

Inflation of Molecular Clock Rates and Dates: Molecular Phylogenetics, Biogeography, and Diversification of a Global Cicada Radiation from Australasia (Hemiptera: Cicadidae: Cicadettini)

DAVID C. MARSHALL^{1,*}, KATHY B. R. HILL¹, MAX MOULDS², DAN VANDERPOOL^{1,3}, JOHN R. COOLEY¹,
ALMA B. MOHAGAN⁴, AND CHRIS SIMON¹

¹Department of Ecology and Evolutionary Biology, 75 N. Eagleville Rd., Storrs, CT 06269, USA; ²Entomology Department, Australian Museum, 6 College Street, Sydney NSW 2010, Australia; ³Division of Biological Sciences, Health Sciences 304, U. Montana, Missoula, MT 59812; ⁴Central Mindanao University, Sayre Highway, Bukidnon, Philippines

*Correspondence to be sent to: Department of Ecology and Evolutionary Biology, 75 N. Eagleville Rd., Storrs, CT 06269, USA;
E-mail: david.marshall@uconn.edu.

Received 4 April 2014; reviews returned 16 September 2015; accepted 17 September 2015
Associate Editor: Karl Kjer

Abstract.—Dated phylogenetic trees are important for studying mechanisms of diversification, and molecular clocks are important tools for studies of organisms lacking good fossil records. However, studies have begun to identify problems in molecular clock dates caused by uncertainty of the modeled molecular substitution process. Here we explore Bayesian relaxed-clock molecular dating while studying the biogeography of ca. 200 species from the global cicada tribe Cicadettini. Because the available fossils are few and uninformative, we calibrate our trees in part with a cytochrome oxidase I (COI) clock prior encompassing a range of literature estimates for arthropods. We show that tribe-level analyses calibrated solely with the COI clock recover extremely old dates that conflict with published estimates for two well-studied New Zealand subclades within Cicadettini. Additional subclade analyses suggest that COI relaxed-clock rates and maximum-likelihood branch lengths become inflated relative to EF-1 α intron and exon rates and branch lengths as clade age increases. We present corrected estimates derived from: (i) an extrapolated EF-1 α exon clock derived from COI-calibrated analysis within the largest New Zealand subclade; (ii) *post hoc* scaling of the tribe-level chronogram using results from subclade analyses; and (iii) exploitation of a geological calibration point associated with New Caledonia. We caution that considerable uncertainty is generated due to dependence of substitution estimates on both the taxon sample and the choice of model, including gamma category number and the choice of empirical versus estimated base frequencies. Our results suggest that diversification of the tribe Cicadettini commenced in the early- to mid-Cenozoic and continued with the development of open, arid habitats in Australia and worldwide. We find that Cicadettini is a rare example of a global terrestrial animal group with an Australasian origin, with all non-Australasian genera belonging to two distal clades. Within Australia, we show that Cicadettini is more widely distributed than any other cicada tribe, diverse in temperate, arid and monsoonal habitats, and nearly absent from rainforests. We comment on the taxonomic implications of our findings for thirteen cicada genera. [Aridification; Australia; branch lengths; calibration; climate change; gamma distribution; maximum-likelihood; relaxed clock.]

The simultaneous development of automated sequencing technology and relaxed-clock dating methods that accommodate among-lineage rate variation (Thorne and Kishino 2002; Drummond et al. 2006; Drummond and Rambaut 2007) has increased confidence in the application of molecular clocks. It is widely known that clock rates of specific genes vary across organisms due to life history and other factors (Ho and Lo 2013), and that careful estimates of clock rates are lacking for most genes and most organisms. At the same time, awareness is growing of additional limitations that go beyond uncertainty in the clocks themselves, problems that appear when estimating the number of molecular substitutions on the phylogeny. This uncertainty has been linked to model choice (e.g., Phillips 2009; Papadopoulou et al. 2010; Schwartz and Mueller 2010; Andujar et al. 2012; Obbard et al. 2012), calibration method (Brandley et al. 2011; Andujar et al. 2012), and other aspects of phylogenetic analyses (Brown et al. 2010; Marshall 2010; Schwartz and Mueller 2010; Nelson et al. 2015).

In our studies of diversification of a global cicada tribe, one for which fossil information is poor, we have encountered new evidence of this problem—estimated relative rates of genes, and therefore divergence times of

clades, are substantially influenced by the taxon sample. We suspect that these patterns could be common in deeper-level molecular dating studies, a concern that is further supported by recent simulation work (Schwartz and Mueller 2010), and we show how this phenomenon was detected and handled in our case. Overall, our results suggest that divergence time analyses based entirely on molecular clocks estimated from other taxa will often be problematic for deeper-level studies.

Diversification of Cicadettini.—Cicadas of the tribe Cicadettini have attracted attention for rapid diversification and adaptive radiation, where Miocene-age (23–5 Ma) ancestors have diversified into a wide range of habitats (Arensburger et al. 2004; Buckley and Simon 2007; Marshall et al. 2008; Hill et al. 2009; Marshall et al. 2012; Owen et al. 2015a). Ongoing surveys of the tribe Cicadettini are discovering many new species (Gogala et al. 2008; Marshall et al. 2011; Popple 2013; Ewart et al. 2015; Hertach et al. 2015; Owen and Moulds 2016) and new genera (Ewart 2005; Ewart and Marques 2008; Puissant and Sueur 2010; Moulds 2012), accelerated by studies emphasizing analysis of male songs (e.g., Sueur and Puissant 2007; Gogala et al. 2012). With about

500 catalogued species, the tribe Cicadettini includes a disproportionate fraction of the 3300 described species from 41 Cicadidae tribes (Sanborn 2013) and, based on work from our group and others, is estimated to contain a large number of undescribed species. Despite its wide distribution, just one small genus is known from North America (Sanborn 2009; Sanborn and Heath 2012; Sanborn and Phillips 2013), and none are found or suspected in South America. Most described genera of Cicadettini (57 out of 88) are found in Australasia, with most of those from Australia (Moulds 1990; Moulds 2012; Sanborn 2013). Smaller centers of diversity are found in the Palearctic (especially Europe) and in South Africa (Villet and van Noort 1999; Puissant and Sueur 2010). This global distribution suggests the possibility of an Australasian origin following isolation of that continent from Antarctica and South America (Hall 2011), a biogeographic pattern for which few examples are known (De Jong 2001; Crisp et al. 2004).

In Australia, most species of Cicadettini occupy temperate, arid, and monsoonal habitats. Few are known from tropical rainforests (in Australia or globally). Dating Australian clades of Cicadettini, a task recently completed for the large *Pauropsalta* group (Owen et al. 2015a), will help to understand if and how they diversified during the well-documented aridification of Australia that intensified following its Oligocene isolation (Hill 1994; Byrne et al. 2008; Fujioka and Chappell 2010). This aridification caused a widespread replacement of rainforest habitat, which is limited today to the eastern fringe of the continent, by *Eucalyptus*-dominated forests, *Acacia* and chenopod shrublands, spinifex grasslands, and desert dunes.

In this study, we sought to test historical hypotheses on the diversification of the Cicadettini using molecular phylogenetic data and molecular clock calibrations. Specifically, we asked: Did the tribe Cicadettini originate in Australasia? Did it originate during the period of post-Gondwanan isolation of Australia? If so, when did Cicadettini reach Eurasia, Africa, and North America, and do these dispersal events match predicted timings implied by historical connections (Hall 2001; Sanmartín 2001; Vila et al. 2011)? Answers to these questions will contribute to growing knowledge of the assembly of global biota on the historical canvas of geology and climatology.

MATERIALS AND METHODS

Taxon Sampling and Biosurveys

Taxon sampling.—We preserved cicada tissue in 95% ethanol during expeditions to Australia, Argentina, Chile, China, Fiji, New Caledonia, New Guinea, New Zealand, Philippines, South Africa, and the USA. Considerable field effort was concentrated on Australia, Papua New Guinea, and South Africa, where the described diversity of Cicadettini is greatest, and on SE Asian tropical regions (Philippines, Vietnam, and

China). Samples from Europe and other regions were obtained through collaborators (Acknowledgments section). We also targeted genera from the related tribes Huechysini, Parnisini, and Taphurini that were suspected based on morphology to belong in Cicadettini. Overall, we obtained samples of 77 of 88 described genera of Cicadettini, including an estimated 45 from Australia (described and undescribed). *Pagiphora*, *Pinheya* and *Stellenboschia* were sequenced at a late stage and are discussed only briefly in the supplementary Results (see the Dryad data repository online at <http://dx.doi.org/10.5061/dryad.5590q>). Type species of genera were sampled in all but five of the 77 genera sampled—*Diemeniana*, *Euryphara*, *Kosemia*, *Myersalna*, and *Tympanistalna*.

Preliminary genetic sequencing (mitochondrial cytochrome oxidase I, or COI), as described below, was completed for over 300 species of Cicadettini plus others from related tribes. By excluding one member of each species pair for which uncorrected mtDNA divergence was less than 5%, we selected 200 in-group taxa. These species are listed in online Appendix 1 with taxonomic and locality information, along with a list of the genera not sampled.

A preliminary maximum-likelihood (ML) analysis of a concatenated phylogenomic Cicadidae data set (unpublished) supports *Chrysocada* (currently a member of Taphurini) as the sister lineage to Cicadettini (100% ML bootstrap). However, other genera in Taphurini are well separated from *Chrysocada* in that study. *Chlorocysta* and *Venustria* (Chlorocystini) form the next sister clade in the phylogenomic tree (100% ML bootstrap). We, therefore, used two genera from Chlorocystini and one genus from Prasiini as the out-group and we included *Chrysocada* and three undescribed species that clustered with it in preliminary analyses (Results section).

At sites sampled in Australia, we identified and audio-recorded (when possible) all cicada species collected or heard singing. This allowed us to estimate geographic variation in the proportion of cicada species from different genera and tribes. Nearly all cicadas can be identified to species on the basis of distinctive male calling songs.

Genetic Sequencing, Alignment, and Model Selection

Genomic DNA was extracted from 1–2 legs using the Qiagen DNeasy Tissue Kit (Qiagen, Valencia, California, USA) or the Clontech kit (Clontech, Mountain View, CA, USA). Using the primers and annealing temperatures shown in Supplementary Table S1, portions of three genes were amplified—mitochondrial COI, mitochondrial COII, and nuclear elongation factor 1 alpha (EF-1 α). The general PCR protocol was as follows: (i) hold at 94°C for 2–2.5 min; (ii) repeat 30 cycles of denature at 94°C for 45 s, anneal at temperature in Supplementary Table S1 for 45 s, and extend at 72°C for 2–2.5 min; (iii) hold at 72°C for 10 min. Annealing

temperatures were lowered as far as 45°C for some taxa. Amplified products were cleaned using the Clontech Extract Kit or ExoSAP-IT (USB Corp., Cleveland, OH). Cycle sequencing was conducted using the Applied Biosystems Big Dye Terminator v1.1 cycle sequencing kit at 1/8- to 1/4-scale reaction volume, and the product cleaned by Sephadex (Millipore) filtration and visualized on an Applied Biosystems ABI 3100 capillary sequencer. A modified sequencing protocol was used as follows: (i) hold at 96°C for 2 min; (ii) repeat 30 cycles of 96°C for 30 s, 50°C for 15 s, and 60°C for 2.5 min; (iii) hold at 60°C for 5 min. Sequences were analyzed using ABI Prism Sequencing Analysis software v3.7 (Applied Biosystems), initially aligned in Sequencher v3.1 (Gene Codes Corp., Ann Arbor, MI), and further aligned and checked by eye.

Partitioning schemes and MrBayes-available substitution models for each partition were selected using the “greedy” search algorithm in Partitionfinder v1.0.1 (Lanfear et al. 2012) with Python v2.7 (Python Software Foundation 2010), with the potential partitions including individual codon positions of protein-coding data (with nuclear and mtDNA data separately treated) and the combined EF-1 α intron data. Indels were coded using Seqstate v1.0 set to the “simple coding” option of Simmons and Ochoterena (2000), and these characters were modeled with the MkV model (Lewis 2001) as a separate binary partition. Data partitions were separately tested for base-composition bias using a chi-square test of homogeneity in Paup v4.0 (Swofford 1998).

Phylogenetic Estimation (Analyses A–D)

A list of phylogenetic and divergence time analyses discussed in this article is found in Table 1. Combined-data (analysis A), EF-1 α -only (analysis B), and mtDNA-only (analysis C) phylogenetic trees were estimated using MrBayes v3.2.1 (Ronquist et al. 2012). Default settings were used for the priors (e.g., *ngammapat*=4, *ratepr*=variable, *brlens*=unconstrained:exponential[50]). Model parameters were estimated separately (*unlinked*) for each partition for up to 90 million generations, for two independent, simultaneous runs (*nruns*=2). Effective sample sizes were checked using Tracer v1.5 (Rambaut and Drummond 2007). MrBayes 50% majority-rule consensus phylograms were used for the working trees. Because tests of compositional bias were significant ($P<0.05$) for the mtDNA third codon position, an additional MrBayes analysis of the mtDNA data was conducted without third position sites (analysis D) to check for any major effects of this bias.

ML analysis of the combined data set was conducted in Garli v2.0 (Zwickl 2006) to obtain bootstrap supports for the tree in analysis A. Model parameters and relative rates were separately estimated for each partition, and random starting trees were used, as in the MrBayes analyses. For each bootstrap

TABLE 1. List of phylogenetic analyses described in this study

Code	Analysis
A	Phylogenetic analysis of combined data, MrBayes (incl. Garli ML bootstrap), tribe Cicadettini
B	Phylogenetic analysis of EF-1 α , MrBayes, tribe Cicadettini
C	Phylogenetic analysis of mtDNA, MrBayes, tribe Cicadettini
D	Phylogenetic analysis of mtDNA—no third positions, MrBayes, tribe Cicadettini
E	Relaxed-clock analysis—COI clock, BEAST, tribe Cicadettini
F	Relaxed-clock analysis—COI clock plus fossils, BEAST, tribe Cicadettini
G	Relaxed-clock analysis—COI clock plus NCAL geology, BEAST, tribe Cicadettini
H1	Relaxed-clock analysis—extrapolating COI to EF-1 α , for Analysis H2, BEAST, MRK
H2	Relaxed-clock analysis—extrapolated-EF-1 α clock, BEAST, tribe Cicadettini
J	<i>Post hoc</i> scaling of BEAST chronogram, tribe Cicadettini (see Methods section)
K	Relaxed-clock analyses—COI clock, BEAST, Cicadettini subclades
L	ML optimization of single-partition COI under varying gamma discretization, Garli, tribe Cicadettini & MRK
M	ML optimization of three-partition COI under varying gamma discretization, Garli, tribe Cicadettini & MRK
N	Analysis E with empirical base frequencies, BEAST
P	Analysis E with HKY substitution model instead of GTR, BEAST

Notes: ML is maximum-likelihood. NCAL refers to a clade of endemic New Caledonian genera used for a geological calibration. MRK refers to the New Zealand *Maoricicada*-*Rhodopsalta*-*Kikihia* subclade of Cicadettini genera.

replicate, two heuristic searches were completed, each with *genthreshfortopterm* set to 10^6 and with *significantttopchange* and *scorethreshforterm* set to 0.01.

Divergence Time Estimation

General approach.—Relaxed-clock divergence times were estimated using BEAST v2.1.3 (Drummond and Rambaut 2007; Bouckaert et al. 2014); each analysis was run 8–10 times to check for repeatability. Because we calibrated our trees using molecular clock estimates, and because branch lengths/gene rates can have weak phylogenetic signal, we needed to minimize the effect of the Bayesian priors on our analyses (see Nelson et al. 2015 for discussion of this problem for branch lengths). Unfortunately, uniform prior distributions (which assign equal probability to all parameter values) are statistically improper and undesirable, and even proper bounded uniform distributions can create unexpected effects (e.g., dos Reis et al. 2014). During trial analyses, we found that some parameter choices influenced date estimates. For example, in analyses with no data, the default tree prior distorted the posterior COI clock rate by favoring shorter trees, while a 1/X tree prior performed better.

Because the molecular clock prior was derived from literature estimates of the whole-gene COI rate (see

below), all of the BEAST analyses were conducted with the COI data partitioned as a single data subset. To simplify these complicated analyses, which could become stuck on inferior solutions (Results section), the COII, EF-1 α intron and exon partitions were also handled as single data subsets; the scored EF-1 α indels were not included. The BEAST analyses were run until effective sample sizes exceeded 200 for most parameters, with any parameters failing that test checked for potential correlation with node ages. Maximum clade compatibility trees with median node heights were calculated using TreeAnnotator v1.7.5.

Supplementary Table S2 shows many of the parameter/prior settings for the analyses; details of the clock settings (*ucl.d.mean*) are given below. In all cases, the site and clock models (but not the tree model) were unlinked across partitions, the uncorrelated relaxed clock for each partition in the Clock Model panel of the BEAUti xml tool was set to “estimate”, and the “substitution rate” parameter in the Site Model panel was not estimated. “Fix mean substitution rate”, “Normalize”, and “Automatic set clock rate” were also not selected. In BEAST 2, the option exists to estimate the values specifying the shapes of some distributions (e.g., gamma, log normal), but we did not use this because the models are already complex.

We assumed a log normal relaxed-clock prior for each gene/partition. Studies have found that auto-correlated rates models fit certain data sets better than uncorrelated models (Lepage et al. 2007; Battistuzzi et al. 2010; but see Christin et al. 2014); we will address this possibility in our ongoing family-level analyses where life-history variation is more likely to be important. Rejection of the strict molecular clock model was confirmed *post hoc* by the observation that the 95% credible interval of *ucl.d.stdev* excluded zero.

Molecular-clock calibration, empirical COI priors (analysis E).—No fossil- or event-calibrated cicada clock estimates are yet published. While many factors are believed to influence evolutionary rates across lineages (Bromham and Penny 2003; Sloan and Taylor 2012), recent studies emphasize generation time as a major factor (Tsantes and Steiper 2009; Thomas et al. 2010) and apparently not body size (Thomas et al. 2006) or metabolic rate (Lanfear et al. 2007) as suggested in earlier work (Gillooly et al. 2005). Emphasizing the COI gene—the focus of most arthropod molecular clock work—we took a conservative approach and applied a relaxed-clock branch rate prior (to *ucl.d.mean*) that included a range of published rates for arthropods, excluding studies that did not account for among-site rate variation (see also Papadopoulou et al. 2010; Ho and Lo 2013) and some outlier estimates derived from parasitic arthropods (e.g., lice – Bess et al. 2014). This “informed prior” (Nelson et al. 2015) for the mean rate ranged from 0.007–0.018 substitutions/site/myr (s/s/myr). These single lineage rates correspond to accumulated “pairwise” divergence rates of 0.014–0.036 s/s/myr; the rates reported in this article are all single-lineage rates. The *ucl.d.mean*

parameter of the log normal COI prior was tuned in BEAUti so that the 95% confidence interval of the distribution just included the range of the empirical COI rates. The median of this distribution (0.0112 s/s/myr) is close to the widely applied “Brower” clock rate of “2.3%/Myr” (Brower 1994), a value that is better reported as 0.0115 s/s/myr because it is strongly influenced by the one corrected genetic distance in the study (Knowlton et al. 1993—COI, K2P). Because cicadas have long life cycles compared to other arthropods (commonly three or more years—Campbell et al. 2015), their mean COI clock rate might be on the slow side of our prior, but knowledge of rate evolution is too incomplete to afford a clear prediction. Furthermore, more recent molecular clock studies have tended to find higher rates, reflecting the increased parameterization of modern evolutionary models (Papadopoulou et al. 2010; Andujar et al. 2012). The conservative approach used here is designed to reduce the dependence of our conclusions on unjustifiably precise specification of the COI rate. Lastly, we believe that incorporating the prior uncertainty in the *ucl.d.mean* parameter is more appropriate here than doing so within the *ucl.d.stdev* parameter (e.g., Marshall et al. 2012), since many of the empirical studies involved clades of multiple genera.

Because BEAST estimates the topology, it was necessary to include the COII and EF-1 α data partitions, which required specifying rate priors for those partitions. Failing to include the slower-evolving nuclear-gene data would have biased the divergence time estimates as well as the topology, by shortening deeper branches relative to shallow ones (Phillips 2009; Brandley et al. 2011). Unfortunately, during trial analyses, we found that applying uniform rate priors to COII and EF-1 α , while using the log normal on COI, specified a biased relative-rate relationship between the data subsets. This was detected as posterior distributions of *ucl.d.mean* for COI that were shifted upward from the prior. In our approach, we needed the posterior confidence interval of the COI rate to match the prior to be sure that the posterior divergence times matched expectations given the empirical clock and the reconstructed substitution process. To escape this problem, we estimated ML relative partition rates for each analysis (in Garli), with the data partitioned and modeled as in the BEAST analyses, and used those relative rates to scale the COI median rate to the other data subsets. The resulting COII, EF-1 α exon, and EF-1 α intron scaled median rates were used in weak log normal distributions with large standard deviations, to allow the COI calibration to control the analysis. This led to posterior distributions of *ucl.d.mean* for COI that matched the prior more closely. Supplementary Table S3 shows the BEAST settings used for the scaled priors of the tribe-level analysis calibrated with COI only (analysis E) and the other tribe-level analyses below. See dos Reis et al. (2014) for more on unintended consequences of combining rate priors in Bayesian analyses.

COI clock calibration plus fossil data (analysis F).—Three fossils are available for the tribe Cicadettini, but none offer strong constraints. *Paracicadetta oligocenica* (Boulard and Nel 1990) has been dated to the Oligocene (Rupelian, Upper Stampian, minimum age of 28.4 Ma). This taxon was originally placed in *Cicadetta*, but its affiliation within Cicadettini is now unclear (MSM unpublished data) and it may belong elsewhere in Cicadettinae. Another species, *Cicada ungeri* Heer from the Miocene (Croatia, minimum age 5.3 Ma) will be transferred to Cicadettini but has no clear affiliation (MSM unpublished data). Finally, a German fossil has been identified as *Cicadetta montana* by Straus (1952) from the Pliocene (minimum age 1.8 Ma). Only the *C. montana* fossil can be confidently assigned within our tree (at the origin of the stem supporting *brevipennis* and *montana*, conservatively), but we decided to explore the consequences of assigning the remaining two fossils in their potentially most informative positions—*Paracicadetta* at the origin of the Cicadettini stem and *C. ungeri* at the origin of the R1 Eurasian clade (see Results) found in our trees—together with the *C. montana* calibration. This analysis (F) was also constrained by the empirical and scaled clock rate priors of analysis E.

Recently, Misof et al. (2014) published a fossil-calibrated insect phylogeny that estimated 105 Ma for the common ancestor of the cicada genus *Okanagana* and its closest relative in the data set, a froghopper (Cercopidae), as well as a mean (170 Ma) and 95% confidence interval estimate (245–90 Ma) for the next deepest node (the split to planthoppers, Fulgoridae). With this information, we multiplied the approximate maximum age of the planthopper node (245 Ma) by the ratio of the two mean ages noted above (0.62), to obtain a conservative fossil-based maximum age for the Cicadidae/Cercopidae split (150 Ma), which we applied to the tree root in analysis F.

COI clock plus geological calibration (analysis G).—Studies suggest that the New Caledonian landmass was submerged before 37 Ma and that the extant New Caledonian species are derived from ancestors that subsequently colonized the island (Grandcolas et al. 2008; Espeland and Murienne 2011; Nattier et al. 2011; Cruaud et al. 2012; Pillon 2012). In our analyses, we reconstructed a clade of New Caledonian endemic genera (*Ueana* Distant, *Kanakia* Distant, and New Caledonian “*Abroma*” Stål, here called the NCAL clade) that is connected to a section of the Cicadettini tree dominated by Australian ancestors (see below). This situation offers a potential maximum-age calibration for the most recent common ancestor (MRCA) of this clade.

Calibration of divergence time analyses with island-emergence data is appropriate only if the ancestor of the focal clade is unlikely to have survived on an unrecognized landmass before arriving on the island (e.g., Fleischer et al. 1998; Jordan et al. 2003; Bonacum et al. 2005). Because the *Abroma*–*Kanakia*–*Ueana* clade in the consensus phylogram is connected to the Cicadettini

phylogeny by an extremely short branch (see below), we argue that alternative refuges are unlikely to have played an important role. In analysis G, we assigned a uniform prior on the MRCA of the NCAL clade ranging from 0 to 37 Ma. This constraint was combined with the empirical COI and scaled COII and EF-1 α clock rate priors.

Molecular clock calibration, extrapolated EF-1 α clock (analyses H1, H2).—In analysis E and other preliminary analyses, we noticed that the divergence times for Cicadettini were much older than expected given published COI-based estimates for certain genera of Cicadettini and other fossil-calibrated Hemiptera dates (Results section). We suspected, in part from subsampling tests given below, that the amount of molecular substitution estimated for COI was inflated compared to that of EF-1 α in the tribe-level analysis. To test this hypothesis, we conducted a tribe-level BEAST analysis using cicada-specific EF-1 α exon and intron clock rates extrapolated from the empirical COI rate in a partitioned ML analysis of the well-studied New Zealand *Maoricicada*–*Rhodopsalta*–*Kikihia* clade (MRK clade). For this procedure, we supplemented the five MRK taxa from our tree with published COI and/or EF-1 α sequences from Genbank (COII sequences were not available for many species), yielding a 54-taxon expanded data set. For *Maoricicada*, where detailed song analyses have not yet been conducted, samples were removed from the available data set when pairwise corrected distances found in ML trees of Buckley and Simon (2007) were less than 0.02 s/s/myr (an approximate threshold for the appearance of song changes in *Kikihia*—Marshall et al. 2011), to avoid combining inter- and intraspecific samples in a single clock analysis (Ho et al. 2005; Ho and Larson 2006). Then, for analysis H1, we estimated relative gene rates for the COI, EF-1 α exon, and EF-1 α intron partitions using ML in a partitioned Garli analysis and used those rates to scale 95% confidence intervals for *ucl.d.mean* for the EF-1 α exon and intron partitions from the empirical COI clock rate. Finally, the extrapolated EF-1 α rates were used to calibrate a tribe-level BEAST analysis as above, but with the COI prior weakened (along with COII) to give the EF-1 α rates priority (analysis H2).

As an additional test to confirm the difference in the amount of COI substitution estimated for the MRK clade in the tribe- and genus-level contexts, ML was used in Garli to optimize the COI data partition alone on (i) a 54-taxon MRK-only topology manually constructed from the results of Marshall et al. (2008) and Buckley and Simon (2007) and (ii) the Cicadettini MrBayes topology from analysis A, using the model structure of the BEAST analyses in both cases.

Post hoc scaling to MRK clade age (analysis J).—In analysis J, as a fourth way of calibrating the Cicadettini tree, we used the age estimate and 95% confidence interval of the MRK MRCA, obtained from a separate BEAST

analysis of the 54-taxon MRK data set, to manually re-scale the tribe-level chronogram obtained from analysis E (while retaining the relative node ages). The MRK BEAST analysis was conducted with the empirical COI prior plus scaled, weak EF-1 α priors as in analysis E. This *post hoc* scaling method was useful because attempts to place a very narrow (i.e., strong) prior on the age of the MRK MRCA led to poor MCMC performance.

Subclade analyses (analysis K).—To explore the relationship between the taxon sample and gene rate estimates, monophyletic or paraphyletic groups of taxa were selected from the tribe-level BEAST tree and separately analyzed in BEAST as above (see Supplementary Fig. S1). Subclades were selected in two ways. For the first series, monophyletic or paraphyletic groups were chosen with different MRCA ages, containing from 27–38 species. In the second series, the large “main radiation” clade (see below) was subsampled to create progressively smaller taxon sets while maintaining the phylogenetic structure as much as possible. Appropriate models for each data set were selected by Partitionfinder. All analyses were calibrated with the empirical COI rate prior from analysis E, and scaled priors were obtained for the other gene partitions in each case as described above (Supplementary Table S4). We then compared the estimated absolute and relative gene rates and their relationships; this group of tests is referred to as analysis K.

Among-site rate variation, substitution model choice, and tree lengths (analyses L–P).—Tree length estimates can vary depending on the handling of among-site rate variation and the degree of data partitioning (Marshall et al. 2006; Papadopoulou et al. 2010). To further demonstrate the significance of the rate variation model for date estimates, we optimized branch lengths on the best mtDNA ML tree using unpartitioned (analysis L) and codon-position partitioned data sets (analysis M) while varying the number of gamma-distributed rate categories. This was done separately for the MRK and Cicadettini COI data sets using Garli v2.0, by starting each analysis with the best ML tree and setting *topoweight* = 0. We considered models with 2–20 gamma distribution categories per partition. Finally, in BEAST analyses modeled on analysis E, two minor variations of the Partitionfinder-selected COI model were employed to illustrate extreme model-dependence of divergence-time estimates—empirical base frequencies (analysis N), and HKY instead of GTR (analysis P).

Diversification Rates

A lineage-through-time (LTT) plot was calculated in R v. 3.0.2 (R Development Core Team, 2011) from the BEAST chronogram in analysis E above using the APE v. 3.0-11 (Paradis et al. 2004) and LASER v. 2.4.1 (Rabosky 2006) packages.

Biogeography

Likelihood-based estimates of ancestral ranges were calculated using the dispersal–extinction–cladogenesis model in Lagrange version 20130526 (Ree and Smith 2008) and the MrBayes phylogram from analysis A. Areas were coded as follows: Australasia (including New Zealand, New Caledonia, Papua New Guinea, and Fiji), Eurasia, Africa (sub-Saharan), and North America. Distributions including up to two areas were allowed, and no time constraints were incorporated. The three out-group taxa were each coded with Australasia+Eurasia to cause Lagrange to use this conservative distribution for the common ancestor. Python scripts for Lagrange were assembled using an online web configurator tool (Ree 2013).

RESULTS

Taxon Sampling and Biosurveys

Australian Cicadettini diversity and distribution.—Using distinguishable songs as a preliminary criterion for unique species, we recorded more than 500 species of Cicadettini (in approximately 45 genera) in Australia, across 1720 locations. These totals are comparable in species diversity to the Australian Gryllidae (crickets) (Ötte and Alexander 1983). From the remaining Australian cicada tribes, approximately 40 genera were sampled (described and undescribed) and over 200 species. Pinned voucher specimens are stored at the University of Connecticut Biological Collections Facility and in the collection of MSM, while frozen vouchers are stored in the Simon lab at the University of Connecticut. Song recordings for undescribed taxa, when available, will be deposited with the BioAcoustica Wildlife Sounds Database (<http://www.bio.acousti.ca>).

We found Cicadettini in many habitats—from grasslands to vegetated coastal and inland dunes to wetlands, shrublands and forests, both tropical and temperate; we found them singing in many plant communities including saltbush flats, shrublands with *Acacia*, *Eremophila*, or subalpine shrubs, and woodlands of *Casuarina*, paperbark, and *Eucalyptus*. The distribution and abundance of the four major Australian cicada groups is shown in Figure 1. Genera of the tribe Cicadettini were most abundant in arid, semiarid, and temperate habitats and scarcest in rainforest regions, with extremely contrasting abundance between tropical north Queensland and southwestern Australia (Fig. 1f).

Genetic Sequencing and Model Selection

Sequencing of COI and COII yielded 1492 base pairs, with no length variants and 813 parsimony-informative in-group sites, after a short tRNA section was trimmed from the COII amplicon. The three EF-1 α segments produced widely varying sequence lengths across taxa due to occasional large insertions/deletions

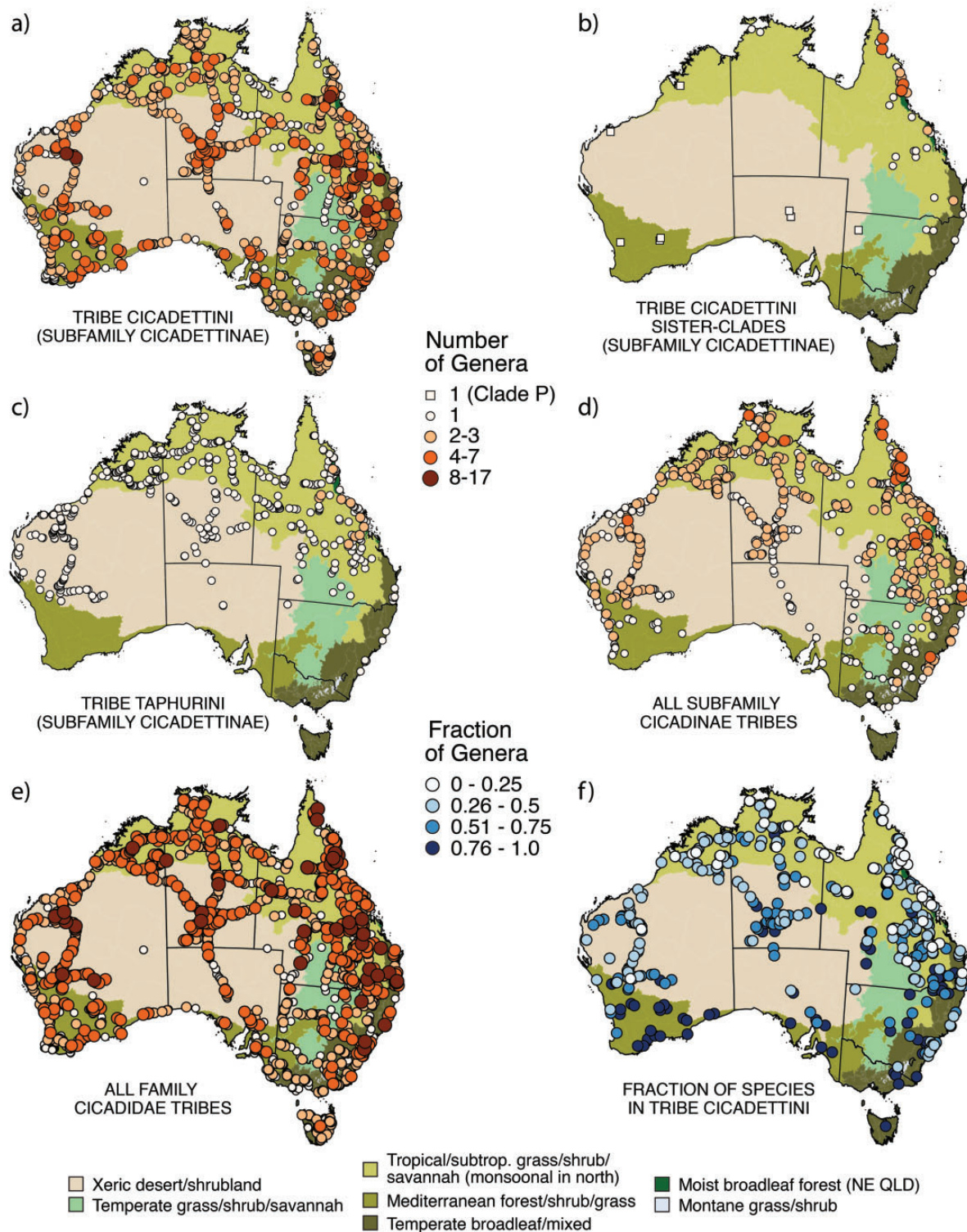


FIGURE 1. Geographic distribution of Australian cicada diversity sampled in this study, comparing habitat affinities of cicadas of the tribe Cicadettini to the sister-clades of the tribe (Clade P, then Prasiini + Chlorocystini, combined in map b) and other Australian cicada groups. Cicadas of the tribe Cicadettini disproportionately dominate arid and southern temperate regions. Family-level genomic data (unpublished) suggest that the clades in maps a+b, c, and d have arisen from three independent colonizations of Australia. Map e shows the total number of genera in all family Cicadidae tribes and map f show the fraction of cicada genera at each location belonging the tribe Cicadettini. Undescribed taxa are coded with likely future taxonomic assignments based on these studies. Two Australian tribes are not shown here: Cicadini (represented by *Diceropyga subapicalis*) reaches Cape York from Papua New Guinea, and Platypleurini (*Oxypleura calypso*) reaches Christmas Island from Indonesia. Inadequate rainfall prevented sampling of some arid-zone regions. Base map is from "Terrestrial Ecoregions in Australia" produced by Environmental Resources Information Network for Parks Australia.

(indels), with 1047 base pairs of coding sequence distributed across five exons. For the in-group, 1266 EF-1 α sites were parsimony-informative (252 in the exons). Seqstate produced a matrix of 1267 binary indel (insertion/deletion) characters from the aligned EF-1 α intron data. Genbank accession numbers are found in online Appendix 1.

For the MrBayes and ML analyses of the tribe-level data set, Partitionfinder selected a seven-partition scheme for the nucleotide component, with first, second, and third codon positions separately estimated for both the mtDNA and EF-1 α exon data (with COI and COII positions combined, except that the COII second-position sites were included in the first-position partition). GTR+I+G sequence-evolution models were selected for the three mtDNA partitions and for the EF-1 α first-position and intron sites. F81+I and HKY+G were selected for the EF-1 α second- and third-position partitions, respectively. Simpler models were occasionally selected for the smaller data sets. Tests of compositional heterogeneity were nonsignificant ($P = 1.0$) for all data partitions except the third codon position mtDNA data ($P < 10^{-8}$).

Phylogenetic Analyses (A–D)

Figure 2 shows the MrBayes combined-data tree for the tribe Cicadettini, midpoint-rooted along the branch to the out-groups. MrBayes trees are shown for mtDNA and EF-1 α separately in Supplementary Figures S2 and S3. A detailed examination of these trees (from analyses A–C) is given in the Supplementary Material, including findings of taxonomic relevance.

The first in-group split in all of the trees separates a large clade containing all of the described species of Cicadettini and many undescribed taxa from a small Australian clade (Clade P) containing two undescribed species plus *Chrysocada* and *Pictila* from the tribe Taphurini. (Most other genera of Taphurini are well separated from the clade studied in this article, as shown in unpublished preliminary family-level trees.) Following a long branch, the larger clade splits again asymmetrically, the smaller sister clade comprising a trans-Australian group containing *Samaecicada subolivacea* (currently placed in Cicadettini) from eastern Australia and an undescribed western species.

All trees show the remaining taxa forming a complex radiation with much shorter internal branches. Important features within this large “main radiation” (Fig. 2) are: (i) multiple large trans-Australian clades containing endemic Australian genera, one of the deepest being *Pauropsalta* and its allies (see Owen et al. 2015a, 2015b); (ii) two clades containing New Zealand species (AmN, for *Amphipsalta-Notopsalta*, and MRK), as previously reported from a much smaller data set (Arensburger et al. 2004); (iii) two clades containing New Caledonian species (the larger one named NCAL here), one including the larger of the two New Zealand

radiations; and (iv) two clades containing all of the non-Australasian species, here termed the R1 (radiation 1) and R2 clades. The R1 clade includes Asian and European genera, an African clade (AFR), and another New Caledonian lineage (*Melampsalta germaini*). The R2 clade contains the remainder of the non-Australian genera, including species from Asia, Europe, and North America (NA) but none from subsaharan Africa.

Divergence Time Analyses

Analysis stability.—In many of the BEAST analyses below, 1–3 of the 8–10 replicates were observed sampling a solution with inferior posterior probability (ca. 150 likelihood points) and inferior site likelihood (ca. 90 points) for the entire run (usually 60–90 million generations). In each case, the mitochondrial substitution parameters and tree height were the most affected, and most of the likelihood difference was traced to the COI partition. The remaining runs always sampled the same superior solution, which we report below in each case. Similar “cryptic failure” was observed by Marshall (2010) for MrBayes phylogenetic analyses containing mtDNA data of the same type as this study.

COI molecular-clock calibrated analyses (E, F, G).—The BEAST chronogram from analysis E, calibrated with a COI clock prior only, closely resembles the topology from the Bayesian combined-data phylogenetic analysis (Supplementary Fig. S4). However, the divergence time estimates from this chronogram (Fig. 3, analysis E) were unexpectedly old given previously published work. Geological calibration points and COI molecular clocks have been used to estimate ages of 15–11 Ma for the MRCA of the New Zealand *Maoricicada-Rhodopsalta-Kikihia* radiation (MRK clade—Buckley and Simon 2007; Marshall et al. 2008) and 12.5 Ma for *Amphipsalta-Notopsalta* (AmN clade—Marshall et al. 2012). The corresponding estimates from the tribe-level analysis were 39 and 32 Ma (Fig. 3, analysis E), despite the fact that our clock prior median was close to the Brower (1994) rate used by the earlier studies. The 95% confidence interval for the root of the analysis E tribe-level tree, which was 260–87 Ma (median 166 Ma), is also difficult to reconcile with a dated insect phylogeny (Misof et al. 2014) that suggests a confidence interval of 150–56 Ma for the divergence of Cicadidae–Cercopidae (Methods section).

Dates from an analysis (F) using the COI clock plus the available fossil calibrations were approximately 25% younger than those with the clock alone, but still unexpectedly old (Fig. 3). In contrast, applying the New Caledonia geological constraint together with the COI clock (analysis G) yielded clade ages about half of those found with the clock alone. With the geological constraint, the 95% confidence interval for the MRK node (25–14 Ma) overlaps some earlier COI-calibrated estimates.

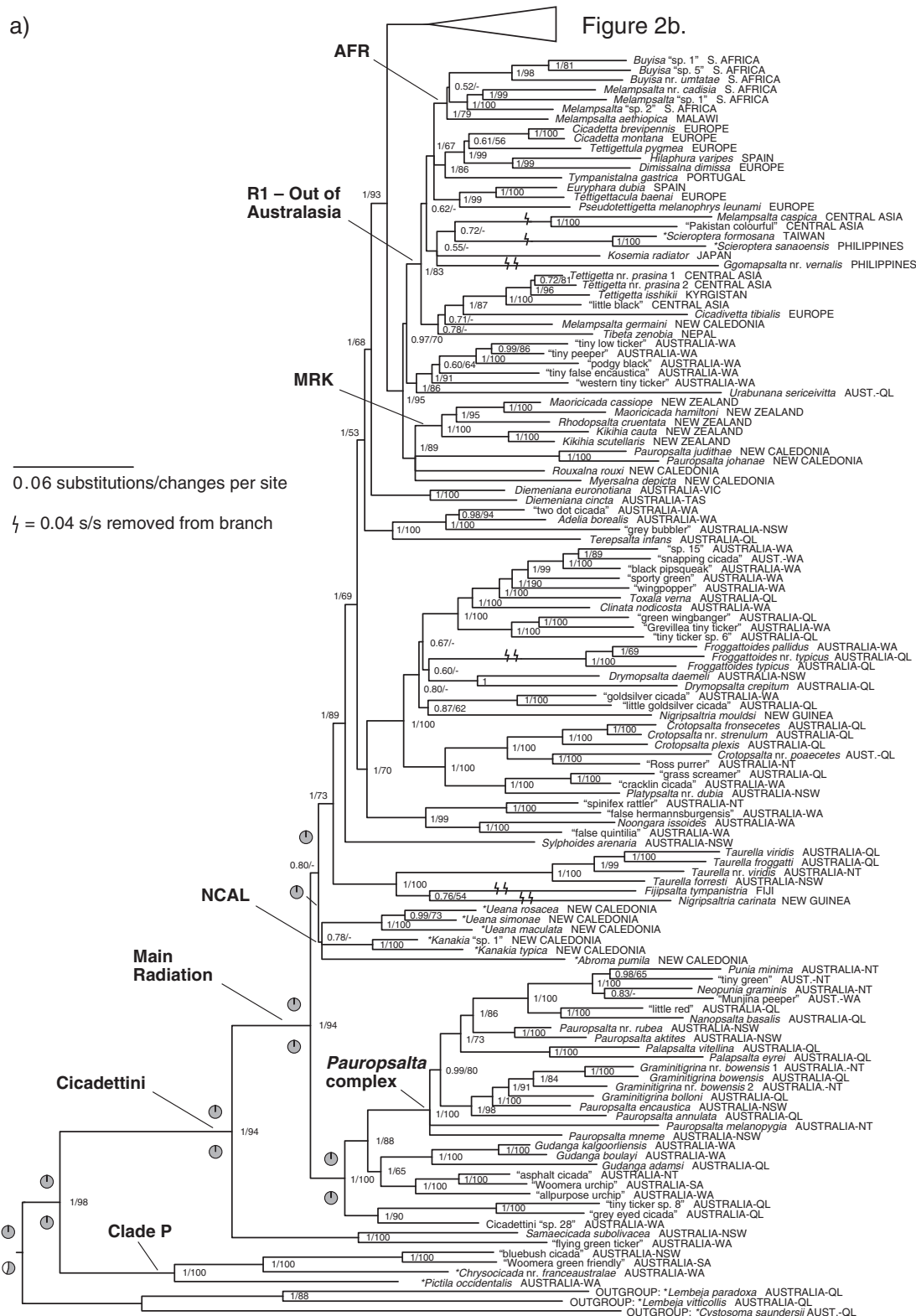


FIGURE 2. MrBayes phylogram for 203 species from the cicada tribe Cicadettini and close allies, from analysis A, based on partitioned mitochondrial and nuclear-gene data (exon and intron). a) Root and first section of tree. b) Distal section of tree (next page). Node supports are Bayesian posterior probabilities followed by ML bootstrap supports. Approximate collecting locations are given at the country level or, for Australian taxa, at the state/territory level (NT = Northern Territory, NS = New South Wales, QL = Queensland, SA = South Australia, TS = Tasmania, VI = Victoria, WA = Western Australia). Taxa with asterisks are currently classified in tribes other than Cicadettini (see online Appendix 1). Pie chart symbols indicate, for selected nodes, probabilities from Lagrange of each ancestral area occurring at each descendant branch (Australasia = light grey). AFR = African clade; AmN = New Zealand *Amphipsalta-Notopsalta* clade; MRK = New Zealand *Maoricicada-Rhodopsalta-Kikihia* clade; NCAL = largest New Caledonia clade; R1 and R2 = Out-of-Australasia radiations.

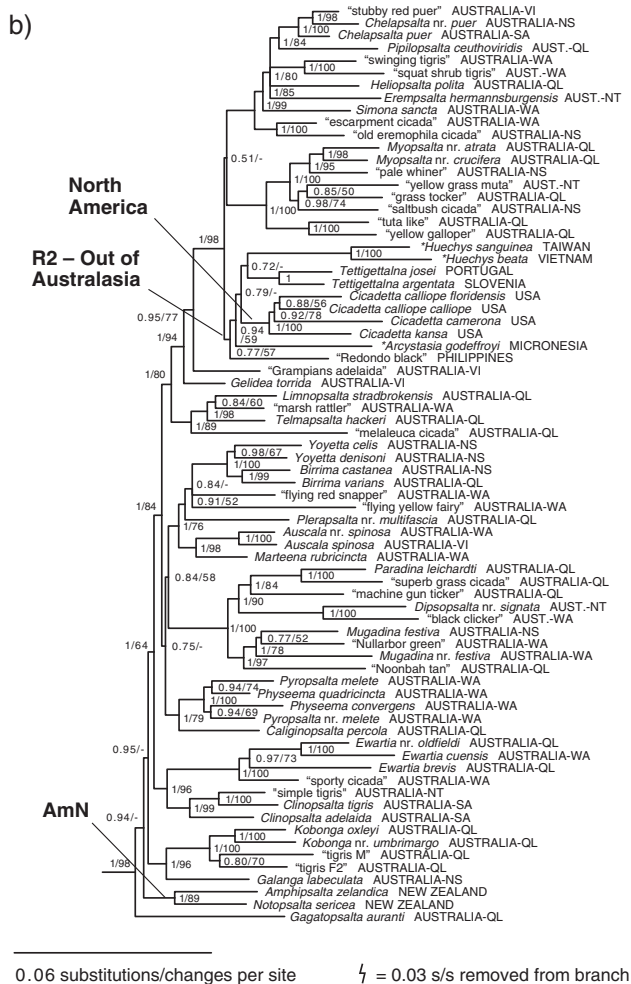


FIGURE 2. Continued.

Extrapolated EF-1 α clock analysis (H2).—The tribe-level BEAST analysis calibrated with the extrapolated EF-1 α intron and exon clocks (analysis H2) yielded younger dates that were similar to those observed when the COI clock was combined with the New Caledonia calibration (analysis G). The in-group and NCAL divergences were estimated at 61 Ma and 34 Ma, respectively, compared to 122 Ma and 68 Ma with only the COI clock.

ML optimization of the COI data partition on MRK- and tribe-level tree topologies confirmed that the tribe-level analysis estimated much more substitution for the MRK subclade even though far fewer representatives of the MRK clade were included in the tribe-level tree (Fig. 4). The mean root-to-tip distance for the MRK subclade was approximately 0.45 s/s when analyzed in the tribe-level context, but only about 0.12 s/s when analyzed in isolation.

Post hoc scaled analysis (J).—The BEAST Cicadettini chronogram from analysis E, when rescaled using the root age and 95% confidence interval of a MRK-only COI-clock calibrated analysis, yielded dates that were much younger (Fig. 3, analysis J) and similar to those of the

geologically calibrated and extrapolated EF-1 α analyses. A second example of this method was recently published in Hill et al. (2015).

Subclade analyses (K).—Perhaps due to the smaller taxon samples, several of the subclade analyses oscillated between alternative combinations of some mtDNA substitution model parameters. In these cases, HKY or TrN was substituted for GTR to obtain a stable result, as noted in Table 2.

Overall, COI-calibrated relaxed-clock analyses of subclades of progressively greater clade age showed that, for younger subclades, ages estimated in isolation tended to be younger than those found when the subclade was analyzed in the context of the whole tribe (Table 2a); the age of the youngest subclade was 25% younger. However, as subclade depth increased, the bias switched direction, with the deepest subclade age estimated to be 50% older in isolation. When the main radiation was subsampled to different degrees, the resulting ages were all somewhat older than those observed in the tribe-level analysis (approximately 85 Ma vs. 73 Ma) except for the smallest subsample (Table 2b). The effects of subsampling on inferred clade ages were reflected in the relative gene rates (Tables 2 and 3). COI/EF-1 α relative rates were more exaggerated in analyses of the deeper subclades. The relative rates (and relative tree lengths) of the EF-1 α exon and intron regions also changed across the subclade analyses, but modestly compared to the change in the estimated relative rates of mtDNA versus nuclear-gene data (Tables 2 and 3).

The relative tree-lengths for the COI and EF-1 α data subsets, estimated using ML to obtain the scaled BEAST priors for each analysis, are also shown in Table 3 and compared to the posterior BEAST relative rate estimates. The COI rate relative to EF-1 α changed much more across the various ML analyses than the relative rate of the two EF-1 α data subsets (Table 3), which remained very stable, and the pattern remained evident in the BEAST posterior rate estimates.

Among-site rate variation, substitution model, and tree lengths (L, M, N, P).—With COI modeled as one partition, the complexity of the gamma distribution substantially affected ML parameter estimates (including tree length) for the tribe-level data set (Supplementary Table S5). Log-likelihood scores improved up to the 20 categories tested, tree length increased by more than 50%, and large changes were observed in several model parameters. In contrast, little change was observed in the genus-level (MRK) analyses. When COI was partitioned by codon position, even greater effects were observed across the tribe-level analyses, while the MRK parameter estimates were again stable (Supplementary Table S6).

When the tribe-level COI-calibrated BEAST analysis (E) was modified to use empirical base frequencies instead of estimated frequencies (analysis N), the date estimates became dramatically younger (e.g., 39 Ma vs. 122 Ma for the in-group, Fig. 3). The estimated

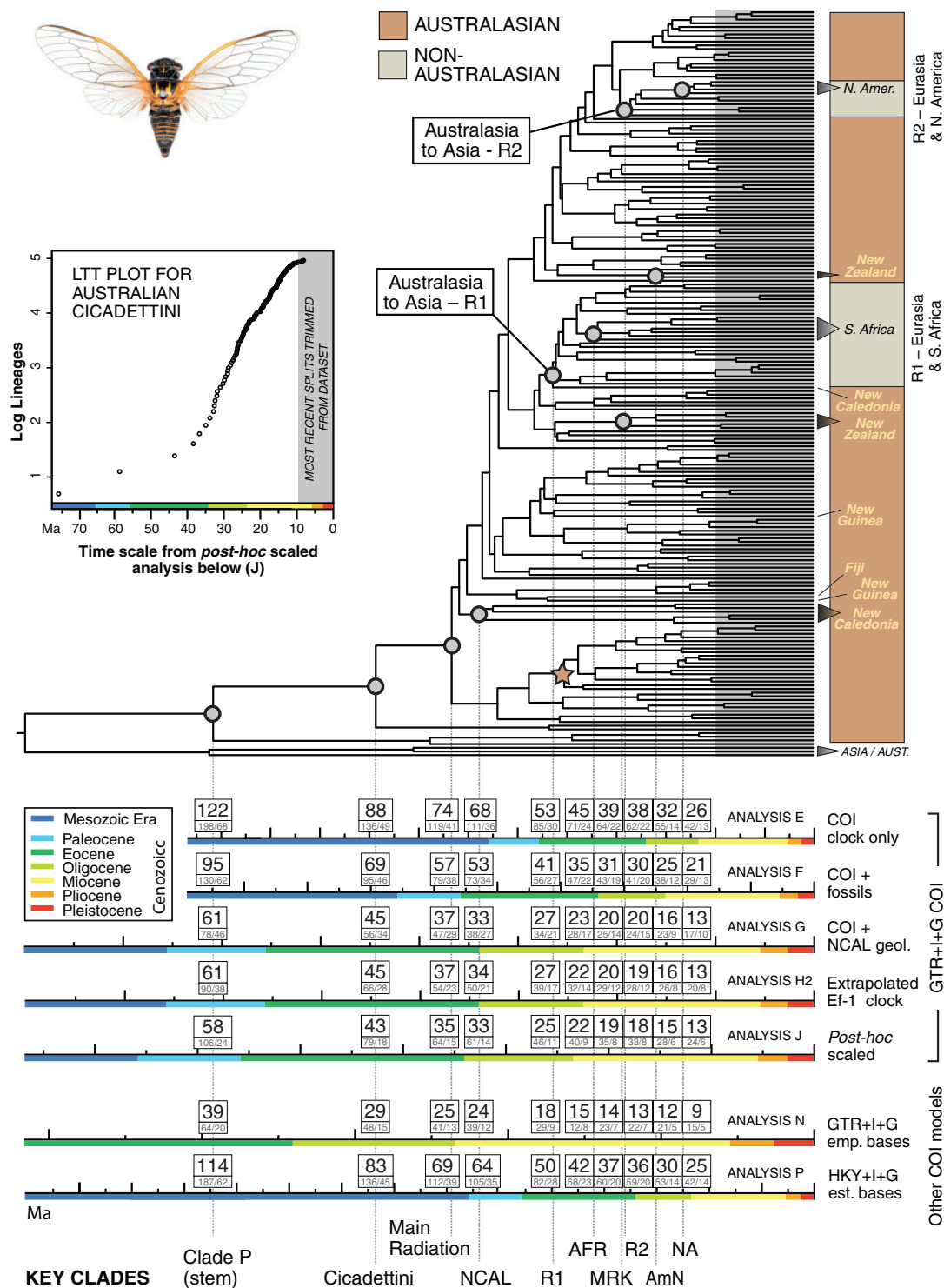


FIGURE 3. Relaxed-clock divergence time results for the cicada tribe Cicadettini and close relatives, showing median node ages and 95% confidence intervals for five different calibration methods and two substitution model variants, together with a representative chronogram from analysis E (all chronograms had nearly identical tree shapes). Analysis E was calibrated directly with a broad mtDNA COI prior encompassing literature estimates for arthropods. In analysis F, the COI prior was combined with the limited available fossil information. In analysis G, the COI prior was combined with a geological calibration on the MRCA of a New Caledonia clade (NCAL). Analysis H2 was calibrated with EF-1 α exon and intron rates that were extrapolated from the COI clock in a separate analysis of the New Zealand *Maoricicada*-*Rhodopsalta*-*Kikihia* (MRK) subclade. In analysis J, the Cicadettini chronogram was calibrated *post hoc* by scaling the age and 95% confidence interval of the MRK MRCA to the results of the MRK-only analysis. Analyses N and P demonstrate the effect of modifications of the COI substitution model on the results from E. The R1 and R2 clades contain all non-Australasian Cicadettini. The brown star labels the large *Pauropsalta* radiation which has been separately dated by Owen et al. (2015a). AFR = African clade; AmN = New Zealand *Amphipsalta*-*Notopsalta* clade; NA = North American clade. Inset: LTT = Lineage through time. Large ticks on the timescales denote 10 myr intervals. Inset photo = *Gagatopsalta auranti*.

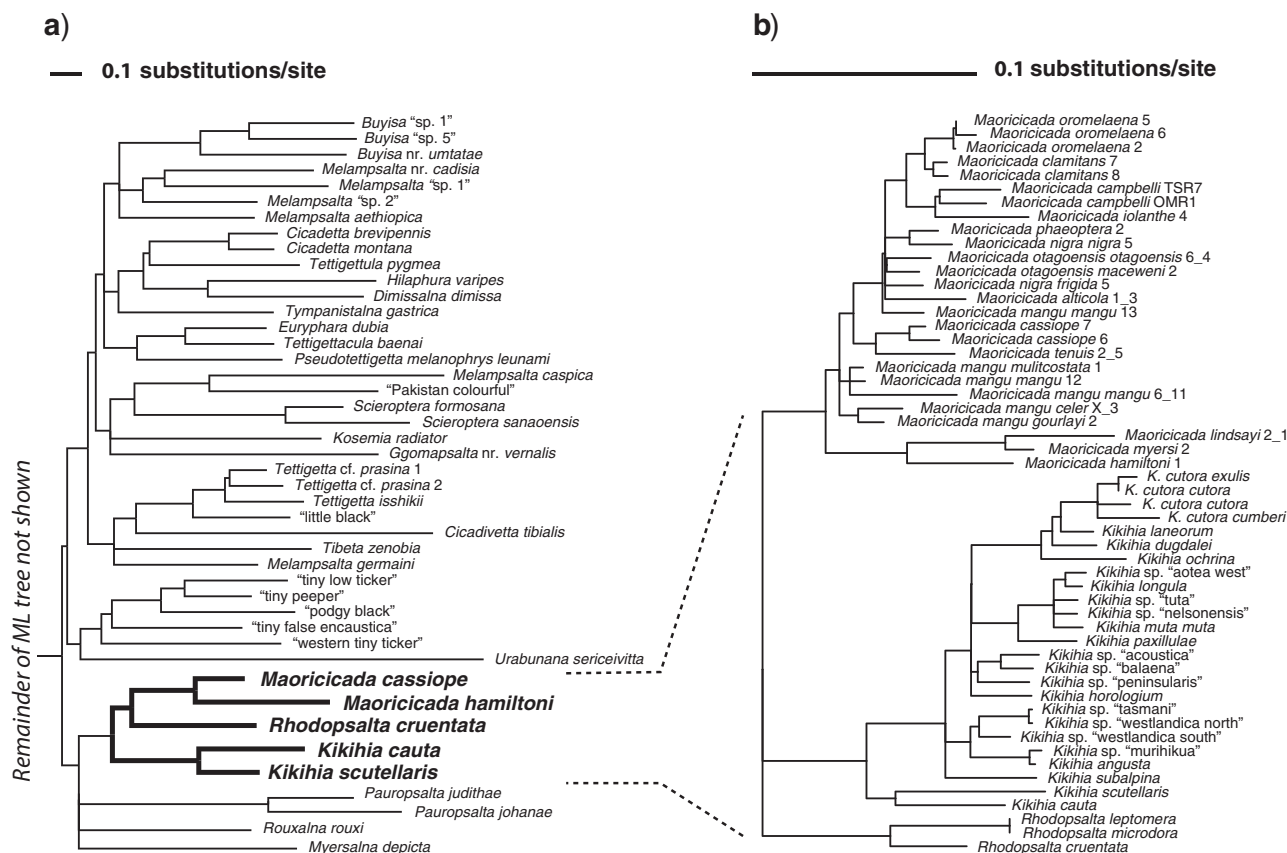


FIGURE 4. ML optimization of COI substitution (tree length) for the New Zealand *Maoricicada*-*Rhodopsalta*-*Kikihia* (MRK) subclade when analyzed (a) within the tribe-level data set and (b) as a separate subclade with more extensive taxon sampling. The greater amount of ML substitution estimated within the MRK clade in the tribe-level analysis with only five MRK taxa present, compared to the 54-taxon MRK-only analysis, is an unexpected outcome because increased taxon sampling usually increases the estimated number of substitutions (the “node-density artifact”, Fitch and Beintema 1990, Venditti et al. 2006).

COI base frequencies of analysis E ($p_A = 0.37$, $p_C = 0.055$, $p_G = 0.062$, $p_T = 0.51$) were much more AT-biased than the observed/empirical base frequencies ($p_A = 0.30$, $p_C = 0.13$, $p_G = 0.16$, $p_T = 0.41$). Reducing the number of parameters in the substitution matrix from six to two by using HKY instead of GTR had a comparatively minor effect on the inferred ages (Fig. 3, analysis P).

Diversification Rates

The lineage-through-time plot (Fig. 3 inset) showed a shallow initial diversification rate corresponding to the deepest half of the Cicadettini tree, followed by a sharp increase in rate that continues through the remainder of the plot (disregarding the most recent 10 myr, for which the splits were not sampled). There is a gradual slowdown apparent in the final third of the plot, beginning about 20–25 Ma if the graph is calibrated according to analysis J.

Biogeography

Lagrange recovered Australasia as the geographic range for all ancestors within the deep- and mid-level

in-group splits of the tree (with $P = 1.0$) (Fig. 2). Moving up the tree from the root, the first non-Australasian ancestors ($P < 0.95$) were: (i) the MRCA of the R1 clade plus its Australian sister clade and (ii) the MRCA immediately preceding the split containing the R2 clade. Detailed results are included in the Supplementary Material.

DISCUSSION

Taxon Sampling and Molecular Clocks

Relaxed-clock analyses conducted in BEAST for the cicada tribe Cicadettini, calibrated solely with an informed clock prior that spanned published mtDNA COI rate estimates (e.g., Brower 1994; Quek et al. 2004; Papadopoulou et al. 2010), returned unexpectedly old divergence times that conflicted with earlier COI-calibrated research on well-studied New Zealand genera of Cicadettini and results from a dated phylogeny of Insecta (see Results and Misof et al. 2014). In contrast, when the tribe-level phylogeny was dated by extrapolating COI-calibrated results from one of the New Zealand clades, or by way of a geological

TABLE 2. Effect of data set subsampling by subclade age and sample size on gene/partition substitution rate estimates in BEAST relaxed-clock analyses

a)	# of taxa	Analysis E age	Subclade Analysis age	COI Mean	COI Stdev	COII Mean	COII Stdev	EF-1 α exon Mean	EF-1 α exon Stdev	EF-1 α intron Mean	EF-1 α intron Stdev
Subclade											
MRK	54	39	20	0.0118	0.879	–	–	0.000470	0.397	0.00168	0.685
Subclade I	29	41	30	0.0115	0.288	0.0107	0.119	0.000384	0.255	0.00183	0.452
Subclade II	35	48	32	0.0115	0.369	0.0132	0.114	0.000372	0.356	0.00171	0.523
Subclade III ^a	38	57	58	0.0115	0.256	0.0132	0.326	0.000348	0.261	0.00170	0.370
Subclade IV ^a	34	65	86	0.0115	0.204	0.0109	0.295	0.000236	0.314	0.00094	0.329
Subclade V ^a	27	64	99	0.0123	0.337	0.0075	0.483	0.000127	0.221	0.00069	0.552
Tribe + out-group (Analysis E)	203	164	N/A	0.0114	0.331	0.0080	0.378	0.000240	0.445	0.00119	0.549
b)	# of taxa	Age E in analysis	Subsample analysis age	COI Mean	COI Stdev	COII Mean	COII Stdev	EF-1 α exon Mean	EF-1 α exon Stdev	EF-1 α intron Mean	EF-1 α intron Stdev
Subclade											
Main Radiation ^b	25	73	65	0.0117	0.139	0.0108	0.324	0.000313	0.186	0.00137	0.391
Main Radiation	42	73	88	0.0122	0.252	0.0091	0.357	0.000198	0.270	0.00090	0.339
Main Radiation ^a	70	73	84	0.0118	0.253	0.0094	0.363	0.000258	0.393	0.00114	0.505
Main Radiation	116	73	84	0.0118	0.323	0.0098	0.373	0.000266	0.345	0.00126	0.507
Main Radiation ^a	194	73	102	0.0123	0.327	0.0089	0.369	0.000257	0.438	0.00125	0.546

Notes: Subclade ages (Ma) and gene rates (substitutions/site/myr) from analysis E, the Cicadettini tribe-level data set calibrated solely with literature COI rates (Fig. 3), are given for reference. a) Results for monophyletic/paraphyletic subclades taken from the Cicadettini tree (for clade identities see Supplementary Fig. S1). For younger subclades, ages estimated in isolation tend to be younger than those found when the subclade is analyzed in the context of the whole tribe, while the pattern is reversed for older subclades. b) Results from subsampling the large “main radiation” clade with different sample sizes while retaining tree structure; subclade age varies erratically. COII sequences were not available for the MRK analysis (Methods section) of the New Zealand *Maoricicada-Rhodopsalta-Kikihia* Cicadettini subclade. Mean is *ucl.d.mean*, and Stdev is *ucl.d.stdev* from BEAST.

^a HKY used for COI and COII substitution models to achieve stable parameter estimates.

^b HKY used for COII substitution model.

TABLE 3. Ratios of mean gene rates from BEAST relaxed-clock analyses compared to tree-length ratios from ML phylogenetic analyses, for Cicadettini tree subclades

a)	ML tree-length ratios			BEAST rate ratios		
Clade	COI/Exon	COI/Intron	Intron/Exon	COI/Exon	COI/Intron	Intron/Exon
MRK (54 taxa)	25.197	5.660	3.568	25.1	7.02	3.57
Subclade I	48.266	11.032	4.375	30.0	6.28	4.77
Subclade II	42.048	10.184	4.129	30.9	6.73	4.60
Subclade III	45.577	9.219	4.944	33.1	6.76	4.89
Subclade IV	73.451	18.634	3.941	48.7	12.2	3.98
Subclade V	142.889	28.791	4.963	96.9	17.8	5.43
Tribe + out-group (Analysis E)	43.146	9.166	4.707	47.5	9.58	4.96
b)	ML tree-length ratios			BEAST rate ratios		
Clade	COI/Exon	COI/Intron	Intron/Exon	COI/Exon	COI/Intron	Intron/Exon
Main rad. 25 taxa	41.742	10.070	4.145	37.4	8.54	4.38
Main rad. 42 taxa	54.778	12.281	4.460	61.6	13.6	4.55
Main rad. 70 taxa	56.301	12.803	4.398	45.7	10.4	4.42
Main rad. 116 taxa	53.121	12.118	4.384	44.4	9.37	4.74
Main rad. 194 taxa	57.818	12.471	4.636	47.9	9.84	4.86

Notes: Results from analysis E, the Cicadettini tribe-level data set calibrated solely with literature COI rates (Fig. 3) are given for reference. a) Results for subclades of varying age (see Table 2 and Supplementary Fig. S1). Although some differences between the Bayesian and ML analyses are apparent, a common pattern of altered mtDNA rates is apparent in both. b) Results from subsampling the large “main radiation” clade with different taxon sets while retaining tree structure. In all cases, the relative intron/exon rate within EF-1 α is more stable than the relative rates of the mtDNA data versus the two EF-1 α partitions.

calibration, much younger ages were estimated (Fig. 3). The older trees obtained at the tribe level appear to be caused by inflated estimates of mtDNA COI substitution, observed as higher clock rates in relaxed-clock analyses and as longer branches in ML analyses (nuclear EF-1 α rate/branch estimation was little affected by comparison). Across all tribe-level analyses, the shape of the underlying chronogram was stable, meaning that younger and older node times were affected proportionally to their depth, a pattern different from that observed by Schwartz and Mueller (2010), who found changes in tree shape—i.e., inflation only in some branch lengths—in their ML-based simulation studies. The results also differ from earlier work showing that highly saturated sites can lead to underestimation of the amount of substitution for deeper branches, and therefore older ages for younger nodes when the root age is constrained (Phillips 2009; Brandley et al. 2011; Dornburg et al. 2014).

Separate analyses of subclades sampled from our tribe-level data set confirmed that divergence times became more inflated as subclade depth increased, again in association with increased COI rate estimates relative to those of the nuclear EF-1 α partitions (Tables 2 and 3). For younger subclades studied in isolation, we estimated divergence times that were younger than when these same subclades were included in the tribe-level analysis. However, the effect was reversed with the deepest subclades, for which older ages were found when analyzed in isolation, suggesting a complicated interaction between clade age and taxon sample size. These effects are caused by differences in the reconstructed mtDNA substitution process, rather than by Bayesian priors. ML analyses, which use no priors, showed the same inflation pattern when the New Zealand subclade was examined within the tribe-level tree versus in isolation (Fig. 4).

Owen et al. (2015a) attempted unsuccessfully to eliminate COI rate/date inflation in BEAST analyses of the *Pauropsalta* generic complex, a major subclade of our tree, by excluding the fastest-evolving sites (as in Dornburg et al. 2014) and by modeling each codon position separately. Lacking fossils and geological calibration points, they instead used the Brower (1994) COI rate to calibrate divergence times estimated using penalized likelihood (Sanderson 2002). The median age for the *Pauropsalta* MRCA in our analyses G, H2, and J (ca. 24 Ma—starred node in Fig. 3) falls on the young side of the range of 44–21 Ma estimated by Owen et al., while the *Pauropsalta* MRCA node age from our apparently biased COI-calibrated tribe-level analysis (51 Ma) was much older and outside their interval.

The approximate concordance of analyses G, H2, J (Fig. 3) and the *Pauropsalta* study, together with the greater taxon sampling of the MRK analyses, increases our confidence that the tribe-level dates obtained from COI calibration alone are too old. The younger dates place the beginning of the main radiation of Cicadettini between the early Eocene (56 Ma, for the slowest clock estimate) and the mid-Miocene (15 Ma,

fastest clock), and the remaining discussion follows this tentative conclusion. However, even the younger dates for the New Zealand radiations remain older than those of earlier studies. Furthermore, extreme shifts in tree length and rate/date reconstructions were observed under different substitution and among-site rate variation models (see Supplementary Tables S5 and S6 here and Schwartz and Mueller 2010), especially when empirical base frequencies were used (Fig. 3). The latter result revealed a much greater A-T bias for COI when base frequencies were estimated, suggesting that heterotachy may be an issue in this data set. If we had not reconstructed a New Caledonian clade at an intermediate depth in our tree, allowing us to exploit a geological calibration, our confidence would be considerably weakened.

Combining relaxed-clock Bayesian methods with empirical data on clock rates is appealing, especially for groups with few informative fossils. However, even a strong rate prior (i.e., a well-tested clock with a narrow confidence interval) is no guarantee of sensible results because so much uncertainty can exist in the estimate of the molecular substitution process. Decisions regarding the taxon sample, substitution model, partitioning scheme, and Bayesian priors can substantially influence the outcome (Brown et al. 2010; Schwartz and Mueller 2010; Andujar et al. 2012; Rannala et al. 2012), even when topology is well-supported (Marshall 2010). We have speculated on the underlying causes of these patterns in our data but the ultimate causes remain unknown; however, subsampling of taxa can help to unmask the problem. These findings are relevant for efforts to calibrate new clocks (e.g., Papadopoulou et al. 2010) using likelihood-based methods (including Bayesian), as well as divergence time studies. Sauquet et al. (2012), while analyzing the effects of calibration on divergence times, called for “increased background research ...at all stages of the calibration process”, a sentiment we strongly echo here.

Historical Biogeography and Diversification of Cicadettini

Global, out-of-Australasia radiations—With all non-Australasian genera clustering in two distal clades (R1 and R2, Fig. 2), biogeographic patterns support an Australasian ancestry for the tribe Cicadettini (Fig. 2). Few studies have identified globally distributed terrestrial animal groups with this pattern. However, fig wasps of the subfamily Sycophaginae exhibit a pattern that is remarkably similar to the Cicadettini case presented here in biogeography, timing, and number of dispersals (Cruaud et al. 2010), and in that case one lineage has also returned to Australasia. *Delias* butterflies originated in the Australian region following the breakup with Antarctica and spread globally in at least seven dispersal events (Braby and Pierce 2007). *Rhantus* diving beetles expanded throughout Eurasia from a New Guinean origin, but only after the ancestor

arrived from Asia ca. 9–7 Ma (Balke et al. 2009). Also, Kayaalp et al. (2013) suggest that *Hylaeus* bees originated in Australia in the Oligocene and subsequently spread globally.

The most well-known animal example of Australasian origin and subsequent dispersal is probably the songbirds, which have been examined both at the order level (Barker et al. 2004; Ericson et al. 2006; Jönsson and Fjeldså 2006) and separately within the bush-shrikes and their allies (Fuchs et al. 2012). The dates obtained by the Barker et al. study are interesting because they reconstructed a long series of exclusively Australian oscine Cenozoic splits between 65 and 50 Ma, when Australia was connected through a temperate Antarctica to South America. The tree was dated in multiple ways, one of which involved a cytochrome b (mtDNA) clock and ML distances for splits up to at least 35 Ma, so it could be useful to check for the possibility of inflated distances.

The rarity of out-of-Australasia cases may support the idea that movement of clades across major adaptive zones or biomes is a slow process (Chapple and Keogh 2004; Crisp et al. 2009), or it may reflect the relatively brief period of Australasian isolation after the final Gondwanan breakup. In our study, dispersal across continents following “escape” from Australasia was rapid for both the R1 and R2 clades of Cicadettini, which reached South Africa and North America only about 3–7 myr later in each case (Fig. 3). We can speculate that they never entered South America because they were blocked by tropical biomes or competition with resident cicadas of other tribes.

Only Australian species are known for the Cicadettini sister clade (Clade P), potentially pushing Australasian ancestry further into the past. Ongoing family-level studies will help to determine if an Australasian or late Gondwanan ancestry can be confidently concluded for deeper nodes. While the out-group tribes Chlorocytiini+Prasiini have the most species in southeast Asia, much of their generic-level diversity is found in the Australian wet tropics (Boer and Duffels 1996). However, geographically biased extinction and poor fossil records can create misleading biogeographic patterns (see discussions of the Casuarinaceae and the genus *Eucalyptus* in McLoughlin 2001; Swenson et al. 2001; Gandolfo et al. 2011).

Divergence timing of Cicadettini and key geological events.—The divergence time estimates we found for the tribe Cicadettini (mainly early Paleocene to the late Oligocene) include the early Oligocene date of the final Australia–Antarctica separation (ca. 33 Ma—Hall 2011), although our median estimates are considerably higher (ca. 45 Ma). The broad range of the empirical COI clock prior, necessary in this study because of the lack of information on cicada rates, weakens our ability to reject hypotheses. *Post hoc* examination of the confidence intervals suggests that the slower end of the clock prior could be less appropriate for Cicadettini. The confidence intervals for R1, the oldest out-of-Australasia clade, include a

period when Australia was still well isolated from Asia (analyses G, H2, and J, Fig. 3) (Hall 2011), which conflicts with the strong biogeographic patterns we found. The older end of the confidence interval for the NCAL node conflicts with the maximum-age assumption of 37 Ma for New Caledonia in analyses H2 and J, and the confidence intervals for the New Zealand MRK and AmN radiations allow for unlikely diversification preceding the Oligocene near-submergence of New Zealand (Knapp et al. 2007; Landis et al. 2008; Bunce et al. 2009; Biffin et al. 2010; Giribet and Boyer 2010; Lee et al. 2012). On the other end of the scale, the youngest end of the date range estimated for the North American Cicadettini clade (in analyses G, H2, and J) is younger than the latest date proposed for temperate species to have crossed Beringia from Asia (ca. 10 Ma—Sanmartín 2001). Our original question—whether the tribe Cicadettini originated after Australia became isolated—remains necessarily uncertain. Either a faster COI rate prior is correct, in which case Cicadettini diverged after the breakup, or—we feel less likely—diversification happened well before and the group failed to reach South America, went extinct there, or has been missed in sampling.

Cenozoic aridification and diversification of Cicadettini.—Most studies suggest that the shift to a modern arid-adapted flora and fauna accelerated in Australia ca. 35 Ma, near the end of the Eocene (Byrne et al. 2011; Crisp and Cook 2013), although there is evidence suggesting seasonally dry habitats much earlier in the record (Crisp et al. 2011). Monsoonal habitats, which also support many species of Cicadettini, developed and expanded during approximately the same period (Bowman et al. 2010).

Phylogenies of major Australian plant groups suggest radiations beginning or accelerating in the Oligocene or later, including *Eucalyptus* sensu stricto, Casuarinaceae, and *Acacia*, and in groups of Australian legumes, while some (e.g., the broader *Eucalyptus* clade and *Nothofagus*) began earlier (Ladiges et al. 2003; Crisp and Cook 2009; Sauquet et al. 2012; Crisp and Cook 2013). Species of Cicadettini are associated with many of these plant groups. Some studies of Australian arid-adapted animal groups have been conducted, with most focusing on mid-Miocene to Pleistocene-age arid-zone radiations (Chapple and Keogh 2004; Jennings and Edwards 2005; Rabosky et al. 2007; Sanders et al. 2008; Shoo et al. 2008). Some animal studies have found major arid-zone diversification events beginning in the Oligocene (e.g., Jennings et al. 2003; Kayaalp et al. 2013).

Our age estimates for the main radiation of Cicadettini (ca. 36 Ma, analyses G, H2, and J) are centered on the Eocene/Oligocene boundary. Thus, diversification either occurred entirely in concert with aridification (see also Owen et al. 2015a), or largely overlapped it, depending on the actual cicada COI rate. Most of the expansion of the R1 clade, the oldest out-of-Australasia group, overlaps estimates for the onset of aridification in Asia (17 Ma—Miao et al. 2012) and South Africa

(8 Ma—Sepulchre et al. 2006). The fact that Cicadettini are nearly absent from Australian rainforests, which dominated Australia in the early Eocene, is expected if their origin and diversification occurred after those habitats began to recede. In contrast, the Chlorocystini and Prasiini sister groups are rare in the arid, temperate, and monsoonal habitats of Australia, and diverse in the rainforests (Fig. 1).

These findings support a view of habitat change (Pena and Wahlberg 2008) and expansion (Mayhew 2007) as drivers of clade diversification. However, the mechanisms involved remain to be determined, and there is evidence that subtle historical and/or ecological factors have played important roles. The large Cicadettini main radiation was preceded by the divergence of two clades (the *Samaecicada* clade and Clade P) that are extremely species-poor despite also being trans-Australian in distribution and present in the temperate and arid zones. Like Cicadettini, these clades have no representatives known from the tropical or southeastern temperate rainforests where the out-group taxa are found (Fig. 1b,c). The survival of these lineages without comparable diversification is an intriguing puzzle since their geographic ranges and habitat associations seem to exclude niche specialization, a factor that can limit diversification (Mayhew 2007). This situation is similar to that observed in New Zealand, where two clades of Cicadettini (AmN and MRK, Fig. 2) have diversified, in concert, to strikingly different degrees, during a period of dramatic geological and climatic changes (Marshall et al. 2012).

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.5590q>.

FUNDING

This work was supported by the National Science Foundation (NSF) [grants DEB-0955849 and DEB-0720664 to C.S.], the New Zealand Marsden Fund, and the University of Connecticut. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

ACKNOWLEDGMENTS

We are indebted to many collectors who provided specimens from around the world: Z. Ahmed, P. Arensberger, D. Britton, T. Britton, T. Buckley, J.-H. Chen, S. Chiswell, S. Cowan, S. Crandall, J. Cryan, A. de Boer, C. Dietrich, F. Douglas, H. Duffels, J. Dugdale, A. Emmott, K. Emmott, T. Ewart, D. Emery, R. Futahashi, G. Gibbs, M. Gogala, K. Green, B. Haywood, P. Hill, B. Hill, S. Hill, R. Hoare, B. Holland, M. Humphrey, L. Jingke, J. Jones, A. Kapla, H. Karuba, D. Lane, M. Lark, Y.-J.

Lee, Z. Lei, C. Mille, D. Mohagan, J. Moss, B. Moulds, T. Moulds, J. Olive, C. Owen, R. Perry, L. Popple, B. Pinchen, L. Popple, B. Price, S. Puissant, J. Quartau, M. Morgan-Richards, M. Robertson, A. Sanborn, M. Schouten, D. Sikes, J. Sueur, H. T. Pham, K. Trilar, T. Trilar, J. Urban, R. Veal, M. Villet, M. Whiles. Remco Bouckaert provided critical advice on the BEAST analyses. Andy Anderson, Karl Kjer, Pete Lockhart, Chris Owen, Tony Ewart, Alan Sanborn, Lindsay Popple, Dan Mozgai, and an anonymous reviewer provided ideas and/or criticism of the article.

Computational resources were provided by the Bioinformatics Cluster at the University of Connecticut and the CIPRES Science Gateway. Permits were obtained from the Australian federal and state governments, the Philippines Protected Areas and Wildlife Management Bureau, the governments of China, Vietnam, South Africa, and the NZ Department of Conservation.

REFERENCES

- Andujar C., Serrano J., Gomez-Zurita J. 2012. Winding up the molecular clock in the genus *Carabus* (Coleoptera: Carabidae): assessment of methodological decisions on rate and node age estimation. *BMC Evol. Biol.* 12:40.
- 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. *J. Biogeogr.* 31:557–569.
- Balke M., Ribera I., Hendrich L., Miller M.A., Sagata K., Posman A., Vogler A.P., Meier R. 2009. New Guinea highland origin of a widespread arthropod supertramp. *Proc. Roy. Soc. Lond. B Biol. Sci.* 276:2359–2367.
- Barker F.K., Cibois A., Schikler P., Feinstein J., Cracraft J. 2004. Phylogeny and diversification of the largest avian radiation. *Proc. Natl. Acad. Sci. USA* 101:11040–11045.
- Battistuzzi F.U., Filipski A., Hedges S.B., Kumar S. 2010. Performance of relaxed-clock methods in estimating evolutionary divergence times and their credibility intervals. *Mol. Biol. Evol.* 27:1289–1300.
- Bess E.C., Catanach T.A., Johnson K.P. 2014. The importance of molecular dating analyses for inferring Hawaiian biogeographical history: a case study with bark lice (Psocidae: *Ptycta*). *J. Biogeogr.* 41:158–167.
- Biffin E., Hill R.S., Lowe A.J. 2010. Did Kauri (Agathis: Araucariaceae) really survive the Oligocene drowning of New Zealand? *Syst. Biol.* 59:594–602.
- Boer A.J.D., Duffels J.P. 1996. Historical biogeography of the cicadas of Wallacea, New Guinea and the West Pacific: a geotectonic explanation. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 124:153–177.
- Bonacum J.P., O'Grady M., DeSalle R. 2005. Phylogeny and age of diversification for the *planitibia* species group of the Hawaiian *Drosophila*. *Mol. Phylogenet. Evol.* 37:73–82.
- Bouckaert R., Heled J., Kuehnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comp. Biol.* 10:e1003537.
- Boulard M., Nel A. 1990. Sur deux cigales fossiles des terrains tertiaires de la France (Homoptera, Cicadoidea). *Revue Fr. Ent.* 12:37–45.
- Bowman D.M.J.S., Brown G.K., Braby M.F., Brown J.R., Cook L.G., Crisp M.D., Ford F., Haberle S., Hughes J., Isagi Y., Joseph L., McBride J., Nelson G., Ladiges P.Y. 2010. Biogeography of the Australian monsoon tropics. *J. Biogeogr.* 37:201–216.
- Braby M.F., Pierce N.E. 2007. Systematics, biogeography and diversification of the Indo-Australian genus *Delias* Hübner (Lepidoptera: Pieridae): phylogenetic evidence supports an 'out-of-Australia' origin. *Syst. Entomol.* 32:2–25.

- Brandley M.C., Wang Y., Guo X., de Oca A.N., Fera-Ortiz M., Hikida T., Ota H. 2011. Accommodating heterogeneous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon* (Eumeces) lizards. *Syst. Biol.* 60:3–15.
- Bromham L., Penny D. 2003. The modern molecular clock. *Nat. Rev. Genet.* 4:216–224.
- Brower A.V.Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. USA* 91:6491–6495.
- Brown J., Hedtke S.M., Lemmon A.R., Lemmon E.M. 2010. When trees grow too long: investigating the causes of highly inaccurate Bayesian branch-length estimates. *Syst. Biol.* 59:145–161.
- 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. *Biol. J. Linn. Soc.* 91:419–435.
- Bunce M., Worthy T.H., Phillips M.J., Holdaway R.N., Willerslev E., Haile J., Shapiro B., Scofield R.P. 2009. The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. *Proc. Natl. Acad. Sci. USA* 106:20646–20651.
- 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. *J. Biogeogr.* 38:1635–1656.
- Byrne M., Yeates D.K., Joseph L., Kearney M., Bowler J., Williams M.A., Cooper S., Donnellan S.C., Keogh J.S., Leys R., Melville J., Murphy D.J., Porch N., Wyrwoll K.H. 2008. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Mol. Ecol.* 17:4398–4417.
- Campbell M.A., Van Leuven J.T., Meister R.C., Carey K.M., Simon C., McCutcheon J.P. 2015. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont *Hodgkinia*. *Proc. Natl. Acad. Sci. USA* 112:10192–10199.
- Chapple D.G., Keogh J.S. 2004. Parallel adaptive radiations in arid and temperate Australia: molecular phylogeography and systematics of the *Egernia whitii* (Lacertilia: Scincidae) species group. *Biol. J. Linn. Soc.* 83:157–173.
- Christin P.A., Spriggs E., Osborne C.P., Stromberg C.A., Salamin N., Edwards E.J. 2014. Molecular dating, evolutionary rates, and the age of the grasses. *Syst. Biol.* 63:153–165.
- Crisp M., Cook L., Steane D. 2004. Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 359:1551–1571.
- Crisp M.D., Arroyo M.T.K., Cook L.G., Gandolfo M.G., Jordan G.J., McGlone M.S., Weston P.H., Westoby M., Wilf P., Linder H.P. 2009. Phylogenetic biome conservatism on a global scale. *Nature* 458:754–756.
- Crisp M.D., Burrows G.E., Cook L.G., Thornhill A.H., Bowman D.M. 2011. Flammable biomes dominated by eucalypts originated at the Cretaceous-Palaeogene boundary. *Nat. Comm.* 2:193.
- Crisp M.D., Cook L.G. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. *Evolution* 63:2257–2265.
- Crisp M.D., Cook L.G. 2013. How was the Australian flora assembled over the last 65 million years? A molecular phylogenetic perspective. *Annu. Rev. Ecol. Evol. Syst.* 44:303–324.
- Cruaud A., Jabbour-Zahab R., Genson G., Couloux A., Yan-Qiong P., Da Rong Y., Ubaidillah R., Pereira R.A.S., Kjellberg F., van Noort S., Kerdelhué C., Rasplus J.-Y. 2010. Out of Australia and back again: the world-wide historical biogeography of non-pollinating fig wasps (Hymenoptera: Sycophaginae). *J. Biogeogr.* 38:209–225.
- Cruaud A., Jabbour-Zahab R., Genson G., Ungricht S., Rasplus J.Y. 2012. Testing the emergence of New Caledonia: fig wasp mutualism as a case study and a review of evidence. *PLoS One* 7:e30941.
- De Jong R. 2001. Faunal exchange between Asia and Australia in the Tertiary as evidenced by recent butterflies. In: Metcalfe I, Smith JMB, Morwood M, Davidson I editors. *Floral and faunal migrations and evolution in SE Asia-Australasia*. Lisse, A. A. Balkema, p. 133–146.
- Dornburg A., Townsend J.P., Friedman M., Near T.J. 2014. Phylogenetic informativeness reconciles ray-finned fish molecular divergence times. *BMC Evol. Biol.* 14:169.
- dos Reis M., Zhu T., Yang Z. 2014. The impact of the rate prior on Bayesian estimation of divergence times with multiple loci. *Syst. Biol.* 63:555–565.
- Drummond A.J., Ho S.Y., Phillips M.J., Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:e88.
- Drummond A.J., Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7:214.
- Ericson P.G., Anderson C.L., Britton T., Elzanowski A., Johansson U.S., Kallersjö M., Ohlson J.I., Parsons T.J., Zuccon D., Mayr G. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. *Biol. Lett.* 2:543–547.
- Espeland M., Muriene J. 2011. Diversity dynamics in New Caledonia: towards the end of the museum model? *BMC Evol. Biol.* 11:254.
- Ewart A. 2005. New genera and species of small ticking and 'chirping' cicadas (Hemiptera: Cicadoidea: Cicadidae) from Queensland, with description of their songs. *Mem. Queensl. Mus.* 51:439–500.
- Ewart A., Marques D. 2008. A new genus of grass cicadas (Hemiptera: Cicadoidea: Cicadidae) from Queensland, with descriptions of their songs. *Mem. Queensl. Mus.* 52:149–202.
- Ewart A., Popple L.W., Marshall D.C. 2015. New species of *Simona* Moulds 2012 and *Chelapsalta* Moulds, 2012 cicadas (Cicadidae: Cicadettinae: Cicadettini) from Australia: comparative morphology, songs, behaviour and distributions. *Zootaxa* 4001:1–65.
- Fitch W.M., Beintema J.J. 1990. Correcting parsimonious trees for unseen nucleotide substitutions: the effect of dense branching as exemplified by ribonuclease. *Mol. Biol. Evol.* 7:438–443.
- Fleischer R.C., McIntosh C.E., Tarr C.L. 1998. Evolution on a volcanic conveyor belt: Using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. *Mol. Ecol.* 7:533–545.
- Fuchs J., Irestedt M., Fjeldsa J., Couloux A., Pasquet E., Bowie R.C. 2012. Molecular phylogeny of African bush-shrikes and allies: tracing the biogeographic history of an explosive radiation of corvid birds. *Mol. Phylogenet. Evol.* 64:93–105.
- Fujioka T., Chappell J. 2010. History of Australian aridity: chronology in the evolution of arid landscapes. In: Bishop P., Pillans B. editors. *Australian Landscapes*. London, Geological Society, p. 121–139.
- Gandolfo M.A., Hermesen E.J., Zamaloa M.C., Nixon K.C., Gonzalez C.C., Wilf P., Cuneo N.R., Johnson K.R. 2011. Oldest known *Eucalyptus* macrofossils are from South America. *PLoS One* 6: e21084.
- Gillooly J.F., Allen A.P., West G.B., Brown J.H. 2005. The rate of DNA evolution: effects of body size and temperature on the molecular clock. *Proc. Natl. Acad. Sci. USA* 102:140–145.
- Giribet G., Boyer S.L. 2010. 'Moa's Ark' or 'Goodbye Gondwana': is the origin of New Zealand's terrestrial invertebrate fauna ancient, recent, or both? *Invert. Syst.* 24:1–8.
- Gogala M., Drosopoulos S., Trilar T. 2008. *Cicadetta montana* complex (Hemiptera, Cicadidae) in Greece – a new species and new records based on bioacoustics. *Dtsch. Entomol. Z.* 55:91–100.
- Gogala M., Drosopoulos S., Trilar T. 2012. *Cicadivetta goumenissa*, a new cicada species from Peloponnesos, Greece (Hemiptera: Cicadidae). *Acta Entomol. Slov.* 20:5–16.
- Grandcolas P., Muriene J., Robillard T., Desutter-Grandcolas L., Jourdan H., Guilbert E., Deharveng L. 2008. New Caledonia: a very old Darwinian island? *Phil. Trans. R. Soc. Lond. B Biol. Sci.* 363:3309–3317.
- Hall R. 2001. Cenozoic reconstructions of SE Asia and the SW Pacific: changing patterns of land and sea. In: Metcalfe I, Smith JMB, Morwood M, Davidson I editors. *Floral and faunal migrations and evolution in SE Asia-Australasia*. Lisse, A. A. Balkema, p. 35–56.
- Hall R. 2011. Australia-SE Asia collision: plate tectonics and crustal flow. In: Hall R, Cottam MA, Wilson MEJ editors. *The SE Asian Gateway: History and Tectonics of the Australia-Asia Collision*. London, Geological Society, p. 75–109.
- Hertach T., Trilar T., Wade E.J., Simon C., Nagel P. 2015. Songs, genetics and morphology: revealing the taxonomic units in the European *Cicadetta cerdaniensis* cicada group with description of new taxa (Hemiptera: Cicadidae). *Zool. J. Linn. Soc.* 173:320–351.

- 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* 3985:219–251.
- 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*. *J. Biogeogr.* 36:675–692.
- Hill R.S. 1994. History of the Australian vegetation: cretaceous to recent. Cambridge, UK: Cambridge University Press.
- Ho S.Y., Larson G. 2006. Molecular clocks: when times are a-changin'. *Trends Genet.* 22:79–83.
- Ho S.Y.W., Lo N. 2013. The insect molecular clock. *Aust. J. Entomol.* 52:101–105.
- Ho S.Y.W., Phillips M.J., Cooper A., Drummond A.J. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22:1561–1568.
- Jennings W.B., Edwards S.V. 2005. Speciation history of Australian grass finches (*Poephila*) inferred from thirty gene trees. *Evolution* 59:2033–2047.
- Jennings W.B., Pianka E.R., Donnellan S. 2003. Systematics of the lizard family Pygopodidae with implications for the diversification of Australian temperate biotas. *Syst. Biol.* 52:757–780.
- Jönsson K.A., Fjeldså J. 2006. A phylogenetic supertree of oscine passerine birds (Aves: Passeri). *Zool. Scripta* 35:149–186.
- Jordan S., Simon C., Polhemus D.A. 2003. Molecular systematics and adaptive radiation of Hawaii's endemic damselfly genus *Megalagrion*. *Syst. Biol.* 52:89–109.
- Kayaalp P., Schwarz M.P., Stevens M.I. 2013. Rapid diversification in Australia and two dispersals out of Australia in the globally distributed bee genus, *Hylaeus* (Colletidae: Hylaeinae). *Mol. Phylogenet. Evol.* 66:668–678.
- Knapp M., Mudaliar R., Wagstaff S.J., Lockhart P.J. 2007. The drowning of New Zealand and the problem of *Agathis*. *Syst. Biol.* 56:862–870.
- Knowlton N., Weight L.A., Solorzano L.A., Mills D.K., Bermingham E. 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science* 260:1629–1632.
- Ladiges P.Y., Udovicic F., Nelmes G. 2003. Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. *J. Biogeogr.* 30:989–998.
- Landis C.A., Campbell H.J., Begg J.G., Mildenhall D.C., Paterson A.M., Treweek S.A. 2008. The Waipounamu erosion surface: questioning the antiquity of the New Zealand land surface and terrestrial fauna and flora. *Geol. Mag.* 145:173–197.
- Lanfear R., Calcott B., Ho S.Y.W., Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29:1695–1701.
- Lanfear R., Thomas J.A., Welch J.J., Brey T., Bromham L. 2007. Metabolic rate does not calibrate the molecular clock. *Proc. Natl. Acad. Sci. USA* 104:15388–15393.
- Lee D.E., Conran J.G., Lindqvist J.K., Bannister J.M., Mildenhall D.C. 2012. New Zealand Eocene, Oligocene and Miocene macrofossil and pollen records and modern plant distributions in the Southern Hemisphere. *Bot. Rev.* 78:235–260.
- Lepage T., Bryant D., Philippe H., Lartillot N. 2007. A general comparison of relaxed molecular clock models. *Mol. Biol. Evol.* 24:2669–2680.
- Lewis P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* 50:913–925.
- Marshall D.C., Simon C., Buckley T. 2006. Accurate branch length estimation in partitioned Bayesian analyses requires accommodation of among-partition rate variation and attention to branch length priors. *Syst. Biol.* 55:993–1003.
- Marshall D.C. 2010. Cryptic failure of partitioned Bayesian phylogenetic analyses: lost in the land of long trees. *Syst. Biol.* 59:108–117.
- Marshall D.C., Hill K.B.R., Cooley J.R., Simon C. 2011. Hybridization, mitochondrial DNA taxonomy, and prediction of the early stages of reproductive isolation: Lessons from New Zealand cicadas of the genus *Kikihia*. *Syst. Biol.* 60:482–502.
- 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 Evol. Biol.* 12:177.
- Marshall D.C., Slon K., Cooley J.R., Hill K.B., Simon C. 2008. Steady Plio-Pleistocene diversification and a 2-million-year sympatry threshold in a New Zealand cicada radiation. *Mol. Phylogenet. Evol.* 48:1054–1066.
- Mayhew P.J. 2007. Why are there so many insect species? Perspectives from fossils and phylogenies. *Biol. Rev. Camb. Phil. Soc.* 82:425–454.
- McLoughlin S. 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Aust. J. Bot.* 49:271–300.
- Miao Y., Herrmann M., Wu F., Yan X., Yang S. 2012. What controlled mid-late Miocene long-term aridification in central Asia? — global cooling or Tibetan Plateau uplift: a review. *Earth-Sci. Rev.* 112: 155–172.
- Misof B., Liu S., Meusemann K., Peters R.S., Donath A., Mayer C., Frandsen P.B., Ware J., Flouri T.S., Beutel R.G., Niehuis O., Petersen M., Izquierdo-Carrasco F., Wappler T., Rust J., Aberer A.J., Aspöck U., Aspöck H., Bartel D., et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346:763–767.
- Moulds M.S. 1990. Australian cicadas. Kensington, NSW/Australia: New South Wales University Press.
- Moulds M.S. 2012. A review of the genera of Australian cicadas (Hemiptera: Cicadoidea). *Zootaxa* 3287:1–262.
- Nattier R., Robillard T., Desutter-Grandcolas L., Couloux A., Grandcolas P. 2011. Older than New Caledonia emergence? A molecular phylogenetic study of the eneopterine crickets (Orthoptera: Grylloidea). *J. Biogeogr.* 38:2195–2209.
- Nelson B.J., Andersen J.J., Brown J.M. 2015. Deflating trees: improving Bayesian branch-length estimates using informed priors. *Syst. Biol.* 64:441–447.
- Obbard D.J., Maclennan J., Kim K.W., Rambaut A., O'Grady P.M., Jiggins F.M. 2012. Estimating divergence dates and substitution rates in the *Drosophila* phylogeny. *Mol. Biol. Evol.* 29:3459–3473.
- Otte D., Alexander R.D. 1983. The Australian Crickets (Orthoptera: Gryllidae). Academy of natural sciences of Philadelphia monograph 22. Lawrence (KS): Allen Press.
- Owen C.L., Marshall D.C., Hill K.B.R., Simon C. Forthcoming 2015a. How the aridification of Australia structured the biogeography and influenced the diversification of a large lineage of Australian cicadas. *Syst. Biol.*
- Owen C.L., Marshall D.C., Hill K.B.R., Simon C. 2015b. The phylogenetic utility of acetyltransferase (ARD1) and glutamyl tRNA synthetase (QtRNA) for reconstructing Cenozoic relationships as exemplified by the large Australian cicada *Pauropsalta* generic complex. *Mol. Phylogenet. Evol.* 83:258–277.
- Owen C.L., Moulds M.S. Forthcoming 2016. Systematics and phylogeny of the Australian cicada genus *Pauropsalta* Goding and Froggatt and allied genera (Hemiptera: Cicadidae: Cicadettini). *Rec. Aust. Mus.*
- Papadopoulou A., Anastasiou I., Vogler A.P. 2010. Revisiting the insect molecular clock: The mid-Aegean trench calibration. *Mol. Biol. Evol.* 27:1659–1672.
- Paradis E., Claude J., Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
- Pena C., Wahlberg N. 2008. Prehistorical climate change increased diversification of a group of butterflies. *Biol. Lett.* 4:274–278.
- Phillips M.J. 2009. Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. *Gene* 441: 132–140.
- Pillon Y. 2012. Time and tempo of diversification in the flora of New Caledonia. *Bot. J. Linn. Soc.* 170:288–298.
- Popple L.W. 2013. A revision of the *Pauropsalta annulata* Goding & Froggatt species group (Hemiptera: Cicadidae) based on morphology, calling songs and ecology, with investigations into calling song structure, molecular phylogenetic relationships and a case of hybridisation between two subspecies. *Zootaxa* 3730: 1–102.
- Puissant S., Sueur J. 2010. A hotspot for Mediterranean cicadas (Insecta: Hemiptera: Cicadidae): new genera, species and songs from southern Spain. *Syst. Biodivers.* 8:555–574.
- PythonSoftwareFoundation. 2010. Python version 2.7, <http://www.python.org/psf/> (Last accessed 9 October 2015).
- Quek S.-P., Davies S.J., Itino T., Pierce N.E. 2004. Codiversification in an ant-plant mutualism: stem texture and the evolution of host use in

- crematogaster (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). *Evolution* 58:554–570.
- Rabosky D.L. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. *Evol. Bioinfo. Online* 2006:257–260.
- Rabosky D.L., Donnellan S.C., Talaba A.L., Lovette I.J. 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proc. Biol. Sci. Lond. B Biol. Sci.* 274:2915–2923.
- Rambaut A., Drummond A.J. 2007. Tracer v1.5, Available from <http://evolve.zoo.ox.ac.uk/> (Last accessed 9 October 2015).
- Rannala B., Zhu T., Yang Z. 2012. Tail paradox, partial identifiability, and influential priors in Bayesian branch length inference. *Mol. Biol. Evol.* 29:325–335.
- R Development Core Team. 2011. R: a language and environment for statistical computing. <http://www.R-project.org>. R Foundation for Statistical Computing, Vienna, Austria.
- Ree R.H. 2013. Lagrange configurator v. 20130526. Available at: <http://www.reelab.net/lagrange/configurator/index>.
- Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57:4–14.
- Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–542.
- Sanborn A.F. 2009. A new species of *Cicadetta* (Hemiptera: Cicadomorpha: Cicadidae) from Hispaniola. *Carib. J. Sci.* 45:1–7.
- Sanborn A.F. 2013. Catalogue of the Cicadoidea (Hemiptera: Auchenorrhyncha). With contributions to the bibliography by Martin H. Villet. San Diego: Elsevier. Inc., Academic Press.
- Sanborn A.F., Heath M.S. 2012. The cicadas (Hemiptera: Cicadoidea: Cicadidae) of North America north of Mexico. Lanham (MD): Entomological Society of America.
- Sanborn A.F., Phillips P. 2013. Biogeography of the cicadas (Hemiptera: Cicadidae) of North America, north of Mexico. *Diversity* 5: 166–239.
- Sanders K.L., Lee M.S., Leys R., Foster R., Keogh J.S. 2008. Molecular phylogeny and divergence dates for Australasian elapids and sea snakes (Hydrophiinae): evidence from seven genes for rapid evolutionary radiations. *J. Evol. Biol.* 21:682–695.
- Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Mol. Biol. Evol.* 19:101–109.
- Sanmartín I. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. *Biol. J. Linn. Soc.* 73:345–390.
- Sauquet H., Ho S.-Y., Gandolfo M.A., Jordan G.J., Wilf P., Cantrill D.J., Bayly M.J., Bromham L., Brown G.K., Carpenter R.J., Lee D.M., Murphy D.J., Sniderman J.M., Udovicic F. 2012. Testing the impact of calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus* (Fagales). *Syst. Biol.* 61:289–313.
- Schwartz R.S., Mueller R.L. 2010. Branch length estimation and divergence dating: estimates of error in Bayesian and maximum likelihood frameworks. *BMC Evol. Biol.* 10:5.
- Sepulchre P., Ramstein G., Fluteau F., Schuster M., Tiercelin J.J., Brunet M. 2006. Tectonic uplift and Eastern Africa aridification. *Science* 313:1419–1423.
- Shoo L.P., Rose R., Doughty P., Austin J.J., Melville J. 2008. Diversification patterns of pebble-mimic dragons are consistent with historical disruption of important habitat corridors in arid Australia. *Mol. Phylogenet. Evol.* 48:528–542.
- Simmons M.P., Ochoterena H. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Syst. Biol.* 49:369–381.
- Simon C., Frati F., Beckenbach A., Crespi B., Liu H., Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved PCR primers. *Ann. Entomol. Soc. Am.* 87:651–701.
- Sloan D.B., Taylor D.R. 2012. Evolutionary rate variation in organelle genomes: The role of mutational processes. In: Bullerwell CE editor. *Organelle Genetics*. Berlin, Springer, p. 123–146.
- Straus A. 1952. Beiträge zur Pliocänflora von Willershausen III. Die niederen Pflanzengruppen bis zu den Gymnospermen. *Paleont. Abteil. B. Paläeophytol.* 93:1–44.
- Sueur J., Puissant S. 2007. Similar look but different song: a new *Cicadetta* species in the *montana* complex (Insecta, Hemiptera, Cicadidae). *Zootaxa* 1442:55–68.
- Sueur J., Vanderpool D., Simon C., Ouvrard D., Bourgoïn T. 2007. Molecular phylogeny of the genus *Tibicina* (Hemiptera, Cicadidae): rapid radiation and acoustic behaviour. *Biol. J. Linn. Soc.* 91:611–626.
- Swenson U., Hill R.S., McLoughlin S. 2001. Biogeography of *Nothofagus* supports the sequence of Gondwana break-up. *Taxon* 50:1025–1042.
- Swofford D.L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sunderland (MA): Sinauer Associates.
- Thomas J.A., Welch J.J., Lanfear R., Bromham L. 2010. A generation time effect on the rate of molecular evolution in invertebrates. *Mol. Biol. Evol.* 27:1173–1180.
- Thomas J.A., Welch J.J., Woolfit M., Bromham L. 2006. There is no universal molecular clock for invertebrates, but rate variation does not scale with body size. *Proc. Natl. Acad. Sci. USA* 103:7366–7371.
- Thorne J.L., Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* 51:689–702.
- Tsantes C., Steiper M.E. 2009. Age at first reproduction explains rate variation in the strepsirrhine molecular clock. *Proc. Natl. Acad. Sci. USA* 106:18165–18170.
- Venditti C., Meade A., Pagel M. 2006. Detecting the node-density artifact in phylogeny reconstruction. *Syst. Biol.* 55:637–643.
- Vila R., Bell C.D., Macniven R., Goldman-Huertas B., Ree R.H., Marshall C.R., Balint Z., Johnson K., Benyamini D., Pierce N.E. 2011. Phylogeny and palaeoecology of *Polyommatus* blue butterflies show Beringia was a climate-regulated gateway to the New World. *Proc. Roy. Soc. Lond. B. Biol. Sci.* 278:2737–2744.
- Villet M.H., van Noort G.S. 1999. Cicadas (Hemiptera, Homoptera: Cicadoidea) of Mkomazi. In: Villet MH, van Noort GS, Packer M editors. *Mkomazi: the ecology, biodiversity, and conservation of a Tanzanian savanna*. London, Royal Geographical Society, p. 223–233.
- Wade E.J., Hertach T., Gogala M., Trilar T. Forthcoming 2015. Molecular species-delimitation methods recover most song- delimited cicada species in the European *Cicadetta montana* complex. *J. Evol. Biol.*
- Zwickl D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence data sets under the maximum likelihood criterion. Austin, TX, Ph.D. Thesis. The University of Texas, p. 115.