribe to which the fossil is assigned) and Gaeanini found in Hill & Marshall et al. (2021)	
Camuracicada aichhorni	11.6	13.8	Tacuini: Tacuini		Moulds 2018
Burbungoides gulgongensis	11	16	Cicadinae (without Tacuini): Burbungini		Moulds et al. 2022
Laopsaltria ferruginosa	11	16	Cicadettinae: Cicadettini		Moulds et al. 2022
Tithopsaltria titan	11	16	Cicadinae (without Tacuini): in a clade with extant Cyclochilini, Macrotristriini, Psaltodini, and Thophini	Placement is based on the close relationship of Arenopsaltriini with these tribes in Marshall et al. (2018)	Moulds et al. 2022
Lyristes renei	8	8.5	Tacuini: *Lyristes*		Moulds 2018
Miocenoprasia grasseti	8	8.5	Cicadettinae: Lamotialnini		Moulds 2018
Tibicina gigantea	8	8.5	Tibicininae: *Tibicina*		Moulds 2018
Auritibicen sp. aff. *japonicus*	5.3	11.6	Tacuini: *Auritibicen*		Moulds 2018
Graptopsaltria inaba	5.3	11.6	Cicadinae (without Tacuini): *Graptopsaltria*		Moulds 2018
Yezoterpnosia sp. aff. *vacua*	5.3	11.6	Cicadinae (without Tacuini): in a clade with extant Gaeanini and extinct Leptopsaltriini	Placement is based on the close relationship of Leptopsaltriini (the tribe to which the fossil is assigned) and Gaeanini found in Hill & Marshall et al. (2021)	Moulds 2018
Lyristes? emathion	5.3	7.2	Tacuini: *Lyristes*		Moulds 2018

Fossil	**Age (Ma)**		**Phylogenetic constraint**	**Notes**	**Reference**
	Min	**Max**			
Cicada sp. aff. *lodosi*	2.6	3.6	Cicadinae (without Tacuini): *Cicada*		Moulds et al. 2023
Cicada sp. aff. *orni*	2.6	3.6	Cicadinae (without Tacuini): *Cicada*		Moulds 2018
Cicada tithonus	2.6	3.6	Cicadinae (without Tacuini): *Cicada*		Moulds et al. 2023
Meimuna sp.	2.6	3.6	Cicadinae (without Tacuini): Dundubiini		Moulds 2018
Tibicina boulardi	2.6	3.6	Tibicininae: *Tibicina*		Moulds et al. 2023
Tibicina sp. aff. *haematodes*	2.6	3.6	Tibicininae: *Tibicina*		Moulds 2018
Tibicina lata	2.6	3.6	Tibicininae: *Tibicina*		Moulds et al. 2023
Tanna? sp.	1.5	1.7	Cicadinae (without Tacuini): in a clade with extant Gaeanini and extinct Leptopsaltriini	Placement is based on the close relationship of Leptopsaltriini (the tribe to which the fossil is assigned) and Gaeanini found in Hill & Marshall et al. (2021)	Moulds 2018
Graptopsaltria sp. aff. *nigrofuscata*	0.5	1.1	Cicadinae (without Tacuini): *Graptopsaltria*		Moulds 2018
Auritibicen bihamatus	0.033	0.038	Tacuini: *Auritibicen*		Moulds 2018
Yezoterpnosia nigricosta	0.033	0.038	Cicadinae (without Tacuini): in a clade with extant Gaeanini and extinct Leptopsaltriini	Placement is based on the close relationship of Leptopsaltriini (the tribe to which the fossil is assigned) and Gaeanini found in Hill & Marshall et al. (2021)	Moulds 2018

Table 2: Priors for FBD model parameters

All lognormal distributions are reported with the mean in linear space instead of log space.

Analysis	**Origin**	**Root**	***ρ* (extant sampling prop.)**	***d* (diversification rate)**	***s* (sampling prop.)**	***r* (turnover)**
Backbone (all three analyses)	LogNorm(54, 1), offset=166	NA	0.005	Uniform(inf)	Beta(1, 250)	Uniform(1, 0)
Cicadettinae	NA	LogNorm(34.1, 0.25), offset=33.9	0.04	Uniform(inf)	Beta(1, 250)	Uniform(1, 0)
Cicadinae (no Tacuini)	NA	LogNorm(40, 0.198), offset=23	0.03	Uniform(inf)	Beta(1, 120)	Uniform(1, 0)
Tacuini	LogNorm(45, 0.17), offset=28	NA	0.04	Uniform(inf)	Beta(1, 32)	Uniform(1, 0)
Tettigomyiinae	LogNorm(90, 0.1)	NA	0.1	Uniform(inf)	0	Uniform(1, 0)
Tibicininae	NA	LogNorm(9, 0.5), offset=56	0.06	Uniform(inf)	Beta(1, 35)	Uniform(1, 0)

Table 3: Differences in divergence times across different StarBeast3 analyses

Common ancestor (CA) heights and highest posterior densities (HPD) for assorted deep clades obtained for each StarBeast3 analysis. The posteriors for the backbone analysis with diversified sampling only (second line of table) were approximated as root or origin prior distributions for the subtree analyses (see Table 2 above).

Analysis	**Cicadoidea CA height (95% HPD)**	**Cicadidae CA height (95% HPD)**	**Cicadettinae CA height (95% HPD)**	**Cicadinae (no Tacuini) CA height (95% HPD)**	**Tacuini CA height (95% HPD)**	**Tettigomyiinae CA height (95% HPD)**	**Tibicininae CA height (95% HPD)**
Backbone, constant FBD	197.5 (166.8-247.2) Ma	169.1 (130.8-221.1) Ma	92.9 (65.2-125.8) Ma	83.0 (59.0-111.9) Ma	45.4 (32.7-61.0) Ma	109.3 (68.1-154.9) Ma	71.9 (57.4-92.3) Ma
Backbone, Skyline-FBD diversified sampling	177.7 (164.8-192.8) Ma	135.2 (105.3-162.5) Ma	67.7 (49.3-86.2) Ma	63.2 (40.1-75.4) Ma	38.2 (31.4-46.8) Ma	80.2 (45.6, 109.8) Ma	65.0 (57.0-75.3) Ma
Backbone, Skyline-FBD diversified sampling, 2 transition points	170.0 (161.3-180.2) Ma	145.2 (123.3-162.3) Ma	61.2 (45.4-80.4) Ma	57.7 (47.2-66.0) Ma	37.8 (25.4-43.7) Ma	67.4 (27.6-104.8) Ma	62.0 (57.9-66.0) Ma
Subtree, constant FBD	—	—	72.6 (57.8-88.7) Ma	61.5 (49.6-73.4) Ma	40.6 (32.0-51.1) Ma	65.4 (45.4-85.3) Ma	65.7 (59.0-74.0) Ma

The StarBeast3 analyses were set up similarly for the individual subtree taxon sets. We The 50 loci used in the backbone tree were excluded from consideration and the locus selection process repeated for each taxon subset. The nuclear AHE loci containing all taxa in each subset

were identified. If there were fewer than 50 loci that contained all taxa for a set, the taxon occupancy threshold was relaxed down to all loci with at least *N*-2 taxa, where *N* is the number of taxa in the subtree taxon set. The top 50 loci with the largest number of parsimony-informative sites were selected from thes high-occupancy loci. Because the subtree taxon sets were not as thinly sampled as the backbone taxon set, all analyses were performed with the standard, constant-rates FBD model. A combined total of 36 fossil tips was used, split among the Cicadettinae, Tibicininae, Tacuini, and Cicadinae without Tacuini datasets; no fossil Tettigomyiinae were included as no fossils from this subfamily have been identified (Table 1). As with the backbone tree, we accounted for stratigraphic age uncertainty in all of the fossils using log-normal priors such that the published fossil age range spanned the central 95% of the prior distribution. We chose the backbone tree estimated using the Skyline FBD model with no transition points and diversified sampling as the source for calibrating the subtrees. The root age or origin age priors for all analyses were set to approximate the posterior ages for the corresponding clades in this backbone tree, with the appropriate offsets based on the oldest fossils (Table 2). Root ages were approximated for clades which we were confident had their deepest split captured in the backbone tree. For clades that we were unsure had their deepest split captured in the backbone tree, the age distribution of the next-oldest split in the backbone tree was approximated as an origin age prior. A lognormal approximation of the backbone Species Tree Relaxed Clock rate posterior was used as the prior distribution for this parameter in each of the subtrees. Values for ρ (extant sampling proportion) were set based on the total number of described species in each subtree clade (Catalog of Life, accessed 13 May 2024) rescaled to an

estimate of 4,000 total extant species for the whole family (Table 2). Uniform priors for *d* (net diversification rate) and *r* (turnover) were used as in the backbone analysis (Table 2). The fossil sampling proportion *s* was again approximated by dividing the number of sampled fossils in a dataset by the estimated extant species in the clade, and a beta distribution with a mean of this value was chosen to be the prior (Table 2). Lacking Tettigomyiinae fossils, analysis for this subfamily was attempted using a Yule model instead of an FBD model, but preliminary results were unsatisfactory, likely due to our sampling being insufficiently dense. Instead, an FBD model was used with *s* set to zero, which is equivalent to a birth-death model.

All analyses, whether backbone or subtree, were performed with two identical runs each for up to 400 million generations. Because the Tettigomyiinae and Cicadinae without Tacuini datasets became stuck in local optima, three heated chains and one cold chain were used with the CoupledMCMC v. 1.3.1 package with the default swapping value for those datasets to improve convergence (Müller and Bouckaert 2020); all other analyses lacked heated chains. Convergence was determined by comparing the two independent runs for each analysis in Tracer v. 1.7.2 and the clade comparison tool in DensiTree (Bouckaert and Heled 2014). The first 10% of tree samples were removed as burn-in from each run and combined the species tree files in LogCombiner v. 2.7.5. Fossil tips were removed from each combined species tree and MCC consensus trees were generated with common ancestor node heights in TreeAnnotator (Heled and Bouckaert 2013). The subtree consensus trees and their associated metadata were grafted onto the backbone tree in R. Because we considered the root age estimates and highest posterior density (HPD) distributions from the different subtrees to be more refined than those from the

backbone tree, subtrees were attached to the backbone at their inferred root node height and the corresponding node on the backbone tree was removed. However, the posterior probabilities from the backbone tree for these nodes were kept because we believed they were closer to the true level of uncertainty for these relationships (the posterior probabilities for these nodes in the subtrees were all 1 due to being their roots).

Ancestral area reconstruction

We removed the Tettigarctidae outgroup and a handful of tips that appeared to be the same species as other tips from the combined BEAST dated phylogeny, resulting in a pruned phylogeny of 156 species tips. We obtained species range data for all species in this pruned tree using the World Auchenorrhyncha Database and iNaturalist; cases where the identity of the geographic area of a particular species was ambiguous were resolved based on the primary literature. For specimens in our tree that were not identified to species (~25% of species), we used the country of collection as the species range. While this certainly restricts their ranges compared to the fully-identified species, this was mitigated by the large geographic areas assigned to species. Ranges were to species according to the following area scheme, loosely based on the biogeographic realms of Udvardy (1975): North America (N; includes Mexico north of the Yucatan peninsula), Central America and Caribbean islands (C; includes the Yucatan Peninsula to Panama), South America (S), Africa (F; includes the Arabian peninsula), Madagascar and Indian Ocean islands (M; includes Mauritius, Rodrigues, Réunion, and Seychelles), West Eurasia (W; Eurasia west of the Ural Mountains and the Caspian Sea), East

Eurasia (E; Eurasia east of the Ural Mountains and the Caspian Sea, north of the Indian subcontinent and southern China), India (I; Indian subcontinent including Pakistan, India, Nepal, Bangladesh, and Sri Lanka), Southeast Asia (D; Indomalayan biogeographic realm west of Wallace's line excepting the Indian subcontinent), and Australasia (Z; Australia, New Guinea, Wallacea, New Zealand, New Caledonia, and nearby Polynesian islands).

To account for the effect of distance on area reconstruction, a distance matrix was built for biogeographic inference analysis. The closest present-day geodesic distance between each pair of areas was measured, whether over water or land. For areas that are directly adjacent on land, a minimum distance of 1 km was assigned since zeros are reserved for diagonals on the distance matrix. To account for distances between areas changing over geologic time, five past time points for additional area measurements were selected: 25 Ma, 50 Ma, 75 Ma, 100 Ma, and 150 Ma. We skipped 125 Ma as a time point since very few lineages were present at that time point (see results). All distances were measured in GPlates v. 2.5 (Müller et al. 2018), with the provided EarthByte Global Rotation Model (Müller et al. 2019; Torsvik et al. 2019; Young et al. 2019; Cao et al. 2022; Zahirovic et al. 2022) and Paleogeography feature set (Cao et al. 2017) used to assist with measuring paleodistances at the chosen time points.

Ancestral area reconstruction analyses were performed in BioGeoBEARS v. 1.1.1 under the DEC, DIVALIKE, and BAYAREALIKE models (Matzke 2013, 2018). A version of each model with a jump-dispersal parameter (+*j*) was also included, resulting in a total of six base models (Matzke 2014). The use of the jump-dispersal parameter in model comparison of DEC-type biogeography models has been critiqued as being statistically invalid (Ree and Sanmartín

2018), but the claims of this critique have been shown to be unfounded: the critique is based on incorrectly calculated likelihoods, the DEC model without a jump parameter is statistically inadequate on datasets which have many species restricted to single areas, and statistical comparison between DEC and DEC+*j* models is equivalent to comparison of their corresponding ClaSSE (Goldberg and Igić 2012) submodels (Matzke 2022). Three analyses for each base model were conducted to test for distance-dependent dispersal (+*x* variants, where dispersal probabilities are multiplied by distance raised to *x*): one incorporating dispersal independent of distance, one incorporating dispersal dependent on fixed modern distances (+*x*), and one incorporating dispersal dependent on paleodistances at the above-mentioned stratified time points (+*x*; time-stratified) (Van Dam and Matzke 2016).To reduce the size of the state space for computational tractability, the maximum number of areas was set to three in all analyses because the maximum number of areas inhabited by any species in the analysis was three (*Huechys sanguinea*, *Quesada gigas*). Model fit was optimized using generalized simulated annealing, and the fit of all models were compared using standard and sample-size corrected Akaike information criterion (AIC, AICc) and the Bayesian information criterion (BIC) (Akaike 1973; Schwarz 1978; Hurvich and Tsai 1989).

Results

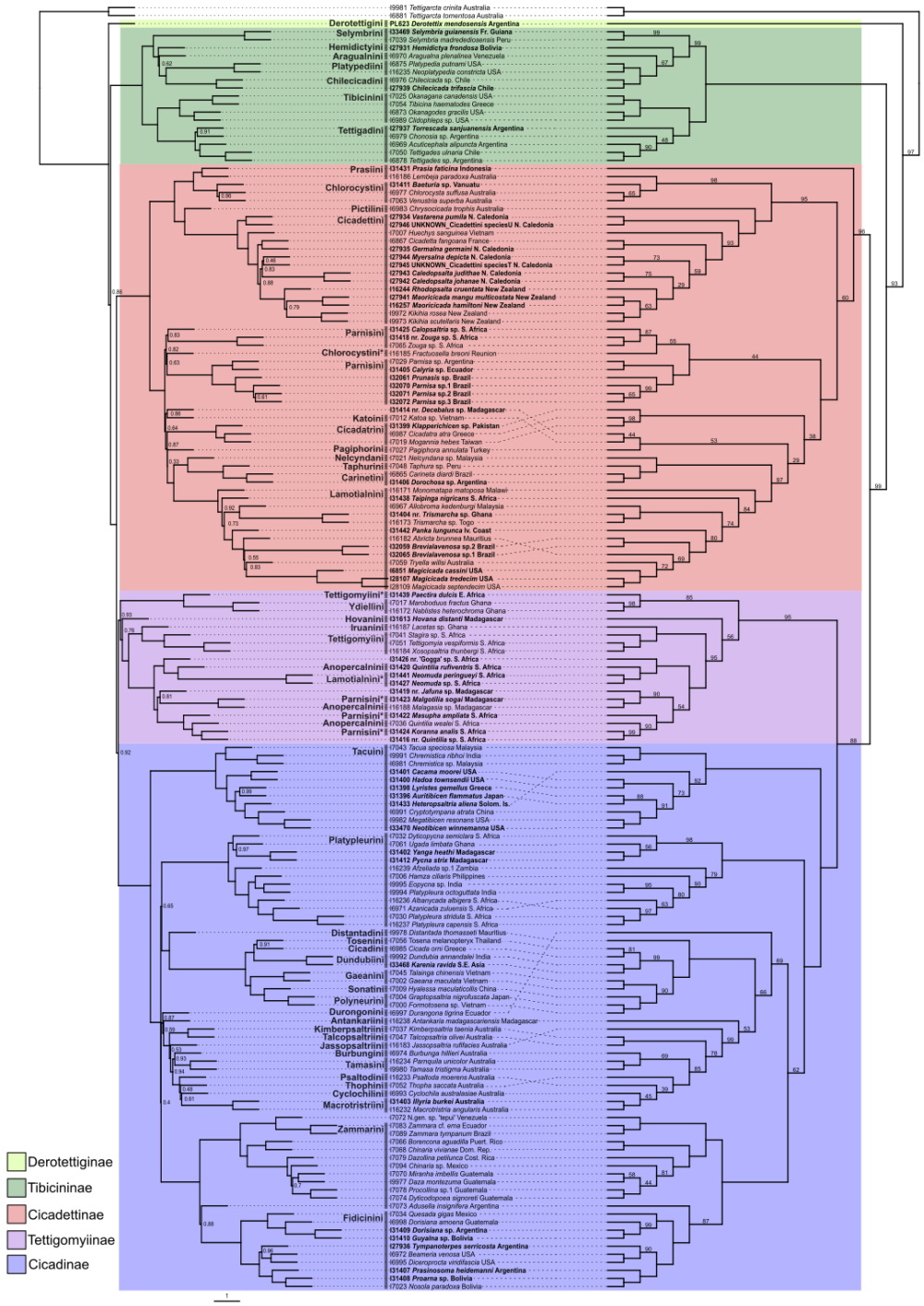
AHE sequencing and processing

Our final dataset recovered 160 taxa and 490 nuclear AHE loci (Table S1). The taxon sampling of Cicadidae from Owen et al. (2022) was expanded from 100 ingroup taxa to 158

443 ingroup taxa while retaining the two Tettigarctidae outgroup taxa (Table S1). Our 158 ingroup
444 taxa comprise 98 taxa carried over from Owen et al. (2022) and 60 additional taxa (*Magicicada*
445 *septendecula* and *Lamotialna condamini* specimens from the previous study were removed as we
446 had difficulty processing their AHE sequence reads). The final taxon sample included all five
447 subfamilies of Cicadidae, 47 of 55 tribes, and 128 of 516 ingroup genera (World
448 Auchennoryncha Database, accessed 17 November 2024). The 60 additional species added three
449 tribes and 37 genera to the study in Owen et al. (2022). After removing taxa from loci that were
450 previously identified as contaminated (Owen et al. 2022), we identified two additional pairs of
451 taxa as contaminated. These pairs were *Dorisiana* sp. and *Heteropsaltria aliena* (64 loci) and
452 *Prasinosoma* sp. and *Prasia faticina* (94 loci). Interestingly, the Lemmon Lab AHE sample IDs
453 for each pair are the same numerical distance apart (24 IDs), offset by two (I31409 & I31433 vs
454 I31407 & I31431). The number of loci recovered for each taxon after removing contamination
455 ranged from 94 to 463, with a mean of 396. The taxon occupancy across the nuclear loci ranged
456 from 74 to 159 taxa, with a mean of 129. The length of the nuclear loci ranged from 153 to 1,381
457 bp, with a mean of 326 bp. The total nuclear alignment length was 159,796 bp. We recovered 114
458 mitochondrial genomes from AHE bycatch, which when including the GenBank sequences
459 brought our entire mitochondrial alignment to 120 taxa. The total alignment length for the
460 mitochondrial genomes was 14,620.
461
462 *Phylogenetic analyses*

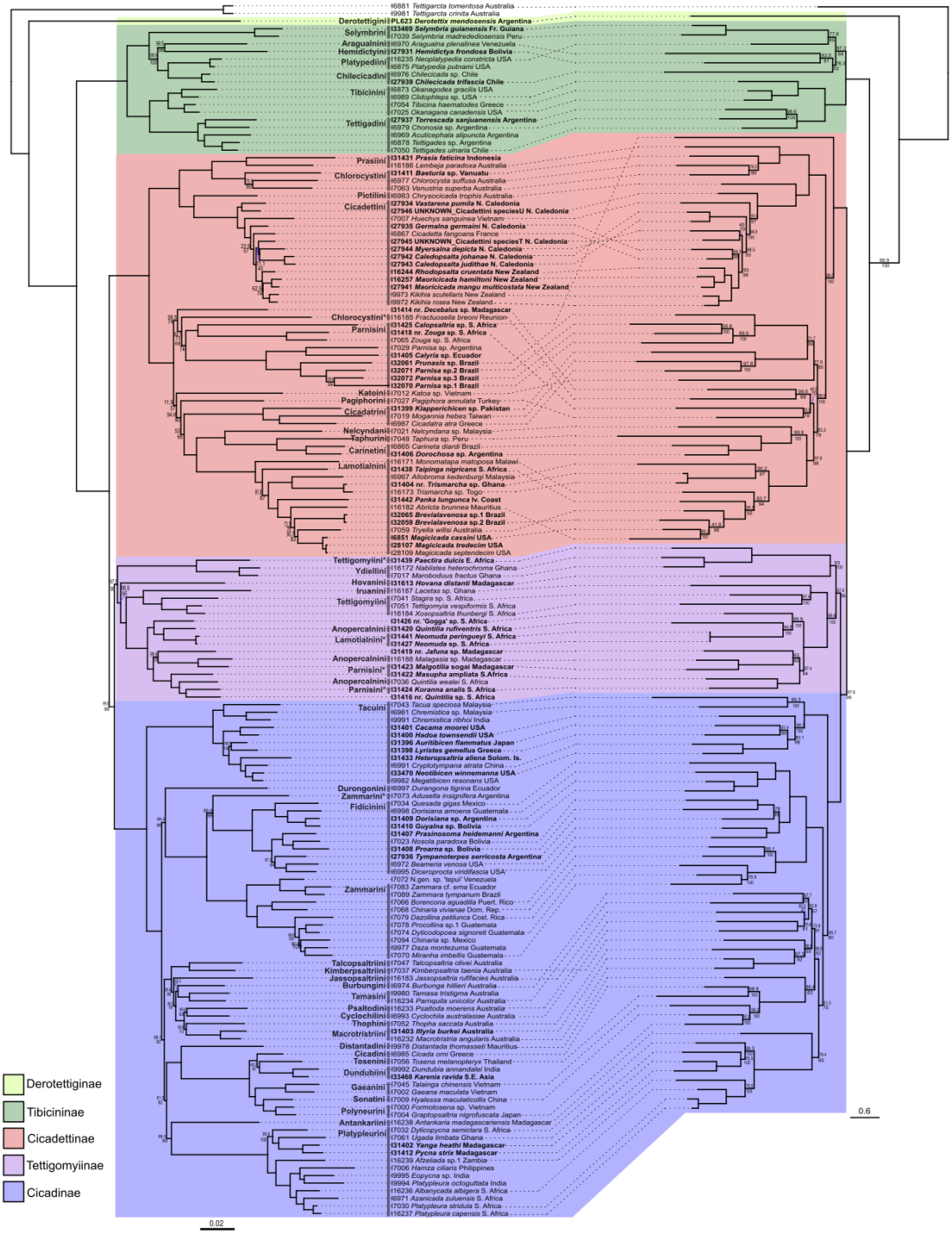
The ASTRAL, SVDQuartets, and concatenated ML nuclear phylogenies, and the ML mitochondrial genome phylogeny, were all largely congruent (Fig. 1-2). All four of these analyses recovered the same topological arrangement of cicada subfamilies. Other uncontentious relationships included deep splits within the subfamilies Cicadettinae and Cicadinae: Cicadettine was divided into two large clades and Tacuini was separated from the rest of Cicadinae. Most disagreements in topology among analyses were relationships that were strongly supported in only some or no analyses. These included the relationships of the deep radiation of Australian Cicadinae tribes, the position of *Antankaria*, and the position of *Durangona*. Although IQ-Tree ultrafast bootstraps (UFB), SH-aLRT scores, ASTRAL local posterior probabilities, and regular bootstraps are different measures of branch support that are difficult to compare directly, the SVDQuartets tree displayed much less overall branch support than the ASTRAL or nuclear concatenated ML trees.

Figure 1: Nuclear ASTRAL and SVDQuartets phylogenies



Nuclear ASTRAL and SVDQuartets phylogenies plotted face-to-face. Taxa in bold are newly added relative to Owen et al. (2022). Clade colors indicate subfamily (color key in lower left). Current tribal classifications are labeled; taxa requiring reclassification marked with asterisks. Left: ASTRAL phylogeny constructed from 490 nuclear gene trees. Internal branch lengths are in coalescent units (scale on bottom left). Terminal branch lengths are arbitrary. Nodal support statistics with ASTRAL local posterior probability below 1 are labeled. Right: SVDQuartets cladogram constructed from concatenated alignment of 490 nuclear loci. Branch lengths are arbitrary. Branch support statistics with bootstrap proportion below 100% are labeled.

Figure 2: Concatenated nuclear and mitochondrial genome IQ-Tree phylogenies



Concatenated nuclear and mitochondrial genome IQ-Tree maximum-likelihood (ML) phylogenies plotted face-to-face. Taxa in bold are newly added relative to Owen et al. (2022). Clade colors indicate subfamily (color key in lower left). Current tribal classifications are labeled; taxa requiring reclassification marked with asterisks. Branch support statistics on both trees are SH-aLRT above, Ultrafast bootstrap support below, with only branches under 100% for either support value labeled. Left: concatenated nuclear IQ-Tree ML phylogeny constructed from 490 nuclear loci. Branch lengths are in substitutions/site (scale on bottom left). Right: IQ-Tree ML phylogeny constructed from mitochondrial genomes obtained from AHE bycatch or NCBI. Branch lengths are in substitutions/site (scale on bottom right). Note that for some taxa, mitochondrial genomes were not recovered.

Bayesian species tree inference and divergence time estimation

The three versions of the backbone analysis (standard FBD; Skyline-FBD with diversified sampling and two transition points; Skyline-FBD with diversified sampling only) yielded identical summary tree topologies, but slightly different divergence time estimates (Table 3). The standard FBD analysis yielded slightly older estimated ages for the root and deeper clades, but considerably younger estimated ages for shallower clades. The two Skyline-FBD analyses estimated similar ages for most clades, but differed slightly for clades that were close to the Jurassic-Cretaceous and Cretaceous-Paleogene transition points. A closer examination of the analysis with transition points showed unusually-shaped posteriors for clade heights very close to the transition points: posterior density was concentrated on one side of the transition and diffused

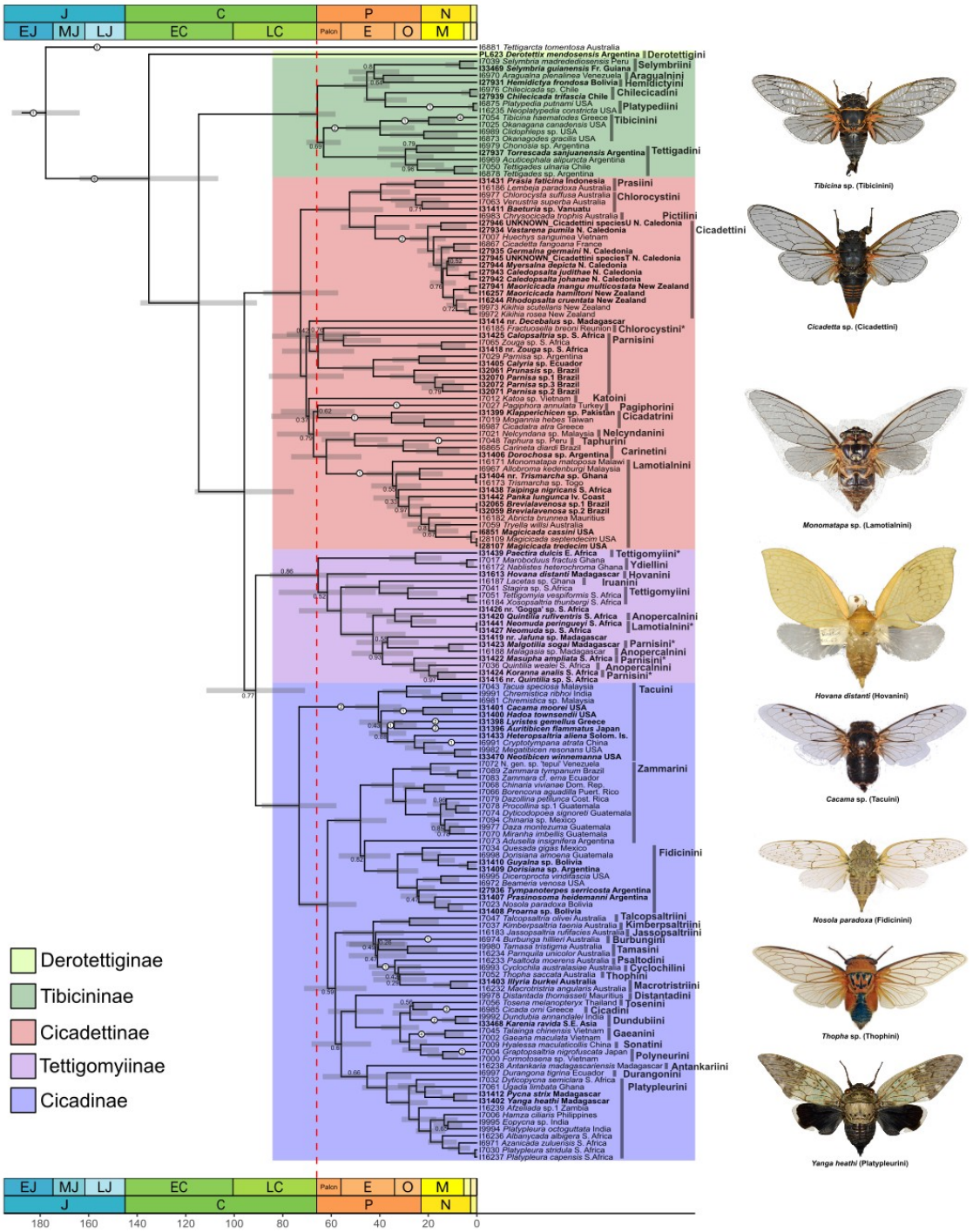
in a long tail on the other side. We believe this is an artifact from the diversification, turnover, and sampling rates changing as they cross those transition points. We therefore selected the clade heights from the Skyline-FBD with diversified sampling only analysis to serve as root or origin age priors for the subtree analysis grafted the subtrees onto it as the backbone phylogeny.

Excluding the branches that were constrained for fossil placement, the combined StarBeast3 tree largely agreed with the ML, ASTRAL, and SVDQuartets phylogenies described above (Fig. 3). Areas that had lower support in the StarBeast3 tree than the previous trees included the sister relationship of Tettigomyiinae + Cicadinae and the monophyly of Tettigomyiinae. However, the posterior probabilities in the StarBeast3 tree are not readily comparable to ML branch support statistics like bootstrap percentages or SH-aLRT scores because, as discussed earlier, branches in the StarBeast backbone or subtree phylogenies were inferred using information from only 50 nuclear genes instead of the full 490 nuclear genes for computational tractability.

The divergence times for our combined StarBeast3 tree showed a crown age for Cicadoidea in the early Jurassic (170 Ma, HPD 164.8-192.8 Ma), a crown age for Cicadidae in the early Cretaceous (135.2 Ma, HPD 105.3-162.5 Ma), and crown ages for the four non-monotypic Cicadidae subfamilies at the end of the Cretaceous or at the very start of the Paleocene (Tibicininae: 65.7 Ma, HPD 59-74 Ma; Cicadettinae: 72.6, HPD 57.8-88.7 Ma; Tettigomyiinae: 65.4 Ma, HPD 45.4-85.3 Ma; Cicadinae: 65.4 Ma, HPD 45.4-85.3 Ma). The subfamilies Cicadinae and Cicadettinae, which had estimated crown ages in the Cretaceous, had major radiations estimated to originate at or very close to the K-Pg transition. The crown ages for

most of the subtree analyses were consistent with corresponding ages in the backbone tree. As the priors for the crown ages in the subtrees were based on the posterior distributions in the backbone tree, this indicates that the data in these subtrees did not overrule the prior. The only exception to this was the Tettigomyiinae subtree, whose crown age posterior covered only the more recent half of its crown age prior (Table 3).

537 **Figure 3: Composite time-calibrated StarBeast3 phylogeny**



538

Time calibrated StarBeast3 phylogeny constructed from backbone and subtrees. Taxa in bold are newly added relative to Owen et al. (2022). Clade colors indicate subfamily (color key in lower left). Branches with posterior probabilities below 1 are labeled. Current tribal classifications are labeled; taxa requiring reclassification marked with asterisks. Node heights are common ancestor heights, with 95% HPD intervals displayed as node bars. White circles with numbers indicate branches constrained with fossils and the number of fossils assigned to the branch. Period and epoch data from International Commission on Stratigraphy (v2023/09) plotted at the top and bottom with standard abbreviations (Quaternary, Pliocene, and Pleistocene not labeled). Red dotted line indicates K-Pg transition. Cicada photos highlight exemplar genera and species from major clades in the tree (not to scale). Photo credits: *Tibicina*, *Cicadetta*, *Monomatapa*, *Hovana*, *Yanga*: S.P.; *Cacama*, *Nosola*, *Thopha*: D. Marshall.

Ancestral area reconstruction

The best-fitting BioGeoBEARS model was DEC + *j* + *x* with time-stratified paleodistances, with an AIC, AICc, and BIC model weight of 79% (Table 4). The second best-fitting model was DIVALIKE + *j* + *x* with time-stratified paleodistances, with a model weight of 18%, indicating strong support in the data for time-stratified distance-dependent dispersal. The third best-fitting model, with a model weight of 2.3%, was DEC + *j* + *x* with fixed present-day distances. All of the model weight accrued to +*x* versions of the models, indicating that distance between areas plays an important role in dispersal. The mean estimate for *x* with time-stratified

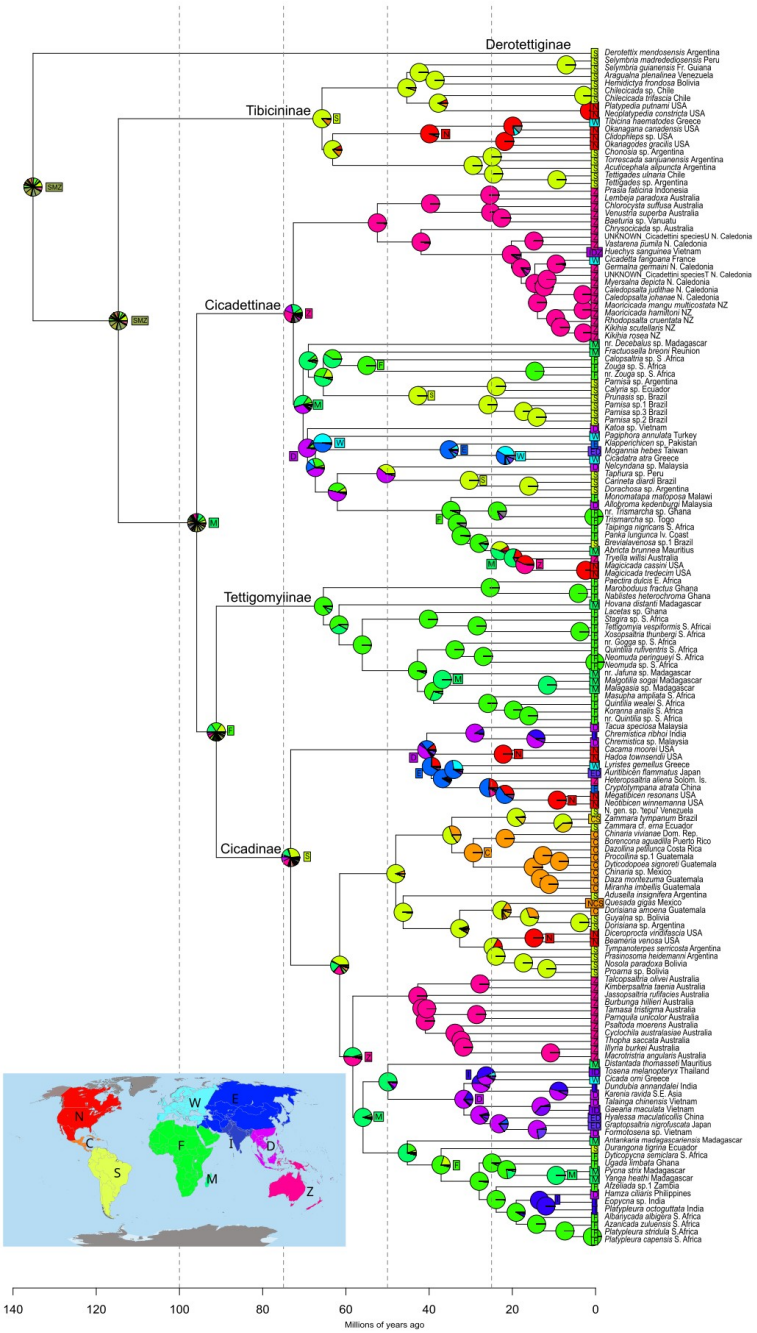
distances weighted by model weights is -0.5013, which corresponds to dispersal probability scaling with the square root of distance, i.e. diffusion.

Ancestral ranges were estimated using the best-fitting BioGeoBEARS model (time-stratified DEC + *j* + *x*). Cicadidae was estimated as having a Gondwanan origin, specifically South America, Africa, and Australasia (Fig. 4). Tibicininae and Cicadinae were both estimated to originate in South America, while Cicadettinae was estimated as originating in Australasia and Tettigomyiinae in Africa. However, the range estimations for Cicadidae, Cicadinae, and Cicadettinae had very high uncertainty.

Table 4: Statistical model comparison of BioGeoBEARS analyses

Model description	Name	lnL	*k*	*d*	*e*	*j*	*x*	AIC	AICc	BIC	AIC weight	AICc weight	BIC weight	LRT (+*j*)	LRT (+*x*)	LRT (+*j*+*x*)
M0: Dispersal independent of distance	DEC	-310.9	2	0.0010	0.0007	0.0	0.0	625.8	621.8	631.9	0.00%	0.00%	0.00%	2.3E-23		
	DEC+J	-261.3	3	0.0003	1.0E-12	0.0196	0.0	528.6	522.6	537.7	0.00%	0.00%	0.00%			
	DIVALIKE	-305.2	2	0.0012	1.0E-12	0.0	0.0	614.4	610.4	620.5	0.00%	0.00%	0.00%	8.6E-20		
	DIVALIKE+J	-263.7	3	0.0003	1.0E-12	0.0194	0.0	533.5	527.5	542.6	0.00%	0.00%	0.00%			
	BAYAREALIKE	-355.2	2	0.0010	0.0170	0.0	0.0	714.4	710.4	720.5	0.00%	0.00%	0.00%	4.5E-40		
	BAYAREALIKE+J	-267.4	3	0.0002	1.0E-07	0.0211	0.0	540.8	534.8	549.9	0.00%	0.00%	0.00%			
M4: Dispersal dependent on fixed modern distances	DEC+x	-300.9	3	0.0030	0.0008	0.0	-0.4004	607.9	601.9	617.0	0.00%	0.00%	0.00%	2.7E-24	9.1E-07	
	DEC+J+x	-249.2	4	0.0009	1.0E-12	0.0645	-0.4151	506.5	498.5	518.7	2.32%	2.32%	2.32%		8.1E-07	1.7E-27
	DIVALIKE+x	-293.0	3	0.0036	1.0E-12	0.0	-0.3947	592.1	586.1	601.2	0.00%	0.00%	0.00%	7.4E-20	7.0E-07	
	DIVALIKE+J+x	-251.4	4	0.0010	1.0E-12	0.0613	-0.4035	510.9	502.9	523.1	0.26%	0.26%	0.26%		1	4.5E-24
	BAYAREALIKE+x	-355.4	3	0.0010	0.0169	0.0	0.0074	716.8	710.8	725.9	0.00%	0.00%	0.00%	2.2E-45	1.0E-06	
	BAYAREALIKE+J+x	-255.5	4	0.0008	1.0E-07	0.0645	-0.3958	518.9	511.0	531.1	0.00%	0.00%	0.00%		3.4E-06	5.0E-44
M3: Dispersal dependent on time-stratified paleodistances	DEC+x	-300.1	3	0.0035	0.0008	0.0	-0.4435	606.2	600.3	615.4	0.00%	0.00%	0.00%	1.8E-25	2.4E-08	
	DEC+J+x	-245.7	4	0.0012	1.0E-12	0.0897	-0.5153	499.4	491.5	511.6	78.64%	78.64%	78.64%		5.5E-08	4.9E-29
	DIVALIKE+x	-290.4	3	0.0044	1.0E-12	0.0	-0.4558	586.9	580.9	596.0	0.00%	0.00%	0.00%	1.4E-20	8.5E-09	
	DIVALIKE+J+x	-247.2	4	0.0014	1.0E-12	0.0859	-0.5114	502.3	494.4	514.5	18.36%	18.36%	18.36%		9.4E-05	6.3E-26
	BAYAREALIKE+x	-347.5	3	0.0038	0.0166	0.0	-0.4439	701.1	695.1	710.2	0.00%	0.00%	0.00%	6.2E-44	9.5E-09	
	BAYAREALIKE+J+x	-250.9	4	0.0010	1.0E-07	0.0904	-0.5110	509.8	501.9	522.0	0.43%	0.43%	0.43%		0.0E+00	5.3E-46

Figure 4: BioGeoBEARS ancestral range estimations using best-fitting time-stratified DEC+*j*+*x* model.



Ancestral range estimates and pie charts showing relative probabilities of ancestral range states plotted on pruned dated phylogeny. Key for areas. N: North America; C: Central America and Caribbean islands; S: South America; F: Africa; M: Madagascar and Indian Ocean islands; W: West Eurasia; E: East Eurasia; I: India; D: Southeast Asia; Z: Australasia.

Discussion

Increased phylogenetic resolution of Cicadidae relationships

This study provides increased resolution for relationships within Cicadidae, particularly of the polytomy of Tettigomyiinae, Cicadettinae, and Cicadinae. Using five genes, Marshall et al. (2018), were first to identify Tettigomyiinae as a subfamily, but recovered the three subfamilies as a polytomy, with Tettigomyiinae itself not resolved as monophyletic. A monophyletic Tettigomyiinae sister to Cicadinettinae + Cicadinae was supported by additional genes (Simon et al. 2019), but phylogenomic analyses provided equal support for Cicadettinae + Tettigomyiinae and Cicadinae + Tettigomyiinae (Owen et al. 2022). Our analyses point to Tettigomyiinae + Cicadinae as the resolution of this polytomy. This is possibly a result of increased taxon sampling in Tettigomyiinae in this study, which went from eight (Owen et al. 2022) to 19 taxa.

The topologies in our concatenated ML, ASTRAL, and SVDQuartets phylogenies are quite similar, clarifying many relationships among taxa within Cicadidae and identifying areas needing taxonomic revision within the family. Our results provide molecular confirmation of the placement of *Hovana* and its tribe Hovanini into Tettigomyiinae (Sanborn et al. 2020). Our results also demonstrate that several African genera currently assigned to tribes in Cicadettinae

also belong to the subfamily Tettigomyiinae, increasing the recognized diversity of this subfamily. These include *Neomuda*, currently assigned to Lamotialnini, *Koranna*, currently assigned to Parnisini, *Malgotilia*, currently assigned to Parnisini, and *Masupha*, also currently assigned to Parnisini. The taxa assigned to Parnisini are especially significant, as Marshall et al. (2018) identified this tribe as having a problematic distribution, with the true members of the tribe probably endemic to South America. Interestingly, Sanborn (2023) removed *Brevialavenosa* from Lamotialnini and placed it in the Taphurini (Cicadettinae) saying, "The transfer of both Brazilian species previously assigned to *Abroma* means there are no members of Lamotialnini remaining in Brazil or South America confirming the tribal distribution as outlined in Marshall et al. (2018)." Our results clearly place *Brevialavenosa* into the Lamotialnini, making it the second genus in this tribe, other than *Magicicada*, to be confirmed in the Americas. During this preparation of this manuscript, Sanborn (2025) moved *Brevialavenosa* back into the Lamotialnini on the basis of our preliminary results and after morphological re-evaluation. One other genus from the Americas, *Chrysolasia*, is currently classified in Lamotialnini (Moulds 2003), but we were unable to extract DNA from the only remaining specimens (loaned by the NHMUK). Our results also agree with Owen et al. (2022) in placing *Aragualna* in Tibicininae and provide phylogenomic support for the placement of *Hemidictya* (Hemidictyini) in Tibicininae by Sanborn et al. (2020). The placement of *Antankaria* (Antankariini) and *Durangona* (Durangonini) differ across analyses and receive low branch support in the ASTRAL and SVDQuartets trees (Fig. 1). The concatenated tree shows Antankariini (from Madagascar) as sister to Platypleurini (primarily from Africa and Madagascar) and Durangonini (from South

America) as sister to Fidicinini + Zammarini (South and Central America), both with high support, which makes the most sense biogeographically. The other phylogenetic analyses place these two taxa elsewhere near taxa from more distant geographic areas with lower support.

Divergence times and the age of Cicadidae

Our results suggest that Cicadidae arose in the early Cretaceous and underwent diversification in the late Cretaceous and early Paleogene, coinciding with the K-Pg mass extinction event. Previous studies have generally assumed an early Paleocene or late Cretaceous origin for Cicadidae based on the family's fossil record (Moulds 2018; Ruschel and Campos 2019; Simon et al. 2019; Wang et al. 2022). The oldest fossil unambiguously assigned to Cicadidae, *Davispia bearcreekensis*, is from the late Paleocene (Cooper 1941); a late Cretaceous fossil of a first-instar nymph, *Burmacicada protera*, ambiguously represents Cicadidae or Tettigarctidae (Moulds 2018). We estimate that Cicadidae may have an early Cretaceous crown age, which would imply a nearly 70 Myr "phylogenetic fuse", meaning a time interval between when diversification began and when it first appeared in the fossil record (Cooper and Fortey 1998; Marshall 2019). The proposed existence of long phylogenetic fuses has recently been criticized as an analyitical artifact stemming from interaction among too-broad calibration densities (Budd and Mann 2020, 2023). Our methods sought to avoid this artifact by using FBD process to calibrate the tree, rather than assigning nodes subjective calibration densities. Furthermore, *D. bearcreekensis* is assigned to an extant tribe (Tibicinini, Tibicininae) that arose

after the split between Derotettiginae and all other Cicadidae subfamilies, which lends more credence to the hypothesis that Cicadidae is considerably older than this fossil.

Our results, likely influenced by our selection of fossil taxa, place the crown age of Cicadoidea in the Early-to-Mid Jurassic. Several fossils assigned to Tettigarctidae have been identified from the Early Jurassic (Shcherbakov 2009), and a few have been identified from the Late Triassic (Lambkin 2019; Jouault et al. 2024). One such fossil, *Sanmai*? *zetavena* (Jouault et al. 2024), has been assigned to a genus that a morphological phylogenetic analysis suggests is a stem-cicadid (Jiang et al. 2024). While these fossils may push the crown age of Cicadoidea into the Triassic, we elected to not use these early tettigarctid fossils in our analysis for several reasons. Jiang et al. (2024) showed that many morphological characters used to assign fossils to Tettigarctidae are plesiomorphic, and that many fossils currently assigned as such may be stem-cicadoids, stem-tettigarctids, or stem-cicadids. Given that these Triassic and Early Jurassic fossils are quite fragmentary, usually consisting of a only complete or partial forewing, we believe their phylogenetic placement is questionable and they may in fact be stem-cicadoids. We therefore elected to include only Mesozoic fossils that had been included in the phylogenetic analysis of Jiang et al. (2024).

Although our use of the tip-dating FBD process avoids artifacts associated with node-dating approaches, there is evidence that the inferred crown age for Cicadidae is slightly inflated. If the inferred crown age of Cicadidae were correct, this would suggest a preservation or discovery bias towards stem-group taxa during the Cretaceous, given that there are several stem-group fossils known from the long interval between the Cicadidae crown age and the oldest

656 fossil belonging to Cicadidae. Our youngest fossils from the stem-Cicadidae lineage are also at
657 least 40 Myr younger than their split from the lineage leading to crown Cicadidae (Jiang et al.
658 2024). However, such long terminal branches for sampled stem lineages might not be entirely
659 problematic, because our included stem-Cicadoidea fossil is 80 Myr younger than the inferred
660 split between Cicadidae and Tettigarctidae, and approximately 60 Myr younger than the oldest
661 stem-fossils assigned to those families.

662 Our FBD analyses were conducted as to minimize any possible artifacts such as deep root
663 attraction, in which violations of model assumptions inflate the ages of deeper nodes (Ronquist
664 et al. 2016; Matschiner 2019; Luo et al. 2023; Zhang et al. 2023). The specific model violations
665 involve the assumption that extant tips are sampled randomly and that the diversification,
666 turnover, and fossil sampling rates of the FBD process are constant through time. In our
667 divergence time analysis strategy, the backbone tree analysis was most likely to be affected by
668 these model violations. To minimize the effect of deep root attraction, we used the age estimates
669 from the backbone tree analysis that incorporated diversified sampling (Zhang et al. 2016) and
670 changing process rates through time (Zhang et al. 2023). The diversified sampling extension of
671 the FBD model incorporates a cut-off age before the present, after which no speciation or fossil
672 sampling events are allowed to occur (Zhang et al. 2016, 2023). However, we did not incorporate
673 the diversified sampling FBD model in the subtree analyses. The subtree datasets incorporated
674 many very recent fossils and also consisted of taxon samples that were intermediate between
675 random and diversified sampling, which would result in the cut-off age being so young as to be

irrelevant. As a result of these analysis choices, we believe that our age estimates for Cicadidae may be inflated by deep root attraction, but only modestly so.

Additional precision for cicada divergence times could be obtained with more sophisticated fossil placements. The fossil placements in our tree are entirely determined by the tree prior through monophyly constraints, which themselves are based on taxonomic assignments by specialists. Incorporating morphological data into a combined molecular and morphological analysis would allow more precise fossil placement, and information from the morphological evolution rate could shrink both the stem-lineage branch lengths and the crown age estimate (Gavryushkina et al. 2017; Luo et al. 2020). While a morphological data matrix was available for the stem-group fossils and many extant species in our dataset (Moulds 2005; Simon et al. 2019; Jiang et al. 2024), we lacked morphological data for the crown fossils for Cicadidae used in this analysis. Scoring the same morphological character data for the crown-group fossils to complement the stem-fossils and extant species would allow for future refinement of divergence times for Cicadidae through a combined data analysis and provide exploration of the history of morphological evolution in Cicadidae. Unfortunately, many of the known fossils assigned to Cicadidae are only fragmentary, making this approach fraught.

Even accounting for modest inflation of divergence times due to deep root attraction, Cicadidae clearly have a Cretaceous origin and present-day diversity within the Cicadidae subfamilies originated after the Cretaceous-Paleogene mass extinction. As is the case for birds (Brocklehurst and Field 2024) and placental mammals (Carlisle et al. 2023), frogs (Feng et al. 2017), and snakes (Klein et al. 2021), the stem-lineages for all five Cicadidae subfamilies were

extant prior to the mass extinction event and crossed the boundary. Four of the five subfamilies diversified around and after the transition boundary, which raises the question why the fifth subfamily, Derotettiginae, did not proliferate. Simon et al. (2019) propose that during a period of mid-Miocene habitat change, the Derotettiginae faced competition from cicadas in the Tibicininae that were better adapted to the new habitat. The sister family to Cicadidae, Tettigarctidae, declined after the K-Pg transition. The extensive Mesozoic fossil record of Tettigarctidae indicates that it was a thriving lineage prior to the K-Pg mass extinction, while the comparatively few Cenozoic fossils suggest that the lineage was in decline after the transition (Shcherbakov 2009; Moulds 2018). However, some Mesozoic fossils currently assigned to Tettigarctidae may actually represent the stem-lineage of Cicadidae (Jiang et al. 2024). Simulation results suggest that stem groups decline rapidly after the origination of the crown group, and that mass extinctions accelerate this process (Budd and Mann 2020; Hoyal Cuthill et al. 2020). The fossil record of insects suggests that mass extinctions are associated with periods of high turnover of insect lineages rather than absolute reductions in insect diversity (Labandeira 2005; Béthoux 2009; Nicholson et al. 2015; Condamine et al. 2016; Wang et al. 2016; Montagna et al. 2019; Schachat and Labandeira 2021), a phenomenon that is compatible with the results presented here. North American placental mammals also experienced a spike in turnover associated with the K-Pg mass extinction (Pires et al. 2018). While not a global mass extinction, the cooling and drying sparked by the establishment of the Antarctic Circumpolar Current and a drop in global CO_2 levels at the Eocene-Oligocene boundary was associated with high turnover in the Australian pygopodoid gecko radiation (Brennan and Oliver 2017) and initiated a period of

exuberant episodic diversification of the Australian desert fauna (e.g., the *Pauropsalta* cicada complex of 11 genera; Owen et al. 2017). The Cenozoic aridification of Australia is associated with extinction of mesic-adaptive lineages (Byrne et al. 2011). Since the extant species of Tettigarctidae are restricted to mesic habitat in southeastern Australia and Tasmania, we speculate that these climatic changes may have contributed to the decline of Tettigarctidae.

Comparison with Cicadidae divergence times estimated through node-dating

Our divergence times for lineages within Cicadidae using the FBD model tip-dating method are largely congruent with previously published studies using node-dating approaches. Our estimated divergence time for Platypleurini (28.0 Ma; HPD: 20.3-35.7 Ma) is very close to previously published ages estimated using node dating and molecular clocks (Price et al. 2019). Similarly, we estimated a divergence time of 40.6 Ma (HPD: 32.0-51.1 Ma) for Tacuini, which is slightly older but with a considerably narrower range than the estimate from Hill et al. (2015). The age estimates for divergences within the tribe are also comparable. Our results are also congruent with the previously published divergence times for the radiation of the tribes within Asian Cicadinae (Hill et al. 2021; Wang et al. 2024). We estimate that the tribes of this radiation (Cicadini, Dundubiini, Gaeanini, Polyneurini, Sonatini, and Tosenini) began to diversify approximately 31.7 Ma (HPD: 24.2-39.8 Ma), which is almost identical to the age inferred by Hill et al. (2021). Our age estimates for shallower divergences within this radiation are also consistent with previous estimates (Hill et al. 2021; Wang et al. 2024). The congruence of our age estimates with previous estimates is likely related to the overlap in fossil calibrations used,

despite the difference in dating method. Our analysis included two fossils for Polyneurini, one of which was used in a previous node-dating estimate (Wang et al. 2024), and we included twelve fossils across all the Asian tribes, five of which were used in a previous study (Hill et al. 2021). Our sampling for Cicadettini excluded some clades that are sister to the rest of the tribe, meaning that our Cicadettini sample does not include the deepest splits in the tribe. Our sampling for Cicadettini corresponds to the large clade descending from the Cicadettini radiation (sister to *Chrysocicada*) that does not include the *Pauropsalta* complex (Marshall et al. 2016). We estimated a mean divergence time of 20.2 Ma (95% HPD: 13.1-28.0 Ma) for this clade, which is on the younger end of estimates from the various molecular clock-based analyses of Marshall et al. (2016). Those authors' estimates suffered from substantial inflation of molecular clock rates, which they tried to rein in using a variety of calibration methods. Our results for Cicadettini have the advantage of the inclusion of a second fossil belonging to the tribe from the Miocene (Moulds et al. 2022), clock rate information from many nuclear loci, the inclusion of many other taxa whose presence can constrain the species tree clock rate in this clade, and a model that accounts for the phylogenetic uncertainty of fossil placement (Gavryushkina et al. 2014; Heath et al. 2014; Zhang et al. 2023).

In a study examining the co-evolution of cicadas and their symbionts, Wang et al. (2022) inferred a crown age for Cicadidae in the late Cretaceous very close to the K-Pg transition (76-66 Ma) based on mitochondrial COI sequences and eight fossil calibrations. This younger date is more consistent with the observed fossil record for Cicadidae, considering the absence of unambiguous crown Cicadidae fossils from the Mesozoic. However, we argue for our older date

based on the advantages of the FBD model's incorporation of fossils into the phylogenetic analysis. Also, Wang et al. (2022) chose to assign *D. bearcreekensis* as a calibration for the root of Tibicininae, despite the fact that this fossil is assigned to an extant tribe within the subfamily. Furthermore, the maximum age of their tree was constrained by the Misof et al. (2014) estimate of 150 Myr for the age of divergence between Cicadoidea and Cercopoidea. This divergence estimate is certainly an underestimate, as the oldest fossils from Cicadoidea have been dated to over 200 Ma (Shcherbakov 2009; Lambkin 2019; Jouault et al. 2024). More recent time-calibrated phylogenies for Hemipteran insects that incorporate the oldest fossils from Cicadoidea push the divergence between Cicadoidea and Cercopoidea to the late Triassic, around 215 Ma (Johnson et al. 2018). Based on these factors, we believe that the younger date for crown Cicadidae obtained by Wang et al. (2022) is an underestimate.

Osozawa and Wakabayashi (2025) inferred crown ages for Cicadoidea and Cicadidae of 200 Ma and 99 Ma, respectively, using mitochondrial COI and 18S rRNA sequences, twelve geologic calibrations, and thirteen fossil calibrations. However, this phylogeny is not calibrated through typical node-dating approaches, in which calibration densities are constructed around minimum ages determined by the fossil age and maximum ages constrained by prior knowledge or other model parameters (Ho and Phillips 2009). Instead, this phylogeny is a minimum age tree, which avoids the perceived subjectivity surrounding selecting maximum ages for calibration densities and instead calibrates nodes directly to their oldest respective fossils (Klopfstein 2021). Because the minimum age tree does not attempt to estimate the mean or maximum age of a lineage, direct comparison with ages inferred from our FBD analyses or other node-dating

studies is difficult. For example, Osozawa and Wakabayashi (2025) estimate the minimum age of Cicadettinae at 30 Ma, which is almost 30 Myr younger than the lower bound of the 95% credible interval in our analyses (Table 3).

Our dating results are congruent with previous node-dating estimates, which raises the question of whether using the FBD model is worth the extra model complexity. We think it is. An advantage of modeling the fossilized birth-death process is that the subjective choices for the node prior calibration densities are replaced with a model incorporating fossil information into the diversification process (Heath et al. 2014). Furthermore, information from all available fossils is incorporated into the estimate rather than simply the oldest fossil assigned to a node. Furthermore, our results for Cicadettini and Tacuini show improvement and refinement of age estimates over previously published dates, demonstrating the benefit of using the FBD model.

Historical biogeography of Cicadidae

The ancestral area reconstruction analyses provide insight into the origin of Cicadidae and the processes that shaped its diversification. The two best-fitting models both involve dispersal dependent on time-stratified paleodistances, which is encouraging because it suggests that the model is capturing actual biological signal in the data. Unsurprisingly, the subfamilies of Cicadidae that are most geographically restricted (Tibicininae and Tettigomyiinae) have the lowest uncertainty in their ancestral reconstructions, reflecting the small amount of range evolution required to produce their current day distributions. In contrast, the two subfamilies with global distributions (Cicadinae and Cicadettinae) along with the family as a whole have

very high uncertainty in their reconstructed origins. This may be due to the diversity of geographic ranges held by their descendant clades, leading to many different scenarios of range evolution that the data are unable to discriminate between. For example, Cicadettinae contains major clades with African, Madagascan, South American, and Eurasian distributions, leading to Australasia, Madagascar, and Southeast Asia being reconstructed as possible ancestral areas for the subfamily. Given that Cicadettini, by far the most speciose tribe of Cicadettinae, is absent from Madagascar despite having an otherwise global distribution and is extremely speciose in Australia (Marshall et al. 2016), Australasia seems the most likely origin for Cicadettinae. Cicadinae contains clades with predominantly Australasian, South and Central American, Southeast Asian, and African and Madagascan distributions, which results in the reconstructed origin of Cicadinae being split between South America, Madagascar, Australasia, and Southeast Asia. Due to the uncertainty of the reconstruction for the deepest clades in the tree, the data do not conclusively support a South American origin for Cicadidae, as proposed by Simon et al. (2019), but they do not rule it out either.

Even with uncertainty regarding the deepest clades, cicadas appear to have been involved in several instances of long-distance dispersal of several thousand kilometers. These include apparent dispersal between South America and Southeast Asia, South America and Australasia, South America and Madagascar, and Madagascar and Australasia. The reconstruction of these long-distance dispersal events may be facilitated by the estimated value of the distance-dependent dispersal parameter x being approximately -0.5. This parameter modifies the rate of anagenetic dispersal and the event weights for cladogenetic jump-dispersal such that $x = -1$

823 corresponds to an inverse distance relationship for dispersal rate, and x = -2 corresponds to an
824 inverse-square relationship (Van Dam and Matzke 2016). Another explanation for the
825 reconstructed long-distance dispersal events is taxon sampling. A major limitation of the DEC-
826 type historical biogeography models used in BioGeoBEARS is that they ignore the process of
827 lineage extinction and unsampled taxa in the phylogeny (Matzke 2014). As the dated phylogeny
828 used in the ancestral area reconstruction contains under 4% of extant Cicadidae species,
829 unsampled lineages may have a very large effect on inference. For example, in the case of the
830 massive Cicadettini radiation, fine-scale sampling provided lineages bridging the gap between
831 the origin of the tribe in Australasia and descendant lineages in Africa, Europe, and North
832 America (Marshall et al. 2016). We believe our taxon sampling maximizes the number of
833 geographically important cicada lineages, but sampling of additional cicada genera, particularly
834 in Africa, may provide additional insight into the biogeographic history of the family. A
835 biogeography model incorporating SSE (state-dependent speciation-extinction), such as a
836 ClaSSE model (Goldberg and Igić 2012), would be able to account for unsampled lineages and
837 lineage extinction, at the cost of computational complexity (Matzke 2014). Although it was not
838 included in our analysis, Antarctica may also have facilitated long-distance dispersal of lineages.
839 The continent would have been inhabitable by cicadas prior to the greenhouse-icehouse
840 transition (Barker and Thomas 2004), and may have played a key role connecting South
841 American and Australasian lineages.

842 Several mechanisms for long-distance dispersal have been proposed, including storms or
843 prevailing winds, transport by birds, or ocean currents (Gillespie et al. 2012). Given that only

cicada adults live above ground, it is unlikely that cicadas could disperse long-distance as nymphs, except possibly through rafting attached to tree roots. Of the usually discussed mechanisms for long-distance dispersal, transport of adults via storms and of eggs on rafting vegetation appear most likely (Osozawa et al. 2021). The absence of cicadas on isolated oceanic islands like those of the Hawaiian chain (Simon et al. 1984; Simon 1987), or even islands near continents like the Canary Islands, suggest that such dispersal is limited.

Funding

This work was supported by the National Science Foundation (grant numbers DEB1655891, DEB0955849, DEB0720664, DEB0529679, and DEB0089946) and Fulbright New Zealand.

Acknowledgements

We thank the numerous collaborators from around the globe who assisted with field work and provided specimens and sequence data for this project, many of whom are listed in the acknowledgements of Owen et al. (2022). We especially thank David Marshall and Kathy Hill, who were invaluable in coordinating specimen collecting efforts. We thank Michelle Kortyna at the Center for Anchored Phylogenomics for assistance with AHE library preparation. We also thank Alexei Drummond and the University of Auckland Centre for Computational Evolution for hosting MS on a Fulbright Fellowship. Computational resources were provided by the University of Connecticut Computational Biology Core. Elizabeth Jockusch, Paul Lewis, and David

Marshall provided helpful suggestions and comments on the manuscript. David Marshall and Max Moulds provided comments on the taxonomic summary appendix.

Data Availability

Data and scripts are available at Dryad Digital Repository:
http://datadryad.org/share/LINK_NOT_FOR_PUBLICATION/DHJLmoEdp3L8QYJeiUO-UiGPCO0INIRmAoNoMxAOA-E

Supplementary Information

Table S1: List of specimens

* Taxonomy needing to be updated based on relationships in this study

† Specimen obtained from Natural History Museum, London, UK (NHMUK)

‡ Specimen obtained from Muséum national d'Histoire naturelle (MNHN)

Online appendix 1: Summary of the subfamilies, tribes and genera including extant and fossil taxa with comments on classification changes since Marshall et al. (2018).

References

Akaike H. 1973. Information theory as an extension of the maximum likelihood principle. Á In: Petrov, BN and Csaki, F. Second International Symposium on Information Theory. Akademiai Kiado, Budapest, pp. 276Á281.

Arensburger P., Buckley T.R., Simon C., Moulds M., Holsinger K.E. 2004. Biogeography and phylogeny of the New Zealand cicada genera (Hemiptera: Cicadidae) based on nuclear and mitochondrial DNA data. Journal of Biogeography. 31:557–569.

Ballesteros J.A., Hormiga G. 2016. A New Orthology Assessment Method for Phylogenomic Data: Unrooted Phylogenetic Orthology. Molecular Biology and Evolution. 33:2117–2134.

Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., Nikolenko S.I., Pham S., Prjibelski A.D., Pyshkin A.V., Sirotkin A.V., Vyahhi N., Tesler G., Alekseyev M.A., Pevzner P.A. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol. 19:455–477.

Barido-Sottani J., Aguirre-Fernández G., Hopkins M.J., Stadler T., Warnock R. 2019. Ignoring stratigraphic age uncertainty leads to erroneous estimates of species divergence times under the fossilized birth–death process. Proceedings of the Royal Society B: Biological Sciences. 286:20190685.

Barker P.F., Thomas E. 2004. Origin, signature and palaeoclimatic influence of the Antarctic Circumpolar Current. Earth-Science Reviews. 66:143–162.

Bator J., Marshall D.C., Hill K.B.R., Cooley J.R., Leston A., Simon C. 2021. Phylogeography of the endemic red-tailed cicadas of New Zealand (Hemiptera: Cicadidae: Rhodopsalta), and molecular, morphological and bioacoustical confirmation of the existence of Hudson's Rhodopsalta microdora. Zoological Journal of the Linnean Society.

Berry V., Gascuel O. 1996. On the Interpretation of Bootstrap Trees: Appropriate Threshold of Clade Selection and Induced Gain. Molecular Biology and Evolution. 13:999.

Béthoux O. 2009. Gaps and nodes between fossil and extant insects. Systematic Entomology. 34:599–609.

Boderau M., Fu Y., Jiang H., Guan S., Peng A., Nel A., Jouault C. 2025a. Bayesian modelling of the fossil record enlightens the evolutionary history of Hemiptera. Proceedings of the Royal Society B: Biological Sciences. 292:20251133.

Boderau M., Nel A., Jouault C. 2025b. Diversification and extinction of Hemiptera in deep time. Commun Biol. 8:352.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., Fourment M., Gavryushkina A., Heled J., Jones G., Kühnert D., Maio N.D., Matschiner M., Mendes F.K., Müller N.F., Ogilvie H.A., Plessis L. du, Popinga A., Rambaut A., Rasmussen D., Siveroni I., Suchard M.A., Wu C.-H., Xie D., Zhang C., Stadler T., Drummond A.J. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLOS Computational Biology. 15:e1006650.

Bouckaert R.R., Drummond A.J. 2017. bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evol Biol. 17:42.

Bouckaert R.R., Heled J. 2014. DensiTree 2: Seeing Trees Through the Forest. :012401.

Brennan I.G., Oliver P.M. 2017. Mass turnover and recovery dynamics of a diverse Australian continental radiation. Evolution. 71:1352–1365.

Brocklehurst N., Field D.J. 2024. Tip dating and Bayes factors provide insight into the divergences of crown bird clades across the end-Cretaceous mass extinction. Proceedings of the Royal Society B: Biological Sciences. 291:20232618.

Buckley T.R., Cordeiro M., Marshall D.C., Simon C. 2006. Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (Maoricicada Dugdale). Systematic Biology. 55:411–425.

Buckley T.R., Simon C. 2007. Evolutionary radiation of the cicada genus Maoricicada Dugdale (Hemiptera: Cicadoidea) and the origins of the New Zealand alpine biota. Biological Journal of the Linnean Society. 91:419–435.

Budd G.E., Mann R.P. 2020. The dynamics of stem and crown groups. Science Advances. 6:eaaz1626.

Budd G.E., Mann R.P. 2023. Two Notorious Nodes: A Critical Examination of Relaxed Molecular Clock Age Estimates of the Bilaterian Animals and Placental Mammals. Systematic Biology.:syad057.

Bushnell B., Rood J., Singer E. 2017. BBMerge – Accurate paired shotgun read merging via overlap. PLOS ONE. 12:e0185056.

Byrne M., Steane D.A., Joseph L., Yeates D.K., Jordan G.J., Crayn D., Aplin K., Cantrill D.J., Cook L.G., Crisp M.D., Keogh J.S., Melville J., Moritz C., Porch N., Sniderman J.M.K.,

Sunnucks P., Weston P.H. 2011. Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. Journal of Biogeography. 38:1635–1656.

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T.L. 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10:1–9.

Cao W., Zahirovic S., Flament N., Williams S., Golonka J., Müller R.D. 2017. Improving global paleogeography since the late Paleozoic using paleobiology. Biogeosciences. 14:5425–5439.

Cao X., Zahirovic S., Li S., Suo Y., Wang P., Liu J., Müller R.D. 2022. A deforming plate tectonic model of the South China Block since the Jurassic. Gondwana Research. 102:3–16.

Carlisle E., Janis C.M., Pisani D., Donoghue P.C.J., Silvestro D. 2023. A timescale for placental mammal diversification based on Bayesian modeling of the fossil record. Current Biology. 33:3073-3082.e3.

Carruthers T., Sun M., Baker W.J., Smith S.A., de Vos J.M., Eiserhardt W.L. 2022. The Implications of Incongruence between Gene Tree and Species Tree Topologies for Divergence Time Estimation. Systematic Biology. 71:1124–1146.

Chernomor O., von Haeseler A., Minh B.Q. 2016. Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. Systematic Biology. 65:997–1008.

Chifman J., Kubatko L. 2014. Quartet Inference from SNP Data Under the Coalescent Model. Bioinformatics. 30:3317–3324.

Condamine F.L., Clapham M.E., Kergoat G.J. 2016. Global patterns of insect diversification: towards a reconciliation of fossil and molecular evidence? Sci Rep. 6:19208.

Cooper A., Fortey R. 1998. Evolutionary explosions and the phylogenetic fuse. Trends in Ecology & Evolution. 13:151–156.

Cooper K.W. 1941. Davispia bearcreekensis Cooper, a new cicada from the Paleocene, with a brief review of the fossil Cicadidae. American journal of Science. 239:286–304.

Degnan J.H., Rosenberg N.A. 2006. Discordance of Species Trees with Their Most Likely Gene Trees. PLOS Genetics. 2:e68.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution. 24:332–340.

Dietrich C.H., Allen J.M., Lemmon A.R., Lemmon E.M., Takiya D.M., Evangelista O., Walden K.K.O., Grady P.G.S., Johnson K.P. 2017. Anchored Hybrid Enrichment-Based Phylogenomics of Leafhoppers and Treehoppers (Hemiptera: Cicadomorpha: Membracoidea). Insect Systematics and Diversity. 1:57–72.

Douglas J., Jiménez-Silva C.L., Bouckaert R. 2022. StarBeast3: Adaptive Parallelized Bayesian Inference under the Multispecies Coalescent. Systematic Biology. 71:901–916.

Douglas J., Zhang R., Bouckaert R. 2021. Adaptive dating and fast proposals: Revisiting the phylogenetic relaxed clock model. PLOS Computational Biology. 17:e1008322.

Edwards S.V. 2009. IS A NEW AND GENERAL THEORY OF MOLECULAR SYSTEMATICS EMERGING? Evolution. 63:1–19.

Erwin D.H. 2001. Lessons from the past: Biotic recoveries from mass extinctions. Proceedings of the National Academy of Sciences. 98:5399–5403.

Feng Y.-J., Blackburn D.C., Liang D., Hillis D.M., Wake D.B., Cannatella D.C., Zhang P. 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous-Paleogene boundary. Proceedings of the National Academy of Sciences of the United States of America. 114:E5864–E5870.

Franco A.D., Poujol R., Baurain D., Philippe H. 2019. Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. BMC Evol Biol. 19:1–17.

Fu L., Niu B., Zhu Z., Wu S., Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 28:3150–3152.

Gavryushkina A., Heath T.A., Ksepka D.T., Stadler T., Welch D., Drummond A.J. 2017. Bayesian Total-Evidence Dating Reveals the Recent Crown Radiation of Penguins. Systematic Biology. 66:57–73.

Gavryushkina A., Welch D., Stadler T., Drummond A.J. 2014. Bayesian Inference of Sampled Ancestor Trees for Epidemiology and Fossil Calibration. PLOS Computational Biology. 10:e1003919.

Gillespie R.G., Baldwin B.G., Waters J.M., Fraser C.I., Nikula R., Roderick G.K. 2012. Long-distance dispersal: a framework for hypothesis testing. Trends in Ecology & Evolution. 27:47–56.

Goldberg E.E., Igić B. 2012. TEMPO AND MODE IN PLANT BREEDING SYSTEM EVOLUTION. Evolution. 66:3701–3709.

Guindon S., Dufayard J.-F., Lefort V., Anisimova M., Hordijk W., Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology. 59:307–321.

Hahn C., Bachmann L., Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. Nucleic Acids Research.

Haji D., Vailionis J., Stukel M., Gordon E., Lemmon E.M., Lemmon A.R., Simon C. 2022. Lack of host phylogenetic structure in the gut bacterial communities of New Zealand cicadas and their interspecific hybrids. Sci Rep. 12:20559.

Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. PNAS. 111:E2957–E2966.

Heled J., Bouckaert R.R. 2013. Looking for trees in the forest: summary tree from posterior samples. BMC Evol Biol. 13:221.

Heled J., Drummond A.J. 2010. Bayesian Inference of Species Trees from Multilocus Data. Mol Biol Evol. 27:570–580.

Hertach T., Puissant S., Gogala M., Trilar T., Hagmann R., Baur H., Kunz G., Wade E.J., Loader S.P., Simon C., Nagel P. 2016. Complex within a Complex: Integrative Taxonomy Reveals Hidden Diversity in Cicadetta brevipennis (Hemiptera: Cicadidae) and Unexpected Relationships with a Song Divergent Relative. PLOS ONE. 11:e0165562.

Hill K.B.R., Marshall D.C., Marathe K., Moulds M.S., Lee Y.J., Pham T.-H., Mohagan A.B., Sarkar V., Price B.W., Duffels J.P., Schouten M.A., Boer A.J. de, Kunte K., Simon C.

2021. The molecular systematics and diversification of a taxonomically unstable group of Asian cicada tribes related to Cicadini Latreille, 1802 (Hemiptera : Cicadidae). Invert. Systematics. 35:570–601.

Hill K.B.R., Marshall D.C., Moulds M.S., Simon C. 2015. Molecular phylogenetics, diversification, and systematics of Tibicen Latreille 1825 and allied cicadas of the tribe Cryptotympanini, with three new genera and emphasis on species from the USA and Canada(Hemiptera: Auchenorrhyncha: Cicadidae). Zootaxa. 3985.

Hill K.B.R., Simon C., Marshall D.C., Chambers G.K. 2009. Surviving glacial ages within the Biotic Gap: Phylogeography of the New Zealand cicada Maoricicada campbelli. Journal of Biogeography. 36:675–692.

Hime P.M., Lemmon A.R., Lemmon E.C.M., Prendini E., Brown J.M., Thomson R.C., Kratovil J.D., Noonan B.P., Pyron R.A., Peloso P.L.V., Kortyna M.L., Keogh J.S., Donnellan S.C., Mueller R.L., Raxworthy C.J., Kunte K., Ron S.R., Das S., Gaitonde N., Green D.M., Labisko J., Che J., Weisrock D.W. 2020. Phylogenomics Reveals Ancient Gene Tree Discordance in the Amphibian Tree of Life. Syst Biol. 70:49–66.

Ho S.Y.W., Phillips M.J. 2009. Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. Syst Biol. 58:367–380.

Hoang D.T., Chernomor O., von Haeseler A., Minh B.Q., Vinh L.S. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution. 35:518–522.

Holder M.T., Sukumaran J., Lewis P.O. 2008. A Justification for Reporting the Majority-Rule Consensus Tree in Bayesian Phylogenetics. Systematic Biology. 57:814–821.

Hoyal Cuthill J.F., Guttenberg N., Budd G.E. 2020. Impacts of speciation and extinction measured by an evolutionary decay clock. Nature. 588:636–641.

Hull P. 2015. Life in the Aftermath of Mass Extinctions. Current Biology. 25:R941–R952.

Hurvich C.M., Tsai C.-L. 1989. Regression and time series model selection in small samples. Biometrika. 76:297–307.

Jablonski D. 2001. Lessons from the past: Evolutionary impacts of mass extinctions. Proceedings of the National Academy of Sciences. 98:5393–5398.

Jiang H., Szwedo J., Labandeira C.C., Chen J., Moulds M.S., Mähler B., Muscente A.D., Zhuo D., Nyunt T.T., Zhang H., Wei C., Rust J., Wang B. 2024. Mesozoic evolution of cicadas and their origins of vocalization and root feeding. Nat Commun. 15:376.

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., Percy D.M., Hardy N.B., Terry I., Liu S., Zhou X., Misof B., Robertson H.M., Yoshizawa K. 2018. Phylogenomics and the evolution of hemipteroid insects. PNAS. 115:12775–12780.

Jouault C., Nam G.-S., Boderau M., Kwon S.-H., Nel A. 2024. A new Triassic Tettigarctidae (Insecta, Hemiptera) from the Amisan Formation (Republic of Korea). Palaeoentomology. 7:132–140.

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

Katoh K., Standley D.M. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution. 30:772–780.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P., Drummond A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28:1647–1649.

Klein C.G., Pisani D., Field D.J., Lakin R., Wills M.A., Longrich N.R. 2021. Evolution and dispersal of snakes across the Cretaceous-Paleogene mass extinction. Nat Commun. 12:5335.

Klopfstein S. 2021. The age of insects and the revival of the minimum age tree. Austral Entomology. 60:138–146.

Knyshov A., Gordon E.R.L., Weirauch C. 2021. New alignment-based sequence extraction software (ALiBaSeq) and its utility for deep level phylogenetics. PeerJ. 9:e11019.

Kubatko L.S., Gibbs H.L., Bloomquist E.W. 2011. Inferring Species-Level Phylogenies and Taxonomic Distinctiveness Using Multilocus Data in Sistrurus Rattlesnakes. Syst Biol. 60:393–409.

Labandeira C.C. 2005. The Fossil Record of Insect Extinction: New Approaches and Future Directions. Am Entomol. 51:14–29.

Lambkin K.J. 2019. Mesodiphthera Tillyard, 1919, from the Late Triassic of Queensland, the oldest cicada (Hemiptera: Cicadomorpha: Cicadoidea: Tettigarctidae). Zootaxa. 4567:358–366.

Lanfear R., Calcott B., Kainer D., Mayer C., Stamatakis A. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol Biol. 14:82.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Molecular Biology and Evolution. 34:772–773.

Lemmon A.R., Emme S.A., Lemmon E.M. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Systematic Biology.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 25:1754–1760.

Liu L., Yu L., Kubatko L., Pearl D.K., Edwards S.V. 2009. Coalescent methods for estimating phylogenetic trees. Molecular Phylogenetics and Evolution. 53:320–328.

Łukasik P., Chong R.A., Nazario K., Matsuura Y., Bublitz D.A.C., Campbell M.A., Meyer M.C., Van Leuven J.T., Pessacq P., Veloso C., Simon C., McCutcheon J.P. 2019. One Hundred Mitochondrial Genomes of Cicadas. Journal of Heredity. 110:247–256.

Luo A., Duchêne D.A., Zhang C., Zhu C.-D., Ho S.Y.W. 2020. A Simulation-Based Evaluation of Tip-Dating Under the Fossilized Birth–Death Process. Systematic Biology. 69:325–344.

Luo A., Zhang C., Zhou Q.-S., Ho S.Y.W., Zhu C.-D. 2023. Impacts of Taxon-Sampling Schemes on Bayesian Tip Dating Under the Fossilized Birth-Death Process. Systematic Biology. 72:781–801.

Maddison W.P. 1997. Gene Trees in Species Trees. Systematic Biology. 46:523–536.

Marshall C.R. 2019. Using the Fossil Record to Evaluate Timetree Timescales. Front. Genet. 10.

Marshall D.C., Cooley J.R. 2000. Reproductive Character Displacement and Speciation in Periodical Cicadas, with Description of a New Species, 13-Year Magicicada Neotredecim. Evolution. 54:1313–1325.

Marshall D.C., Hill K.B.R., Cooley J.R., Simon C. 2011. Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: Lessons from New Zealand cicadas (Genus Kikihia). Systematic Biology. 60:482–502.

Marshall D.C., Hill K.B.R., Fontaine K.M., Buckley T.R., Simon C. 2009. Glacial refugia in a maritime temperate climate: Cicada (Kikihia subalpina) mtDNA phylogeography in New Zealand. Molecular Ecology. 18:1995–2009.

Marshall D.C., Hill K.B.R., Marske K.A., Chambers C., Buckley T.R., Simon C. 2012. Limited, episodic diversification and contrasting phylogeography in a New Zealand cicada radiation. BMC Evolutionary Biology. 12:1.

Marshall D.C., Hill K.B.R., Moulds M., Vanderpool D., Cooley J.R., Mohagan A.B., Simon C. 2016. Inflation of Molecular Clock Rates and Dates: Molecular Phylogenetics, Biogeography, and Diversification of a Global Cicada Radiation from Australasia (Hemiptera: Cicadidae: Cicadettini). Systematic Biology. 65:16–34.

Marshall D.C., Moulds M., Hill K.B.R., Price B.W., Wade E.J., Owen C.L., Goemans G., Marathe K., Sarkar V., Cooley J.R., Sanborn A.F., Kunte K., Villet M.H., Simon C. 2018. A molecular phylogeny of the cicadas (Hemiptera: Cicadidae) with a review of tribe and subfamily classification. Zootaxa. 4424:1–64.

Marshall D.C., Slon K., Cooley J.R., Hill K.B.R., Simon C. 2008. Steady Plio-Pleistocene diversification and a 2-million-year sympatry threshold in a New Zealand cicada radiation. Molecular Phylogenetics and Evolution. 48:1054–1066.

Matschiner M. 2019. Selective Sampling of Species and Fossils Influences Age Estimates Under the Fossilized Birth–Death Model. Front. Genet. 10.

Matzke N. 2018. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis with R scripts. version 1.1. 1. San Francisco, CA: GitHub. Available online at: https://github. com/nmatzke/BioGeoBEARS.

Matzke N.J. 2013. BioGeoBEARS: BioGeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts. R package, version 0.2.

Matzke N.J. 2014. Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. Systematic Biology. 63:951–970.

Matzke N.J. 2022. Statistical comparison of DEC and DEC+J is identical to comparison of two ClaSSE submodels, and is therefore valid. Journal of Biogeography. 49:1805–1824.

Mendes F.K., Hahn M.W. 2016. Gene Tree Discordance Causes Apparent Substitution Rate Variation. Syst Biol. 65:711–721.

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution. 37:1530–1534.

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., Blanke A., Berger S., Böhm A., Buckley T.R., Calcott B., Chen J., Friedrich F., Fukui M., Fujita M., Greve C., Grobe P., Gu S., Huang Y., Jermiin L.S., Kawahara A.Y., Krogmann L., Kubiak M., Lanfear R., Letsch H., Li Y., Li Z., Li J., Lu H., Machida R., Mashimo Y., Kapli P., McKenna D.D., Meng G., Nakagaki Y., Navarrete-Heredia J.L., Ott M., Ou Y., Pass G., Podsiadlowski L., Pohl H., von Reumont B.M., Schütte K., Sekiya K., Shimizu S., Slipinski A., Stamatakis A., Song W., Su X., Szucsich N.U., Tan M., Tan X., Tang M., Tang J., Timelthaler G., Tomizuka S., Trautwein M., Tong X., Uchifune T., Walzl M.G., Wiegmann B.M., Wilbrandt J., Wipfler B., Wong T.K.F., Wu Q., Wu G., Xie Y., Yang S., Yang Q., Yeates D.K., Yoshizawa K., Zhang Q., Zhang R., Zhang W., Zhang Y., Zhao J., Zhou C., Zhou L., Ziesmann T., Zou S., Li Y., Xu X., Zhang Y., Yang H., Wang J., Wang J., Kjer K.M., Zhou X. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science (New York, N.Y.). 346:763–7.

Montagna M., Tong K.J., Magoga G., Strada L., Tintori A., Ho S.Y.W., Lo N. 2019. Recalibration of the insect evolutionary time scale using Monte San Giorgio fossils suggests survival of key lineages through the End-Permian Extinction. Proceedings of the Royal Society B: Biological Sciences. 286:20191854.

Moulds M., Frese M., McCurry M.R. 2022. New cicada fossils from Australia (Hemiptera: Cicadoidea: Cicadidae) with remarkably detailed wing surface nanostructure. Alcheringa: An Australasian Journal of Palaeontology. 46:264–276.

Moulds M., Kaulfuss U., Gehler A. 2023. A review of the singing cicada fossils from the Pliocene Fossil-Lagerstätte Willershausen, Germany, with the description of three new species (Hemiptera: Cicadidae: Tibicinae and Cicadinae). Palaeoentomology. 6:459–471.

Moulds M.S. 2003. An appraisal of the cicadas of the genus Abricta Stål and allied genera (Hemiptera: Auchenorrhyncha: Cicadidae). Rec. Aust. Mus. 55:245–304.

Moulds M.S. 2005. An Appraisal of the Higher Classification of Cicadas (Hemiptera: Cicadoidea) with Special Reference to the Australian Fauna. Records of the Australian Museum. 57:375–446.

Moulds M.S. 2018. Cicada fossils (Cicadoidea: Tettigarctidae and Cicadidae) with a review of the named fossilised Cicadidae. Zootaxa. 4438.

Moulds M.S. 2020. Reassessment of the cicada fossil Oncotympana lapidescens Zhang, 1989 with the description of a new genus (Hemiptera, Cicadoidea, Cicadidae). Palaeoentomology. 3:333–337.

Müller N.F., Bouckaert R.R. 2020. Adaptive Metropolis-coupled MCMC for BEAST 2. PeerJ. 8:e9473.

Müller R.D., Cannon J., Qin X., Watson R.J., Gurnis M., Williams S., Pfaffelmoser T., Seton M., Russell S.H.J., Zahirovic S. 2018. GPlates: Building a Virtual Earth Through Deep Time. Geochemistry, Geophysics, Geosystems. 19:2243–2261.

Müller R.D., Zahirovic S., Williams S.E., Cannon J., Seton M., Bower D.J., Tetley M.G., Heine C., Le Breton E., Liu S., Russell S.H.J., Yang T., Leonard J., Gurnis M. 2019. A Global Plate Model Including Lithospheric Deformation Along Major Rifts and Orogens Since the Triassic. Tectonics. 38:1884–1907.

Nguyen H.Q., Ho P.-T., Kong S., Bae Y., Pham T.H., La H.T., Jang Y. 2022. A time-calibrated mitogenomic phylogeny suggests that Korean Hyalessa fuscata is a bridge between Chinese and Japanese H. maculaticollis. J Genet. 102:9.

Nicholson D.B., Mayhew P.J., Ross A.J. 2015. Changes to the Fossil Record of Insects through Fifteen Years of Discovery. PLOS ONE. 10:e0128554.

Ogilvie H.A., Bouckaert R.R., Drummond A.J. 2017. StarBEAST2 Brings Faster Species Tree Inference and Accurate Estimates of Substitution Rates. Mol Biol Evol. 34:2101–2114.

Ogilvie H.A., Heled J., Xie D., Drummond A.J. 2016. Computational Performance and Statistical Accuracy of *BEAST and Comparisons with Other Methods. Syst Biol. 65:381–396.

Osozawa S., Kanai K., Fukuda H., Wakabayashi J. 2021. Phylogeography of Ryukyu insular cicadas: Extensive vicariance by island isolation vs accidental dispersal by super typhoon. PLOS ONE. 16:e0244342.

Osozawa S., Wakabayashi J. 2025. Cicada minimum age tree: Cryptic speciation and exponentially increasing base substitution rates in recent geologic time. .

Owen C.L., Marshall D.C., Hill K.B.R., Simon C. 2017. How the Aridification of Australia Structured the Biogeography and Influenced the Diversification of a Large Lineage of Australian Cicadas. Systematic Biology. 66:569–589.

Owen C.L., Marshall D.C., Wade E.J., Meister R., Goemans G., Kunte K., Moulds M., Hill K., Villet M., Pham T.-H., Kortyna M., Lemmon E.M., Lemmon A.R., Simon C. 2022. Detecting and Removing Sample Contamination in Phylogenomic Data: An Example and its Implications for Cicadidae Phylogeny (Insecta: Hemiptera). Systematic Biology.:syac043.

Owen C.L., Miller G.L. 2022. Phylogenomics of the Aphididae: Deep relationships between subfamilies clouded by gene tree discordance, introgression and the gene tree anomaly zone. Systematic Entomology. 47:470–486.

Peng J., Swofford D.L., Kubatko L. 2022. Estimation of speciation times under the multispecies coalescent. Bioinformatics. 38:5182–5190.

Pires M.M., Rankin B.D., Silvestro D., Quental T.B. 2018. Diversification dynamics of mammalian clades during the K–Pg mass extinction. Biol. Lett. 14:20180458.

Price B.W., Barker N.P., Villet M.H. 2010. A watershed study on genetic diversity: Phylogenetic analysis of the *Platypleura plumosa* (Hemiptera: Cicadidae) complex reveals catchment-specific lineages. Molecular Phylogenetics and Evolution. 54:617–626.

Price B.W., Marshall D.C., Barker N.P., Simon C., Villet M.H. 2019. Out of Africa? A dated molecular phylogeny of the cicada tribe Platypleurini Schmidt (Hemiptera: Cicadidae), with a focus on African genera and the genus Platypleura Amyot & Audinet-Serville. Systematic Entomology. 44:842–861.

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature. 526:569–573.

Puissant S., Gurcel K. 2023. The genus Tibicina Kolenati, 1857 in Morocco (Hemiptera: Cicadidae: Tibicininae): taxonomic assessment from integrative research. Zootaxa. 5330:535–560.

Ree R.H., Sanmartín I. 2018. Conceptual and statistical problems with the DEC+J model of founder-event speciation and its comparison with DEC via model selection. Journal of Biogeography. 45:741–749.

Ronquist F., Lartillot N., Phillips M.J. 2016. Closing the gap between rocks and clocks using total-evidence dating. Philosophical Transactions of the Royal Society B: Biological Sciences. 371:20150136.

Ruschel T.P., Campos L.A. 2019. Phylogeny and biogeography of the leaf-winged cicadas (Hemiptera: Auchenorrhyncha: Cicadidae). Zoological Journal of the Linnean Society. 185:1150–1187.

Salichos L., Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature. 497:327–331.

Sanborn A.F. 2023. Resolving taxonomic issues of cicadas (Hemiptera: Cicadidae) including new combinations, new synonymies, and revised status, with updates on the diversity of the Brazilian cicada fauna and new records for four South American countries. Zootaxa. 5318:339–362.

Sanborn A.F. 2025. Redescription, illustration and higher taxonomy of the Lesser Antilles cicada Chalumalna martinesisBoulard, 2001 (Hemiptera: Cicadidae: Cicadettinae). Journal of Insect Biodiversity. 62:33–41.

Sanborn A.F., Marshall D.C., Moulds M.S., Puissant S., Simon C. 2020. Redefinition of the cicada tribe Hemidictyini Distant, 1905, status of the tribe Iruanini Boulard, 1993 rev. stat., and the establishment of Hovanini n. tribe and Sapantangini n. tribe (Hemiptera: Cicadidae). Zootaxa. 4747.

Sayyari E., Mirarab S. 2016. Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. Molecular Biology and Evolution. 33:1654–1668.

Schachat S.R., Labandeira C.C. 2021. Are Insects Heading Toward Their First Mass Extinction? Distinguishing Turnover From Crises in Their Fossil Record. Annals of the Entomological Society of America. 114:99–118.

Schulte P., Alegret L., Arenillas I., Arz J.A., Barton P.J., Bown P.R., Bralower T.J., Christeson G.L., Claeys P., Cockell C.S., Collins G.S., Deutsch A., Goldin T.J., Goto K., Grajales-Nishimura J.M., Grieve R.A.F., Gulick S.P.S., Johnson K.R., Kiessling W., Koeberl C., Kring D.A., MacLeod K.G., Matsui T., Melosh J., Montanari A., Morgan J.V., Neal C.R., Nichols D.J., Norris R.D., Pierazzo E., Ravizza G., Rebolledo-Vieyra M., Reimold W.U., Robin E., Salge T., Speijer R.P., Sweet A.R., Urrutia-Fucugauchi J., Vajda V., Whalen M.T., Willumsen P.S. 2010. The Chicxulub Asteroid Impact and Mass Extinction at the Cretaceous-Paleogene Boundary. Science. 327:1214–1218.

Schwarz G. 1978. Estimating the dimension of a model. The annals of statistics.:461–464.

Seabra S.G., Quartau J.A., Bruford M.W. 2009. Spatio-temporal genetic variation in sympatric and allopatric Mediterranean Cicada species (Hemiptera, Cicadidae). Biological Journal of the Linnean Society. 96:249–265.

Shcherbakov D.E. 2009. Review of the fossil and extant genera of the cicada family Tettigarctidae (Hemiptera: Cicadoidea). Russian Entomological Journal. 17:343–348.

Simmons M.P., Gatesy J. 2021. Collapsing dubiously resolved gene-tree branches in phylogenomic coalescent analyses. Molecular Phylogenetics and Evolution. 158:107092.

Simon C. 1987. Hawaiian evolutionary biology: An introduction. Trends in Ecology & Evolution. 2:175–178.

Simon C., Gordon E.R.L., Moulds M.S., Cole J.A., Haji D., Lemmon A.R., Lemmon E.M., Kortyna M., Nazario K., Wade E.J., Meister R.C., Goemans G., Chiswell S.M., Pessacq

P., Veloso C., McCutcheon J.P., Łukasik P. 2019. Off-target capture data, endosymbiont genes and morphology reveal a relict lineage that is sister to all other singing cicadas. Biological Journal of the Linnean Society. 128:865–886.

Simon C.M., Gagné W.C., Howarth F.G., Radovsky F.J. 1984. Hawai'i: A Natural Entomological Laboratory. Bulletin of the Entomological Society of America. 30:9–17.

Smith B.T., Merwin J., Provost K.L., Thom G., Brumfield R.T., Ferreira M., Mauck W.M. III, Moyle R.G., Wright T.F., Joseph L. 2023. Phylogenomic Analysis of the Parrots of the World Distinguishes Artifactual from Biological Sources of Gene Tree Discordance. Systematic Biology. 72:228–241.

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:e0197433.

Smith S.A., O'Meara B.C. 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics. 28:2689–2690.

Stadler T., Kühnert D., Bonhoeffer S., Drummond A.J. 2013. Birth–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). Proceedings of the National Academy of Sciences. 110:228–233.

Stadler T., Smrckova J. 2016. Estimating shifts in diversification rates based on higher-level phylogenies. Biology Letters. 12:20160273.

Steenwyk J.L., Buida T.J. III, Labella A.L., Li Y., Shen X.-X., Rokas A. 2021. PhyKIT: a broadly applicable UNIX shell toolkit for processing and analyzing phylogenomic data. Bioinformatics. 37:2325–2331.

Stukel M., Porczak A.E., Gordon E.R.L., Vailionis J., Haji D., Buckley T.R., Lemmon A.R., Lemmon E.M., Simon C. 2024. Phylogenomics improves the phylogenetic resolution and provides strong evidence of mito-nuclear discordance in two genera of a New Zealand cicada (Hemiptera: Cicadidae) species radiation. Systematic Entomology. 49:237–257.

Tamura K., Battistuzzi F.U., Billing-Ross P., Murillo O., Filipski A., Kumar S. 2012. Estimating divergence times in large molecular phylogenies. Proceedings of the National Academy of Sciences. 109:19333–19338.

Tobella C., Franch M., Bas J.M., Brotons L., Pons P. 2025. Spatial Dynamics and Lifespan of Adult Cicadas After Fire and Logging: A Radiotracking Study. Integrative Zoology. n/a.

Torsvik T.H., Steinberger B., Shephard G.E., Doubrovine P.V., Gaina C., Domeier M., Conrad C.P., Sager W.W. 2019. Pacific-Panthalassic Reconstructions: Overview, Errata and the Way Forward. Geochemistry, Geophysics, Geosystems. 20:3659–3689.

Udvardy M.D.F. 1975. A Classification of the Biogeographic Provinces of the World. International Union for Conservation of Nature and Natural Resources. No. 18:48-pp.

Van Dam M.H., Matzke N.J. 2016. Evaluating the influence of connectivity and distance on biogeographical patterns in the south-western deserts of North America. Journal of Biogeography. 43:1514–1532.

Villet M. 1995. Intraspecific variability in SMRS signals: some causes and implications in acoustic signaling systems. In: Lambert D.L., Spencer H.G., editors. Speciation and the recognition concept: theory and application. Baltimore, MD: John Hopkins University Press. p. 422–439.

Wang D., Huang Z., Billen J., Zhang G., He H., Wei C. 2022. Complex co-evolutionary relationships between cicadas and their symbionts. Environmental Microbiology. 24:195–211.

Wang J., Zhang W., Yang J., Hayashi M., Wei C. 2024. Phylogeny and evolution of the cicada tribe Polyneurini (Hemiptera, Cicadidae). Systematic Entomology. n/a.

Wang Y., Engel M.S., Rafael J.A., Wu H., Rédei D., Xie Q., Wang G., Liu X., Bu W. 2016. Fossil record of stem groups employed in evaluating the chronogram of insects (Arthropoda: Hexapoda). Sci Rep. 6:38939.

Whitfield J.B., Kjer K.M. 2008. Ancient Rapid Radiations of Insects: Challenges for Phylogenetic Analysis. Annu. Rev. Entomol. 53:449–472.

Wilgenbusch J.C., Swofford D. 2003. Inferring Evolutionary Trees with PAUP*. Current Protocols in Bioinformatics. 00:6.4.1-6.4.28.

Young A., Flament N., Maloney K., Williams S., Matthews K., Zahirovic S., Müller R.D. 2019. Global kinematics of tectonic plates and subduction zones since the late Paleozoic Era. Geoscience Frontiers. 10:989–1013.

Zahirovic S., Eleish A., Doss S., Pall J., Cannon J., Pistone M., Tetley M.G., Young A., Fox P. 2022. Subduction and carbonate platform interactions. Geoscience Data Journal. 9:371–383.

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

Zhang C., Ronquist F., Stadler T. 2023. Skyline Fossilized Birth–Death Model is Robust to Violations of Sampling Assumptions in Total-Evidence Dating. Systematic Biology. 72:1316–1336.

Zhang C., Stadler T., Klopfstein S., Heath T.A., Ronquist F. 2016. Total-Evidence Dating under the Fossilized Birth–Death Process. Systematic Biology. 65:228–249.