Unrecognized lineages transform our understanding of diversification in a clade of lizards

Jason G. Bragg[1,2,*], Sally Potter[1,3,*], Ana Afonso-Silva[1], Mozes P. K. Blom[1,4], Conrad J. Hoskin[5], Huw A. Ogilvie[1,6], Craig Moritz[1]

[1] Research School of Biology, Australian National University, Canberra, Australia

[2] Research Centre for Ecosystem Resilience, Botanic Gardens of Sydney, Sydney NSW 2000

[3] School of Natural Sciences, Macquarie University

[4] Leibniz Institut für Evolutions- und Biodiversitätsforschung, Museum für Naturkunde, 10115 Berlin, Germany. https://orcid.org/0000-0002-6304-9827

[5] College of Science & Engineering, James Cook University, Townsville, Qld, 4811. https://orcid.org/0000-0001-8116-6085

[6] Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, Texas

* contributed equally, and considered joint first authors

Correspondence –
JGB: Jason.Bragg@botanicgardens.nsw.gov.au
SP: Sally.Potter@mq.edu.au
CM: Craig.Moritz@anu.edu.au

Abstract

Evolutionary lineages at the tip of the tree of life can be genetically diverged yet phenotypically similar and therefore unrecognized by traditional morphology-based taxonomy. Such lineages, spanning the "grey zone of speciation" [1], are increasingly uncovered using genomic analyses. Here we show that incorporating this unrecognized lineage diversity into macro-evolutionary analyses yields novel insights into the speciation process. Examining a major clade of Australian skinks with extensive sampling of both unrecognised lineages and described species (199 lineages across 124 species) we find that lineages of this group have been forming at a constant net rate over time. In contrast, when including only the described species we see a slow-down in the net rate of diversification. Simulations of lineage formation via a protracted speciation model, extended to include multivariate trait evolution, indicate that phenotypic conservatism can explain the dynamics of taxonomically recognized diversity over time. Including intraspecific lineages in macroevolutionary analyses has provided new insights about the diversification process. In this case, it points to higher net rates of lineage than species formation, and a role for phenotypic constraint in generating cryptic lineage diversity.

Introduction

Taxonomically unrecognized lineages are pervasive and increasingly revealed by phylogeographic analyses [2,3]. While some are insufficiently diverged to warrant recognition as biological species, others may prove to be diagnosable based on subtle morphological differences and some may not have any recorded morphological differences at all. As such, the prevalence of phenotypically similar lineages challenges the view that most speciation is driven by ecological differentiation rather than relatively simple allopatric divergence [4,5]. Despite growing recognition that such lineage diversity is common, we often examine it one

54 species complex at a time, which limits understanding of general processes. An alternative is
55 to take a macroscopic view, and to exploit the commonness of unrecognized lineages to
56 quantitatively characterize patterns in their occurrence across a large clade, and to make
57 inferences about the processes that generate them.

58

59 Including this layer of taxonomically cryptic diversity in macroevolutionary analyses of
60 speciation is expected to improve insights into the dynamics of diversification [6] and the
61 protracted speciation process [7], which models the formation and fate of nascent species [8].
62 Understanding how species traits and environmental history affect the development and
63 persistence of intraspecific lineages will also shed light on speciation processes [9] and enable
64 better links between microevolutionary and macroevolutionary processes. For example,
65 intraspecific lineages should be more prevalent (and older) in species occurring in the tropics,
66 due to less volatile past climates and hence more stable geographic ranges [10], or at mid
67 latitudes as found in Pelletier and Carstens [11], and in those species with large ranges relative
68 to dispersal distance [12].

69

70 To explore additional insights into speciation from including intraspecific lineages in
71 macroevolutionary analyses, we apply phylogenomic and evolutionary analyses
72 (Supplementary file, 1.1) to extensively sampled phylogeographic lineages and recognized
73 species in a diverse group of scincid lizards, the ~25 Myr old, Australian Eugongylinae. This
74 clade, which has been thoroughly examined using morphology-based taxonomy, contains 18
75 genera, and spans the full latitudinal range and biome diversity of Australia (Fig. 1). As
76 ectotherms with generally low dispersal rates and high sensitivity to climate variation, these
77 taxa are expected to have extensive intraspecific variation, especially in tropical or

78 widespread species. Regarding diversification dynamics, we expect to infer different models
79 of rates when including intraspecific lineages, reflecting differences in dynamics of initiation
80 versus completion of speciation. We also explore whether it is necessary to invoke constraints
81 on trait divergence, additional to ephemeral divergence [13], to explain the high prevalence of
82 deep, taxonomically unrecognised intraspecific lineages in these lizards.

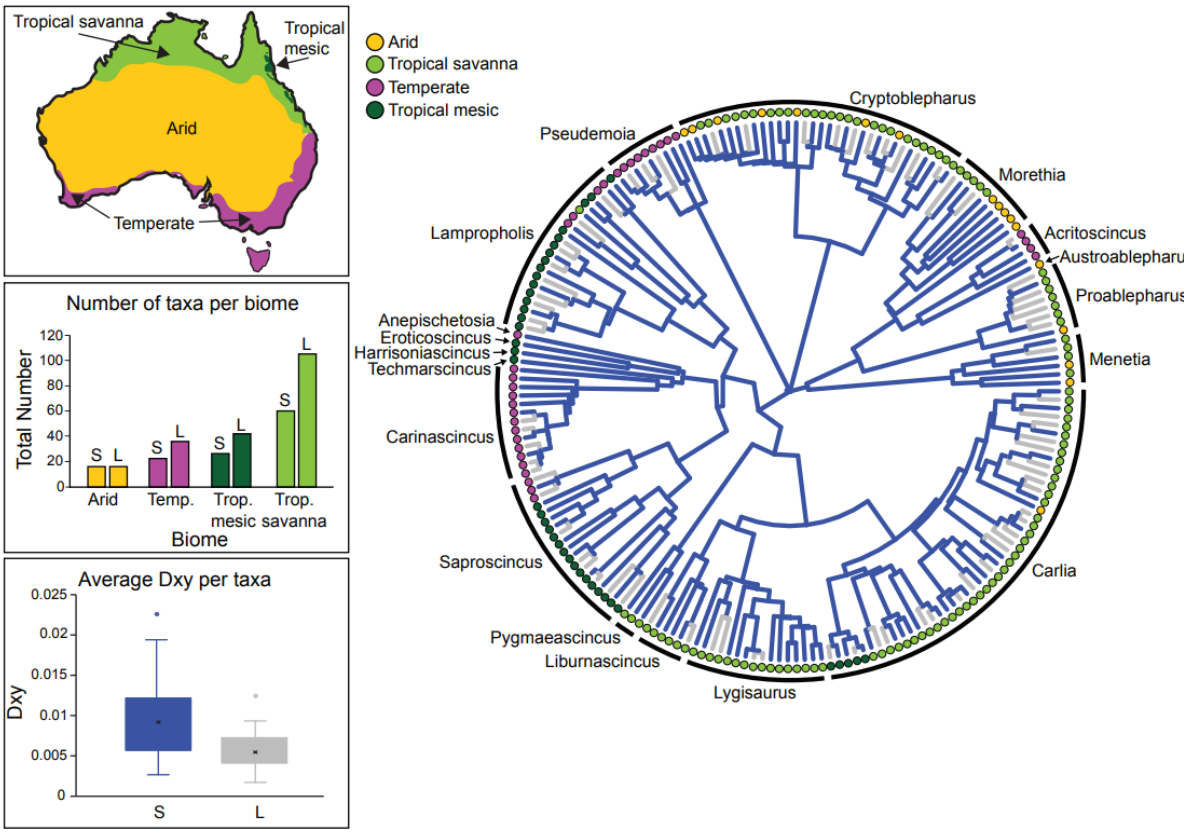


Figure 1. (Left, top) Map of Australia with distribution of major biomes, (Left, middle) comparison of number of species (S) versus lineages (L) across each biome (Trop. = Tropical), (Left, bottom) comparison of average Dxy of exons between lineages versus Dxy between species, and (Right) multi-locus phylogeny of the Australian Eugongylinae. The samples used in this study were distributed across the Australian landmass. The phylogeny shows one (randomly chosen) lineage for each species in blue and other intraspecific lineages for that species in grey, to illustrate the distribution of taxonomically unrecognised diversity across the tree. Genera are outlined on the phylogeny and the associated biome classification for each lineage and species are coloured.

83

84 We examine lineage diversity in this clade in the context of the alpha-taxonomy of species
85 that was inferred largely based on morphological evidence, and before multi-locus nuclear

datasets revealed substantial cryptic diversity in many species. Comparative hybrid zone studies have confirmed stronger isolation in more divergent lineage-pairs, reinforcing the view that this lineage diversity spans the grey zone of speciation [14]. Several of the species complexes revealed by recent genomic analyses have been subject to taxonomic revision [15,16] wherein some, but not all, lineages were elevated to species status. The same is likely to be true of intraspecific lineages in taxa not yet assessed: not all represent true cryptic species, but some may warrant recognition based on high genomic divergence or evidence of clear trait divergence via post-hoc morphological analyses.

Results - sampling and phylogeny

Across the Australian continental Eugongylinae, we assessed 124 species (98% of taxa recognised by morphology) and identified 199 lineages. This was based on extensive mitochondrial DNA sequencing and, in most cases, multi-locus sequence data (Fig. 1; Supplementary file, 1.1). In general, we defined 'lineages' as intraspecific populations that were mostly >3% divergent for mtDNA from their closest relative (range: 1-12%), reciprocally monophyletic for concatenated nDNA genes, and geographically cohesive. Levels of sequence divergence at exons (mean Dxy between species = 0.9% versus between lineages = 0.5%; Fig 1; Supplementary file, 1.2) are consistent with the "grey zone" of speciation as defined by Roux et al. [1]. Of the 75 intraspecific lineages additional to described species, about half have divergence times greater than that accompanied by strong reproductive isolation in analyses of contact zones between cryptic lineages (Supplementary files, 1.4; [14]) and so represent candidate cryptic species [15]. By contrast, several species (typically with distinct male breeding colours [17,18]) had divergence times less than this empirical threshold for cryptic species. The number of intraspecific lineages varies from one to seven and was greater in species with larger geographical ranges (Supplementary file, 1.4,

phyloglm, z = 3.1, p= 0.002). There were also more lineages within species living at lower latitudes (phyloglm, z = 2.9, p = 0.004) and warmer climates (phyloglm, z = 2.5, p = 0.01) (Supplementary file, Fig 2 in 2.1).

We estimated the phylogeny of Eugongylinae skinks based on a phylogenomic dataset including (usually) at least two samples per lineage. The data were generated using a targeted exon capture design, sequenced to high coverage, and were highly complete in terms of both taxa and loci. To avoid problems with inflation of tip lengths in recent radiations [19], the phylogeny was estimated using the full Bayesian species tree implemented in StarBEAST2 [20] for two independent sets of 100 loci and 410 total samples. Given that the data set is massive relative to computational requirements, we used a hierarchical approach, first analysing 13 clades that were well supported in preliminary analyses and then sampling most divergent taxa from these together with phylogenetically uncertain taxa in a backbone tree (Supplementary files, 2.1).

Intergeneric relationships are well supported in most cases and consistent between the two sets of loci and with estimates from concatenation of the same loci, and larger datasets (1270 exonic loci, [21]; and 2-3 loci including non-Australian species of the Eugongylinae radiation, [22]; Supplementary file, 2.1 and 2.2). The main exceptions were three genera – *Cryptoblepharus, Menetia* and *Pseudemoia,* which branched deep and rapidly in the phylogeny and differed in their placement across the two species tree analyses and concatenation analyses. These genera aside, the species tree analyses revealed two well supported, major clades: one with five genera more strongly represented in temperate and arid biomes (*e.g.*, *Acritoscincus, Morethia, Proablepharus, Austroablepharus,* and *Menetia*),

and another with 11 genera that represent the bulk of diversity from tropical to eastern temperate biomes (Fig. 1). Of note in the latter clade is the monophyly of three highly divergent monotypic mesic-forest genera (*Anepischetosia*, *Eroticosincus*, *Harrisoniascincus*) and the sister relationship of another monotypic genus, *Techmarscincus*, from montane tropical rainforest, with the "snow skinks" (*Carinascincus*) of south-eastern Australia. These are sister to a clade containing *Lampropholis* and *Saproscincus*, most of which are eastern rainforest-restricted species. Sister to all these is a clade of tropical savanna genera (*Carlia*, *Liburnascincus*, *Lygisaurus*, *Pygymaeascincus*).

Results – macroevolutionary inferences

The branching patterns of phylogenetic trees can, with care, be examined to make inferences about the processes of diversification that generated them [23]. In the following, we compare inferences about diversification processes between a tree that included all extant lineages (hereafter 'lineage tree') and a tree in which each taxonomic species was represented by a single lineage (hereafter 'taxonomy tree') (Fig. 1). The latter represents the level of sampling that is available for most macroevolutionary studies that are based on recognised species. We expect that including intraspecific lineages will support different diversification models or yield different parameter estimates for rates of speciation and/or extinction.

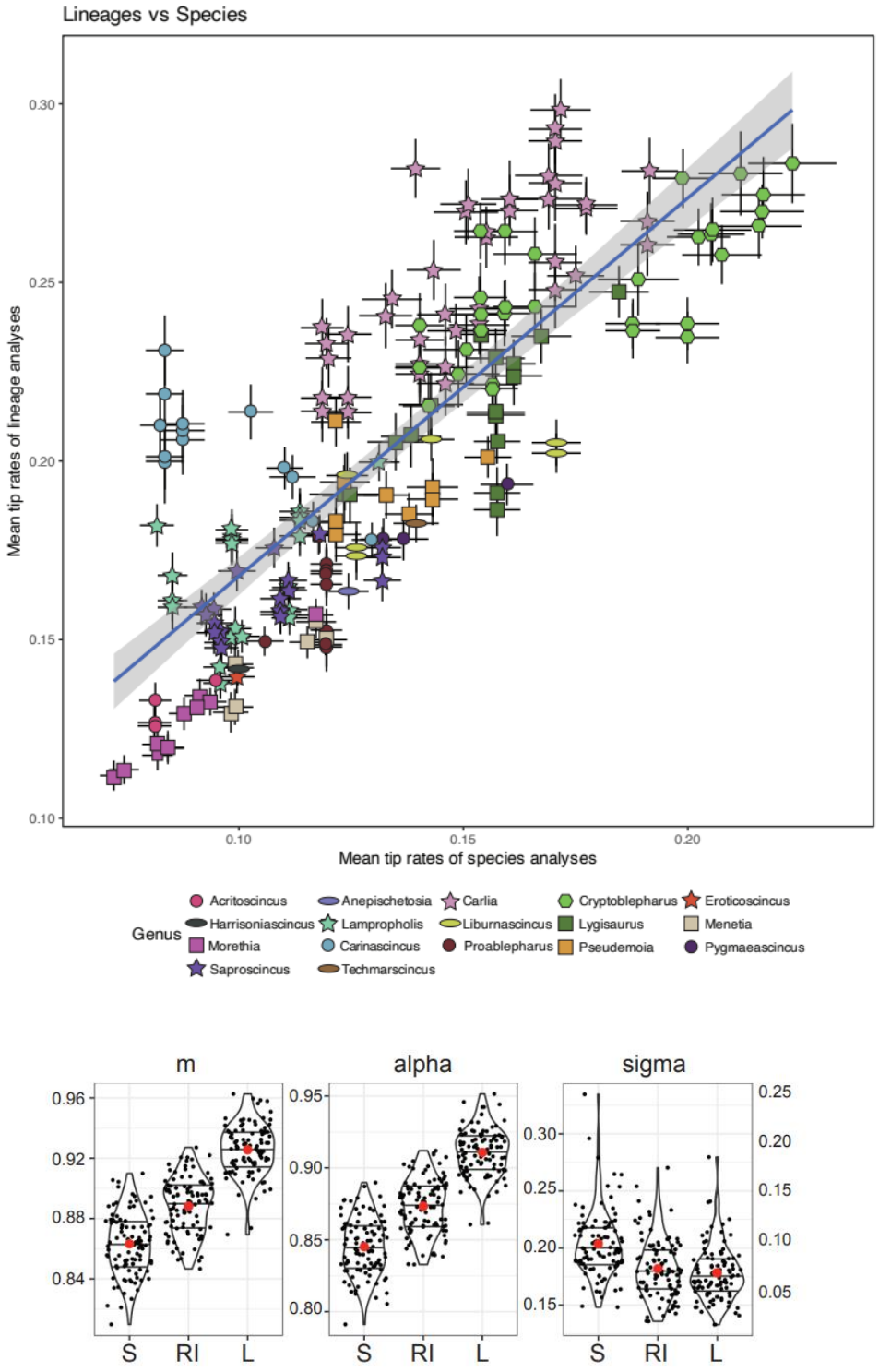


Figure 2. Results from diversification analyses using CLaDS. Top: Estimated values of branch specific diversification rates for tree tips estimated using the lineage tree (y-axis) and the taxonomy tree (x-axis). Bottom: estimates of CLaDS model parameters from posterior distributions of StarBEAST2 trees including species (one lineage per taxon), lineages expected to show reproductive isolation (RI), and all lineages (L).

153 The influence of intraspecific lineages on inferred diversification processes was
154 heterogeneous across the tree. As expected from sampling alone, tip-specific rates are lower
155 in the taxonomy than lineage tree overall, but there were also notable differences among
156 genera (Fig. 2). The snake-eyed skinks (*Cryptoblepharus*) and rainbow skinks (*Carlia*) had
157 the highest tip-rates overall, consistent with the recent radiations of these diverse genera [24,25].
158 The former genus has diversified by both cryptic and ecological speciation [24] and, based on
159 divergent breeding colours, there is a strong possibility of divergent sexual selection in *Carlia*
160 [26]. Further, the rainbow skinks had the highest excess of tip rates for the lineage tree relative
161 to the taxonomy tree (Fig. 2), reflecting the pronounced phylogeographic structuring within
162 species of this largely tropical genus [16,27,28]. Estimates of branch-specific speciation rates [29]
163 inferred more decline (lower m) and more constraint (higher sigma) in rates towards the
164 present for the taxonomy trees than for the lineage trees (Fig. 2). This remains true when
165 including only lineages that are likely to represent cryptic species (RI; divergence greater
166 than the threshold of reproductive isolation outlined in [15]). These differences in
167 macroevolutionary parameters were consistent using phylogenies generated from independent
168 sets of loci and estimated using concatenation or species tree approaches (Supplementary file,
169 2.3).

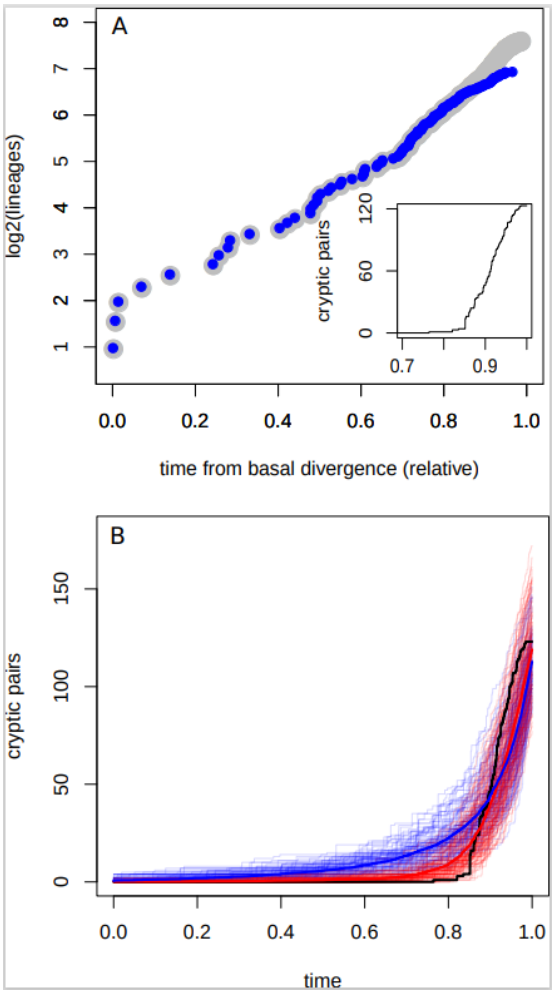


Figure 3. Dynamics of species and lineage diversification. (A) Lineages through time plot depicting the (log) number of lineages in the lineage tree (grey) and the taxonomy tree (blue) at different times from the basal divergence. Inset, the accumulation of 'cryptic' pairs of lineages in the tree at different times from the basal divergence. (B) The numbers of cryptic lineage pairs as a function of the time from the basal divergence. Black line is empirical data for the lineage tree. Red lines are simulations that produced best fits to data (posterior distribution) when *p* was allowed to vary between values of 0.5 and 1. Blue lines are simulations that provided best fits to data when *p* was set to 0. In the absence of suppressed morphological divergence between nascent and parent lineages ($p = 0$), model solutions tend to produce cryptic lineages with much greater ancestor depths than observed, or underestimate the number of extant cryptic lineages.

The taxonomy tree exhibited a decline in net diversification near the present (relative to the lineage tree), manifesting as deceleration in the plot of the (log) number of tips over time (Fig. 3A). These observations are corroborated by two statistical approaches. First, we compared models of diversification where rates of speciation and extinction had different

functional forms (constant or exponential). For the taxonomy tree, a diversification model featuring speciation rates that decline towards the present explained the data (ancestor depths in the tree) better than a model with constant rates of speciation and extinction (Supplementary file, Table 1 in 2.1). Conversely, for the lineage tree, a model with constant speciation and extinction rates provided a better explanation than one with time-dependent rates. Second, using the robust 'pulled speciation rate' approach [30], we observed that the taxonomy tree had estimated values of recent net diversification rate that were smaller than those from the lineage tree (Supplementary file, Fig 3 in 2.1). In sum, the taxonomy tree supported a stronger recent decline in net diversification than the lineage tree. Slowing net diversification has often been observed and is potentially explained by a range of mechanisms [31]. These include changes over time in the availability or suitability of niches, and factors that influence the detection of lineages. Our analyses potentially help disentangle these mechanisms, suggesting that the absence of recently formed lineages that are not represented in species-level taxonomy partly explains the inferred slowing of net diversification in the species tree.

Finally, we used macroevolutionary simulations to study processes of lineage formation, divergence and detection more explicitly. Our approach was based on a protracted speciation model of lineage diversification, in which new lineages pass through a 'nascent' stage on their way to completing the process of speciation [32]. To this we added a simple representation of multivariate trait diversification, with a parameter (p) linking change in the traits of a nascent lineage to change in its parent lineage, allowing for morphological variation to be modelled within the protracted speciation model (Supplementary file, Fig. 5 in 2.1). At the end of a simulation, we identified pairs of lineages whose distance from each other in trait space was less than a threshold value (a parameter of the model) and defined these as

intraspecific lineages. This was intended to represent a process in which lineages become genealogically separated but have not acquired phenotypic differences that would result in an alpha-taxonomic diagnosis.

Several interesting observations emerge from the simulations (Fig. 3). First, 'protracted speciation,' here combined with conservative trait evolution, helped explain the rapid, recent, growth in lineage diversity in Australian Eugongylinae, some of which represents cryptic species. In particular, we observed that models providing the best explanation for our data had relatively large values of *p* (Supplementary file, Fig. 5C in 2.1), which increase the probability that nascent lineages inherited the same changes in trait values as their parent lineage. To help explain this, we ran additional simulations in which *p* was set to zero (i.e., no trait covariation between nascent and parent lineages following divergence). Here the simulations that best fit the data tended to have fewer, and older, cryptic pairs than observed in our empirical dataset (blue lines in Fig. 3B). Taken together, these results suggest that to capture the rapid and recent appearance of taxonomically unrecognised lineages in this radiation, we needed to impose covariation in the morphological changes of nascent lineages, or to constrain their rate of morphological divergence. Second, we observed that species delimited in simulations based on morphospace distances were sometimes not monophyletic, in terms of their divergence history. The numbers and frequencies of these non-monophyletic taxa from simulations were consistent with those observed in the empirical tree of the Australian Eugongylinae (Supplementary file, Fig 6 in 2.1).

Discussion

Macroevolutionary analyses typically use taxonomically recognised species as units of analysis which, for most radiations, correspond to morphologically distinct taxa. This has limited our ability to understand speciation as an outcome of the intraspecific processes generating new lineages, as represented in the protracted speciation model [32]. This model has unidentifiable parameters when fit to species-level phylogenies [33], but it is possible to gain new insights from it by including intraspecific lineages [8] because these provide additional information on diversification. Here, extensive sampling of intraspecific lineages and robust estimates of topology and branching times across a major clade yielded novel insights into patterns in the distribution of diversity, and the processes that generate it. These include evidence of higher rates of lineage formation in the tropics, biologically informative differences among genera in rates of lineage formation relative to speciation, that accumulation of lineages is constant versus a slowdown for taxonomically recognised species, and that constrained trait evolution is necessary to explain the dynamics of lineage formation and cryptic speciation.

Previous studies of individual taxa within this group have pointed to differences in geographic scales of phylogeographic divergence [16,27, 28] and in speciation processes [15,24,34]. Taking a clade-wide approach enables generalities to emerge, overcoming the otherwise idiosyncratic nature of speciation. As accurate phylogenies with good sampling of evolutionarily independent tips (whether identified taxonomically as species or not) in large radiations become commonplace, we can expect more such insights into how species form over time and space. In particular, including cryptic species and unrecognized lineages will improve our ability to merge micro- and macro-evolutionary perspectives [8,13] and so understand why some lineages and regions have more diversity than others [9].

Materials and Methods

Sampling. We set out to obtain multiple samples for all lineages of continental Australian Eugongylinae skink species that are known or suspected to have strong phylogeographic structure. In general, we sampled well supported lineages in taxa surveyed genetically, sometimes lumping described shallow mtDNA lineages to avoid oversampling shallow intraspecific structure. For several widespread taxa we generated new mtDNA sequence data to guide sampling for exon-capture sequencing. See Supplementary files 1.1 and 2.1 for details of sampling and analyses of mtDNA data.

Data generation. For each sample, we performed target capture sequencing across 3320 protein coding exons using laboratory and bioinformatic workflows that have been described previously[35] (see Supplementary file, 2.1). These data have previously been used in clade specific analyses (e.g. 24, 25) and to explore rates of molecular evolution across species using concatenated sequences [21]. We estimated phylogenetic relationships among taxa using StarBEAST2, a multi-species coalescent species tree method [20], with a hierarchical approach to allow for parallel analyses of the two independent 100 locus datasets across 400 samples and ~ 200 taxa. For comparison, we used BEAST to estimate phylogenies from concatenations of the same 100 locus datasets.

Analyses. We tested associations between the number of cryptic lineages observed per species and environmental variables for those species mean annual temperature, geographic range size, and the mean latitude. Values for these environmental variables were obtained by intersecting species' distribution records with environmental data from the Atlas of Living Australia. We tested these associations using Phylogenetic Generalized Linear Models (R package phylolm [36]). These analyses were performed using different trees to ensure they were robust to uncertainty in phylogenetic inference. This included the two trees estimated with

different sets of loci (described above), as well as posterior samples from the estimation of these trees (see Supplementary file 2.1 for details).

We fit models of clade diversification to species and lineage trees (R package RPANDA, [37]). We performed several analyses using three sets of trees: one with all lineages represented, another downsampled to include just lineages expected to have reproductive isolation (RI taxa + species) and another including just a single lineage from each taxonomically recognised species. Each set included time-trees derived from independent sets of 100 loci and estimated using species tree methods (StarBEAST2) and concatenation (BEAST). We did this using the taxonomy that had been established largely using morphological characters, sometimes in combination with small molecular datasets, but before phylogenomic data were used extensively to delimit and describe lineages that are essentially cryptic morphologically (as in [15,39]).

The diversification models had different functions for speciation rate and extinction rate, which were either constant, or varied over time according to an exponential function. For each tree, we estimated the parameters for four models (all combinations of constant and exponential speciation and extinction rates) by Maximum Likelihood, and inferred which model best fit the data for each tree using an AICc criterion.

Simulations. We performed simulations to better understand processes attending the accumulation of cryptic species over evolutionary time. To do this, we used models of lineage diversification (protracted speciation[31]), coupled with a simple representation of phenotypic evolution. Here new lineages arise through evolutionary divergence, and can be placed in the same or different 'species' depending on the amount of multivariate trait distance they exhibit. We used a model in which new lineages arise from existing lineages at a prescribed rate (lineage 'birth'), and are lost at a prescribed rate (lineage 'death'), but new lineages remain 'nascent' until they complete the processes of speciation[31]. Concomitantly, a

297 set of five idealised phenotypic traits evolve randomly over time in each lineage. That is, at
298 each time step, a change occurs in each trait value in each lineage, with the size of the change
299 (δ) drawn from a normal distribution. At the end of a simulation, a pair of lineages was
300 designated 'cryptic' if their distance in trait space was less than a threshold value. When a
301 lineage was born, it inherited the trait values of its parent lineage, and while it was 'nascent',
302 the evolution of its morphological traits was linked to the morphological evolution of its
303 parent lineage. It was linked in that, at each time step, the nascent lineage received changes in
304 its trait (δ) values equal to those assigned to the parent lineage with some probability, p, and
305 received their own values with 1-p. In this way, p controlled the strength of the link between
306 nascent and parent lineage trait evolution. After speciation completion, the trait evolution of
307 the nascent lineage became independent of its parent. We fitted the model to the
308 Eugongylinae lineage dataset by performing simulations with model parameters drawn at
309 random from plausible ranges (prior distributions), and then identified a set of simulations (a
310 posterior sample) and associated parameter values that fit the data based on four summary
311 statistics: (i) the total number of lineages, (ii) the number of cryptic pairs (i.e. pairs
312 unrecognised by alpha taxonomy), and normalized [40] representations of the accumulation of
313 (iii) lineages and (iv) cryptic lineage pairs over time [32] (see Supplementary file, 2.1 for more
314 detail).
315
316
317
318
319
320
321
322
323
324

Acknowledgements

We thank Stephen Donnellan, Mark Adams, Paul Doughty, Mark Hutchinson, Maggie Haines, Adnan Moussalli, Charlotte Jennings, Erik Wapstra, Gaye Bourke, Jessica Worthington Wilmer and Joanna Sumner for providing samples and/or unpublished data. We are grateful to landowners, including indigenous communities for access to their lands and to curators at state museums for access to tissue collections. Ke Bi, Sonal Singhal and Lydia Smith provided help with initial experiments and discussions; and Sonal Singhal and Rhiannon Schembri provided insightful comments on the manuscript. Field work was conducted with permits from relevant State Conservation Agencies and Animal Ethics permits from ANU. This research was supported by an ARC Laureate Fellowship to CM and an ARC Future Fellowship to SP.

References

1. Roux, C. et al. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biol.* **14**, e2000234 (2016). https://doi.org/10.1371/journal.pbio.2000234
2. Struck, T.H. et al. Finding evolutionary processes hidden in cryptic species. *TREE* **33**, 153–163 (2018). https://doi.org/10.1016/j.tree.2017.11.007
3. Fišer, C. et al. Cryptic species as a window into the paradigm shift of the species concept. *Mol. Ecol.* **27**, 613-635 (2018). https://doi.org/10.1111/mec.14486
4. Rundle, R.J. & Price, T.D. Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *TREE* **24**, 394-399 (2009). https://doi.org/10.1016/j.tree.2009.02.007
5. Anderson, S.A.S. & Weir, J.T. The role of divergent ecological adaptation during allopatric speciation in vertebrates. *Science* **378**, 1214-1218 (2022). DOI: 10.1126/science.abo771
6. Faurby, Strong effects of variation in taxonomic opinion on diversification analyses? Methods in ecology and evolution. *Methods Ecol. Evol.* **7**, 4-13 (2016). https://doi.org/10.1111/2041-210X.12449
7. Rosindell, J. et al. Protracted speciation revitalizes the neutral theory of biodiversity. *Ecol. Lett.* **13**, 716–727 (2010). https://doi.org/10.1111/j.1461-0248.2010.01463.x
8. Hua, X. et al. Protracted speciation under the state-dependent speciation and extinction approach. *Syst. Biol.* **71**, 1362-1377 (2022). https://doi.org/10.1093/sysbio/syac041
9. Harvey, M.G. et al. Beyond reproductive isolation: demographic controls on the speciation process. *Annu. Rev. Ecol. Evol. Syst.* **50**, 75-95 (2019). https://doi.org/10.1146/annurev-ecolsys-110218-024701
10. Smith, B.T, et al. A latitudinal phylogeographic diversity gradient in birds. *PLoS Biol*. 15:e2001073 (2017) https://doi.org/10.1371/ journal.pbio.2001073
11. Pelletier, T.A. & Carstens, B.C. Geographical range size and latitude predict population genetic structure in a global survey. *Biol. Lett*. **14**, 20170566 (2018). https://doi.org/10.1098/rsbl.2017.0566
12. Parsons, D.J. et al. Analysis of biodiversity data suggests that mammal species are hidden in predictable places. *PNAS* **119**, e2103400119 (2022). https://doi.org/10.1073/pnas.2103400119
13. Futuyma, D.J. Evolutionary constraint and ecological consequences. *Evolution* **64**, 1865–1884 (2010). doi:10.1111/j.1558-5646.2010.00960.x

14. Singhal, S. & Bi, K. History cleans up messes: the impact of time in driving divergence and introgression in a tropical suture zone. *Evolution* **71**, 1888-1899 (2017). https://doi.org/10.1111/evo.13278
15. Singhal, S. et al. A framework for resolving cryptic species: a case study from the lizards of the Australian wet tropics. *Syst. Biol.* **67**, 1061-1075 (2018). https://doi.org/10.1093/sysbio/syy026
16. Afonso Silva, A.C. et al. Tropical specialist vs. climate generalist: diversification and demographic history of sister species of *Carlia* skinks from northwestern Australia. *Mol. Ecol.* **26**, 4045-4058 (2017). https://doi.org/10.1111/mec.14185
17. Couper, P.J. et al. Skinks currently assigned to *Carlia aerata* (Scincidae: Lygosominae) of north-eastern Queensland: a preliminary study of cryptic diversity and two new species. *Aust. J. Zool.* **53**, 35-49 (2005). https://doi.org/10.1071/ZO04010
18. Hoskin, C.J. & Couper, P.J. Description of two new *Carlia* species (Reptilia: Scincidae) from north-east Australia, elevation of *Carlia pectoralis inconnexa* Ingram & Covacevich 1989 to full species status, and redescription of *Carlia pectoralis* (de Vis 1884). *Zootaxa* **3546**, 1-28 (2012). https://doi.org/10.11646/zootaxa.3546.1.1
19. Ogilvie, H.A. et al. Computational performance and statistical accuracy of* BEAST and comparisons with other methods. *Syst. Biol.* **65**, 381-396 (2016). https://doi.org/10.1093/sysbio/syv118
20. Ogilvie, H.A. et al. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Mol. Biol. Evol.* **34**, 2101-2114 (2017). https://doi.org/10.1093/molbev/msx126
21. Ivan, J. et al. Temperature predicts the rate of molecular evolution in Australian Eugongylinae skinks. *Evolution* **76**, 252-261 (2022). https://doi.org/10.1111/evo.14342
22. Chapple, D.G. et al. Phylogenetic relationships in the Eugongylini (Squamata: Scincidae): generic limits and biogeography. *Aust. J. Zool.* **70**, 165-203 (2023). https://doi.org/10.1071/ZO23007
23. Morlon, H. et al. Phylogenetic Insights into Diversification. *Annu. Rev. Ecol. Evol. Syst.* **55** (2024). https://doi.org/10.1146/annurev-ecolsys-102722-020508
24. Blom, M.P. et al. Convergence across a continent: adaptive diversification in a recent radiation of Australian lizards. *Proc. R. Soc. Lond. B. Biol. Sci.* **283**, 20160181 (2016). https://doi.org/10.1098/rspb.2016.0181
25. Bragg, J.G. et al. Phylogenomics of a rapid radiation: the Australian rainbow skinks. *BMC Evol. Biol.* **18**, 1-12 (2018). https://doi.org/10.1186/s12862-018-1130-4
26. Dolman, G and Stuart-Fox, D. Processes driving male breeding colour and ecomorphological diversity in rainbow skinks: a phylogenetic comparative test. Evol. Ecol. 24, 97-113 (2009). https://doi.org/10.1007/s10682-009-9293-5
27. Dolman, G. & Moritz, C. A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution* **60**, 573–82 (2006). https://doi.org/10.1111/j.0014-3820.2006.tb01138.x
28. Potter, S. et al. Pleistocene climatic changes drive diversification across a tropical savanna. *Mol. Ecol.* **27**, 520-532 (2018). https://doi.org/10.1111/mec.14441
29. Maliet, O. et al. A model with many small shifts for estimating species-specific diversification rates. *Nat. Ecol. Evol.* **3**, 1086-1092 (2019). https://doi.org/10.1038/s41559-019-0908-0
30. Louca, S. & Pennell, M.W. Extant timetrees are consistent with a myriad of diversification histories. *Nature* **580**, 502-505 (2020). https://doi.org/10.1038/s41586-020-2176-1

31. Moen, D. & Morlon, H. Why does diversification slow down? *TREE* **29**, 190-197 (2014). https://doi.org/10.1016/j.tree.2014.01.010
32. Etienne, R.S. & Rosindell, J. Prolonging the past counteracts the pull of the present: protracted speciation can explain observed slowdowns in diversification. *Syst. Biol.* **61**, 204 (2012). https://doi.org/10.1093/sysbio/syr091
33. Simonet, C. et al. Robustness of the approximate likelihood of the protracted speciation model. *J. Evol. Biol.* **31**, 469-479 (2018). https://doi.org/10.1111/jeb.13233
34. Haines, M.L. et al. Phylogenetic evidence of historic mitochondrial introgression and cryptic diversity in the genus *Pseudemoia* (Squamata: Scincidae). *Mol. Phylogenet. Evol.* **81**, 86-95 (2014). https://doi.org/10.1016/j.ympev.2014.09.006
35. Bragg, J.G. et al. Exon capture phylogenomics: efficacy across scales of divergence. *Mol. Ecol. Res.* **16**, 1059-1068 (2016). https://doi.org/10.1111/1755-0998.12449
36. Atlas of Living Australia at http://www.ala.org.au. Accessed 8 June 2022.
37. Ho, L.S.T. & Ane, C. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol.* **63**, 397-408 (2014). https://doi.org/10.1093/sysbio/syu005
38. Morlon, H. et al. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol.* **7**, 589-597 (2016). R package version 1.4, https://CRAN.R-project.org/package=RPANDA.
39. Afonso Silva, A.C. et al. Validation and description of two new north-western Australian Rainbow skinks with multispecies coalescent methods and morphology. *PeerJ* **5**, *e3724* (2017). https://doi.org/10.7717/peerj.3724
40. Janzen, T. et al. Approximate Bayesian Computation of diversification rates from molecular phylogenies: introducing a new efficient summary statistic, the nLTT. *Methods Ecol Evol*, 6: 566-575 (2015). https://doi.org/10.1111/2041-210X.12350

Supplementary file 1.1. Sampling of taxa and evidence for intraspecific lineage diversity.

Species sampled

We focussed on species of Eugongyline skinks from continental Australia, which are a subset of the Sahul clade of this family (Chapple et al. 2023). This larger clade also includes *Emoia* (largely New Guinea and Pacific islands) and a diverse assemblage from New Zealand and New Caledonia. We excluded the two species of *Emoia* and *Eugongylus rufescens*, all from Cape York, as these have most of their geographic ranges to the north of Australia. Of the target taxa, tissues were not available for four species with very small ranges - *Austroablepharus barryloni*, *A. naranjicauda*, *Menetia amaura* and *Saproscincus oriarus* (which for mtDNA is nested within the lineages of *S. mustelinus*, Moussalli et al. 2005). Species of *Cryptoblepharus* and *Carlia* external to continental Australia were also excluded. In total we included samples from 124 of the 128 target taxa.

We represent four recently described species (*Carlia isostriacantha*, *C. insularis*, *Lampropholis similis* and *L. bellendenkerensis*) as lineages, not species in these analyses as they were recognised primarily from genome-scale evidence and are near identical morphologically (Afonso Silva 2017; Singhal et al. 2018).

Recognition and sampling of lineages within species.

Many (46) of the species have been subject to prior range-wide mtDNA phylogeographic analyses, several (30 species) with accompanying multi-locus evidence. For these we mostly followed the given lineage delineations unless mtDNA divergences (here represented as simple p-distances) were very low, in which case we mostly collapsed lineages as described. Fifteen species have very small geographic ranges and were assumed to have a single lineage. Four geographically widespread species had insufficient samples to undertake mtDNA sequencing and are represented as a single lineage (see below). For several widespread species of snake-eyed skinks (*Cryptoblepharus*) we relied on extensive analyses of allozyme variation (Horner & Adams 2007), which were then used to guide taxonomic revisions (Horner 2007) to identify taxa warranting more detailed screening with mtDNA and exon sequencing. For the remainder (37 species) we sequenced mtDNA from geographically distant samples to identify candidate lineages (see below) and then aimed to include a least two samples per candidate lineage in the exon sequencing and subsequent phylogenetic analyses.

Following preliminary phylogenetic analyses (using a combination of SVDquartets, ASTRAL and NTree) of 700 exons sequenced from these candidate lineages we then ensured that samples from each candidate lineage were monophyletic, collapsing lineages where this was not the case. Given computational limits for StarBEAST2 analyses, we also subsampled known lineage diversity in fire-tailed skinks (*Morethia*) including two of four lineages in *M. storri* (Potter et al. 2019) and one of three in *M. boulengeri* and one of two in each of *M. lineoocellata* and *M. butleri*. In all, while recognising that our discovery and sampling of lineage diversity is inevitably incomplete, we are confident of a high level of sampling across genera, biomes and other variables considered here. A summary of the prior and new evidence for intraspecific lineage diversity across species is summarised in the following Table and Supplementary file, 1.1a.

Species	mtDNA lineages	notes	reference
Acritoscincus duperreyi	13, with 2-5% divergence	Also supporting SNP data. Only 2 most divergent lineages included here.	Dubey & Shine 2010; Dissanayake et al. 2022
Acritoscincus platynota	4 with 3-5% divergence	Collapsed to 1 lineage (IV + V)	Dubey & Shine 2010
Acritoscincus trilineata	2 with 3% divergence disjunct WA and SA samples	Included SA and WA as separate lineages	Dubey & Shine 2010
Anepischetosia maccoyi	1, with < 3% divergence	Undescribed taxon from NE of range not included	This paper; R. Schembri, P. Oliver & G. Shea pers. comm.
Australoblepharus kinghorni	1, with < 3% divergence	Sparse sampling	Potter et al. 2019
Carinascincus coventryi	1, with < 3% divergence	Sparse sampling	This paper
Carinascincus greeni	1, with < 3% divergence		This paper
Carinascincus metallicus	3, with 4-8% divergence	Divergent mtDNA lineages not monophyletic for exons; collapsed to 1 lineage	Kreger et al. 2020
Carinascincus microlepidotus	3, with 4 % divergence; 3% from C. palfreymani	Additional sampling for exons. Paraphyletic with C. palfremani	Melville & Swain 2003; C. Jennings & S. Potter unpubl. data
Carinascincus ocellatus	3, with low mtDNA divergence	Collapsed to 1 lineage	Cliff et al. 2015
Carinascincus orocryptus	3, with 1 to 3% divergence; mtDNA polyphyletic with C. microlepidotus.	Additional sampling for exons – orocryptus monophyletic for nDNA. 3 lineages included	Melville & Swann 2003; Jennings & Potter unpubl data
Carinascincus pretiosus	1, with <3% divergence		This paper
Carlia amax	4, with 7-10% divergence	4 lineages included; supported by extensive exon sequencing and SNP screening at contact zones.	Potter et al. 2016; Fenker et al. 2023
Carlia decora	1, with <3% mtDNA divergence		This paper
Carlia sp carnarvon	Highly divergent (>7%) for mtDNA, but similar to C. pectoralis for nDNA	Requires further sampling and analysis	Hoskin unpubl.
Carlia dogare	2, with 3% divergence	2 lineages included; Limited sampling	This paper
Carlia gracilis	5, with 1-7% divergence	5 lineages included; supported by extensive exon sequencing;	Potter et al. 2018
Carlia inconnexa	1, with <3% divergence		Hoskin unpubl.
Carlia jarnoldiae	3, with 4-8% divergence	reduced to 2 lineages with exon sequencing; probable RI	This paper
Carlia johnstonei	3, with 5-9% divergence; reduced to 2 with exon sequencing	Included newly described C. insularis as Johnsonei_B lineage	Afonso Silva et al. 2017
Carlia longipes	1, with < 3% divergence	Sampled across range	This paper

Carlia munda	6, with 11-19% divergence	Reduced to 3 with extensive exon sequencing; 3 lineages included	Potter et al. 2018
Carlia pectoralis	1, with < 3% divergence		Hoskin unpubl.
Carlia quinquecarinata	1, with < 3% divergence		Donnellan et al. 2009
Carlia rhomboidalis	3, with 4-6% divergence	2 most divergent lineages included; supported by multilocus sequencing,	Dolman & Moritz 2006
Carlia rubigo	1, with <3 % divergence		Hoskin, unpubl data
Carlia rubrigularis	2, with 8% divergence	2 lineages included, supported by multilocus sequencing, paraphyletic wih C. rhomboidalis, strong RI in contact zone	Dolman & Moritz 2006; Singhal & Bi 2017
Carlia rostralis	1, with <3% divergence	Sampled across range	This paper
Carlia rufilatus	3, with 6-8% divergence	3 lineages included, supported by extensive exon sequencing; most divergent with probable RI	Potter et al. 2018
Carlia schmeltzi	3, with 5-7% divergence,	3 lineages included	This paper
Carlia sexdentata	2, with 5% divergence	Polyphyletic with NG taxa in exon phylogeny; 2 lineages included	Donnellan et al. 2009; Rittmeyer et al. unpubl; This paper
Carlia storri	2, with 8% divergence	Sparse sampling; 2 lineages included	This paper
Carlia tetradactyla	1, with <3% divergence	Limited sampling but spans geographic range	This paper
Carlia triacantha	2, with 9% divergence	2 lineages with newly described C. isostriacantha as triacantha_A lineage; supported by extensive exon sequencing;	Afonso Silva et al. 2017
Carlia vivax	2, with 5% divergence	Sparse sampling; 2 lineages included	This paper
Cryptoblepharus cygnatus	2, with 4% divergence	2 lineages included; probable RI	Blom unpublished, This paper.
Cryptoblepharus juno	mtDNA polyphyletic with C. metallicus but species are monophyletic for exons	Kimberley and NT populations distinct for exons in more extensive dataset; 2 lineages included with probable RI	Blom unpublished, This paper.
Cryptoblepharus mertensi	2, with 4% divergence and paraphyletic with C. zoticus	2 lineages included with probable RI	Blom unpublished, This paper.
Cryptoblepharus megastictus	2, with 6% divergence	1 lineage for nDNA	Blom unpublished, This paper.
Cryptoblepharus metallicus	mtDNA polyphyletic with C. juno/daedalos	Monphyletic for exons with 4 lineages (6-10% mtDNA divergence) included to represent diversity; probable RI	Blom unpublished, This paper.
Cryptoblepharus plagiocephalus	No mtDNA	2 non-monophyletic lineages for exons; 2 included	Blom unpublished, This paper.
Cryptoblepharus ruber	2 with 6% divergence. Buckle Head sample introgressed with metallicus mtDNA	Paraphyletic with megastictus. 3 lineages included with probable RI	Horner & Adams 2007; Blom unpublished, This paper.

Cryptoblepharus tytthos		2 lineages for allozymes & exons; probable RI	Horner & Adams 2007; Blom unpublished, This paper.
Cryptobelpharus virgatus	2 with 6% divergence.		
Lampropholis adonis	3 with 8-12% divergence	3 lineages included; probable RI	Hoskin unpublished, This paper
Lampropholis coggeri	4 with 5-9% divergence	4 lineages included with potential RI in 3; supported by multilocus sequencing. 2 southern lineages later recognised as separate species	Bell et al. 2010; Singhal & Bi 2017; Singhal et al. 2018
Lampropholis delicata	9, with 4-7% divergence	2 lineages spanning deepest node (1 vs 7+9) included here; probable RI	Chapple et al. 2011a
Lampropholis guichenoti	2 major lineages, 7% divergence	2 lineages included with probable RI	Chapple et al. 2011b
Lampropholis mirabilis	2, with 4% divergence	Very similar for exons; treated as 1 lineage here	This paper
Lampropholis robertsi	4, with 7-9% divergence	3 lineages included, 1 with potential RI. Supported by multilocus sequencing. 2 southern lineages later recognised as separate species	Bell et al. 2010; Singhal et al. 2018
Liburnascincus coensis	2, with 9% divergence	2 lineages included here	This paper
Liburnascincus mundivensis	2, with 9% divergence	Sparse sampling; 2 lineages with probable RI	This paper
Lygisaurus absconditus	2 , non-monophyletic	Only the A lineage included	Couper et al. 2005; This paper
Lygisaurus aerata	2, with up to 10% divergence	Further sampling needed in north Cape York. Collapsed to 1 lineage here.	This paper
Lygisaurus foliorum	3, with 7–10% divergence	Paraphyletic with L. absconditus; further sampling needed. Collapsed to 1 lineage here	Couper et al. 2005
Lygisaurus laevis	1, with < 3% mtDNA divergence		This paper
Lygisaurus macfarlani	3, with 6-9% divergence and polyphyletic with other species	monophyletic for exons; 3 lineages included.	This paper
Lygisaurus malleolus	1, with < 3% divergence		This paper
Lygisaurus rimula	3, with 8-11% divergence	monophyletic for exons; 3 lineages included	Hoskin unpubl., this paper
Lygisaurus sesbrauna	2, with >5% divergence	More sampling needed. Collapsed to one lineage here.	Couper et al. 2005, This paper
Lygisaurus tanneri	2, polyphletic with malleolus	1 lineage included here	This paper
Lygisaurus zuma	2, with 3% divergence	More sampling needed; only 1 lineage included here.	This paper
Menetia alanae	2, with 10% divergence, paraphyletic with *M. concinna*	2 lineages included here; probable RI	This paper
Menetia greyii	>4 lineages with <10% divergence	Allozymes - Includes sexual and asexual forms, more sampling required; represented as one lineage here.	Adams et al. 2003
Morethia adelaidensis	1 with < 6% divergence	Extensive exon sequencing; 1 lineage for exons	Ivan et al., unpublished

Morethia boulengeri	2, with 6% divergence	2 lineages with extensive exon sequencing; only 1 lineage included here	Ivan et al., unpublished
Morethia butleri	2, with 4% divergence	2 lineages with extensive exon sequencing; only 1 lineage included here	Ivan et al., unpublished
Morethia lineoocellata	2, with >5%% divergence	Only 1 lineage included	Ivan et al., unpublished
Morethia obscura	1 with < 3% divergence	Extensive exon sequencing; 1 lineage for exons	Ivan et al., unpublished
Morethia ruficauda	2, with 7% divergence	Extensive exon sequencing; Multiple minor lineages from Kimberley collapsed into 1; 2 lineages included	Potter et al. 2019
Morethia storri	4, with 8% divergence	Extensive exon sequencing; Multiple minor lineages from Kimberley collapsed into 1; 2 lineages included with probable RI	Potter et al. 2019
Morethia taeniopleura	3, with > 5% divergence	Sparse sampling; only 1 lineage included	Ivan et al., unpublished
Proablepharus reginae	2, with <5% divergence	Extensive exon sampling; Only 1 lineage included	Potter et al. 2019
Proablepharus tenuis	7, with 7-12% divergence	Extensive exon sampling; all 7 lineages included; probable RI	Potter et al. 2019
Pseudemoia baudini	1 with < 5% divergence	Sparse sampling across range	This paper
Pseudemoia cryodrama	2, but introgressed from other species.	Extensive msat and exon sequencing; Species monophyletic for exons and very similar; collapsed to 1 here.	Haines et al. 2014, 2017; Haines & Potter unpublished
Pseudemoia entrecasteauxii	2, with introgression 10% divergence	Extensive msat and exon sequencing; Species is monophyletic for exons; 2 lineages included	Haines et al. 2014, 2017; Haines & Potter unpublished
Pseudemoia pagenstecheri	3, with 7-10% divergence	Extensive msat and exon sequencing; NE NSW lineage is highly divergent, making this species paraphyletic with other taxa; 3 lineages here – 1 with probable RI	Haines et al. 2014, 2017; Haines & Potter unpublished
Pseudoemoia rawlinsoni	1 with < 5% divergence	Sparse sampling across range	This paper
Pseudeomoia spenceri	3 with <5% divergence	Further sampling needed, treated as 1 lineage here	This paper
Saproscincus basiliscus	3, with 3-10% divergence	Extensive multilocus sequencing; nDNA replacement in 4th, highly divergent southern mtDNA lineages; 3 lineages here – extensive gene flow between N and C lineages	Moussalli et al. 2009, Singhal & Moritz 2012; Singhal & Bi 2017
Saproscincus challengeri	1 lineage with < 5% divergence		Moussalli, unpublished data
Saproscincus czechurai	3 lineages with 5-6% divergence	3 lineages included here	Moussalli et al. 2009
Saproscincus hannahae	1 lineage with < 5% divergence		Moussalli, unpublished data
Saproscincus lewisi	1 lineage with < 5% divergence		Moussalli et al. 2009
Saproscincus mustelinus	3 lineages with 2-4% divergence	3 lineages included here	Moussalli, unpublished data

Saproscincus rosei	3 lineages with 4% divergence	2 lineages included here	Moussalli, unpublished data
Saproscincus spectabilis	1 lineage with< 5% divergence		Moussalli, unpublished data
Saproscincus tetradactylus	2 lineages with 4% divergence	2 lineages included here.	Moussalli et al. 2009

Small-range species assumed to have one lineage: *Eroticoscincus graciloides*; *Carinascincus palfreymani*; *Carlia wandalthini*; *Harrisoniascincus zia*; *Liburnascincus artemis*, *L. scirtetis*; *Lygisaurus parrhasius*, *L. rococo*, *L. tanneri*; *Menetia concinna*; *Pygmaeascincus koshlandae*, *P. sadlieri*; *Saproscinus eungellensis*, *S. saltus*; *Techmarscincus jiggurru*.

Widespread species with insufficient evidence to resolve lineages and represented in phylogeny by one lineage: *Carinascincus coventryi*; *Menetia maini*, *M. surda*; *Pygmaeascincus timlowi.*

References

Adams, M. et al. The Australian scincid lizard Menetia greyii: a new instance of widespread vertebrate parthenogenesis. *Evolution* **57**, 2619-2627 (2003). https://doi.org/10.1111/j.0014-3820.2003.tb01504.x

Afonso Silva, A.C. et al. Validation and description of two new north-western Australian Rainbow skinks with multispecies coalescent methods and morphology. *PeerJ* **5**, *e3724* (2017). https://doi.org/10.7717/peerj.3724

Bell, R. C. et al. Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. *Mol. Ecol.* **19**, 2531-2544 (2010). https://doi.org/10.1111/j.1365-294X.2010.04676.x

Chapple, D. G. et al. Phylogeographic divergence in the widespread delicate skink (*Lampropholis delicata*) corresponds to dry habitat barriers in eastern Australia. *BMC Evol. Biol.* **11**, 1-18 (2011a). https://doi.org/10.1186/1471-2148-11-191

Chapple, D. G. et al. Biogeographic barriers in south-eastern Australia drive phylogeographic divergence in the garden skink, *Lampropholis guichenoti*. *J. Biogeog.* **38**, 1761-1775 (2011b). https://doi.org/10.1111/j.1365-2699.2011.02531.x

Chapple, D.G. et al. Phylogenetic relationships in the Eugongylini (Squamata: Scincidae): generic limits and biogeography. *Aust. J. Zool*. **70**, 165-203 (2023). https://doi.org/10.1071/ZO23007

Cliff, H. B. et al. Persistence and dispersal in a Southern Hemisphere glaciated landscape: the phylogeography of the spotted snow skink (*Niveoscincus ocellatus*) in Tasmania. *BMC Evol. Biol.* **15**, 1-13 (2015). https://doi.org/10.1186/s12862-015-0397-y

Couper, P.J. et al. Skinks currently assigned to *Carlia aerata* (Scincidae: Lygosominae) of north-eastern Queensland: a preliminary study of cryptic diversity and two new species. *Aust. J. Zool.* **53**, 35-49 (2005). https://doi.org/10.1071/ZO04010

Dissanayake, D. S. et al. Lineage diversity within a widespread endemic Australian skink to better inform conservation in response to regional-scale disturbance. *Ecol. Evol.* **12**, e8627 (2022). https://doi.org/10.1002/ece3.8627

Dolman, G. & Moritz, C. A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution* **60**, 573–82 (2006). https://doi.org/10.1111/j.0014-3820.2006.tb01138.x

Donnellan, S. et al. Systematics of the *Carlia* '*fusca*' complex (Reptilia: Scincidae) from northern Australia. *Zootaxa* **2227**, 1–31 (2009). https://doi.org/10.11646/zootaxa.2227.1.1

Dubey, S., & Shine, R. Evolutionary diversification of the lizard genus *Bassiana* (Scincidae) across southern Australia. *PLoS One* **5**, e12982 (2010). https://doi.org/10.1371/journal.pone.0012982

Haines, M.L. et al. Phylogenetic evidence of historic mitochondrial introgression and cryptic diversity in the genus *Pseudemoia* (Squamata: Scincidae). *Mol. Phylogenet. Evol.* **81**, 86-95 (2014). https://doi.org/10.1016/j.ympev.2014.09.006

Haines, M. L. et al. A complex history of introgression and vicariance in a threatened montane skink (*Pseudemoia cryodroma*) across an Australian sky island system. *Conserv. Genet.* **18**, 939-950 (2017). https://doi.org/10.1007/s10592-017-0945-7

Horner, P. Systematics of the snake-eyed skinks, *Cryptoblepharus* Wiegmann (Reptilia: Squamata: Scincidae)–an Australian-based review. *Beagle Suppl.* ***3***, 21-198 (2007).

Horner, P. & Adams, M.A. Molecular-systematic assessment of species boundaries in Australian *Cryptoblepharus* (Reptilia: Squamata: Scincidae): A case study for the combined use of allozymes and morphology to explore cryptic biodiversity. *Beagle Suppl.* **3**, 1–21 (2007).

Kreger, K. M. et al. Phylogeographic parallelism: Concordant patterns in closely related species illuminate underlying mechanisms in the historically glaciated Tasmanian landscape. *J. Biogeog.* **47**, 1674-1686 (2020). https://doi.org/10.1111/jbi.13831

Melville, J. & Swain, R. Evolutionary correlations between escape behaviour and performance ability in eight species of snow skinks (*Niveoscincus*: Lygosominae) from Tasmania. *J. Zool.* **26**, 79-89 (2003). https://doi.org/10.1017/S0952836903003984

Moussalli, A. et al. A mitochondrial phylogeny of the rainforest skink genus *Saproscincus*, Wells and Wellington (1984). *Mol. Phylogenet. Evol.* **34**, 190-202, (2005). https://doi.org/10.1016/j.ympev.2004.08.022

Moussalli, A. et al. Variable responses of skinks to a common history of rainforest fluctuation: concordance between phylogeography and palaeo-distribution models. *Mol. Ecol.* **18**, 483–499 (2009). https://doi.org/10.1111/j.1365-294X.2008.04035.x

Potter, S. et al. Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. *Mol. Ecol*. **25**, 1367-1380 (2016). https://doi.org/10.1111/mec.13546

Potter, S. et al. Pleistocene climatic changes drive diversification across a tropical savanna. *Mol. Ecol.* **27**, 520-532 (2018). https://doi.org/10.1111/mec.14441

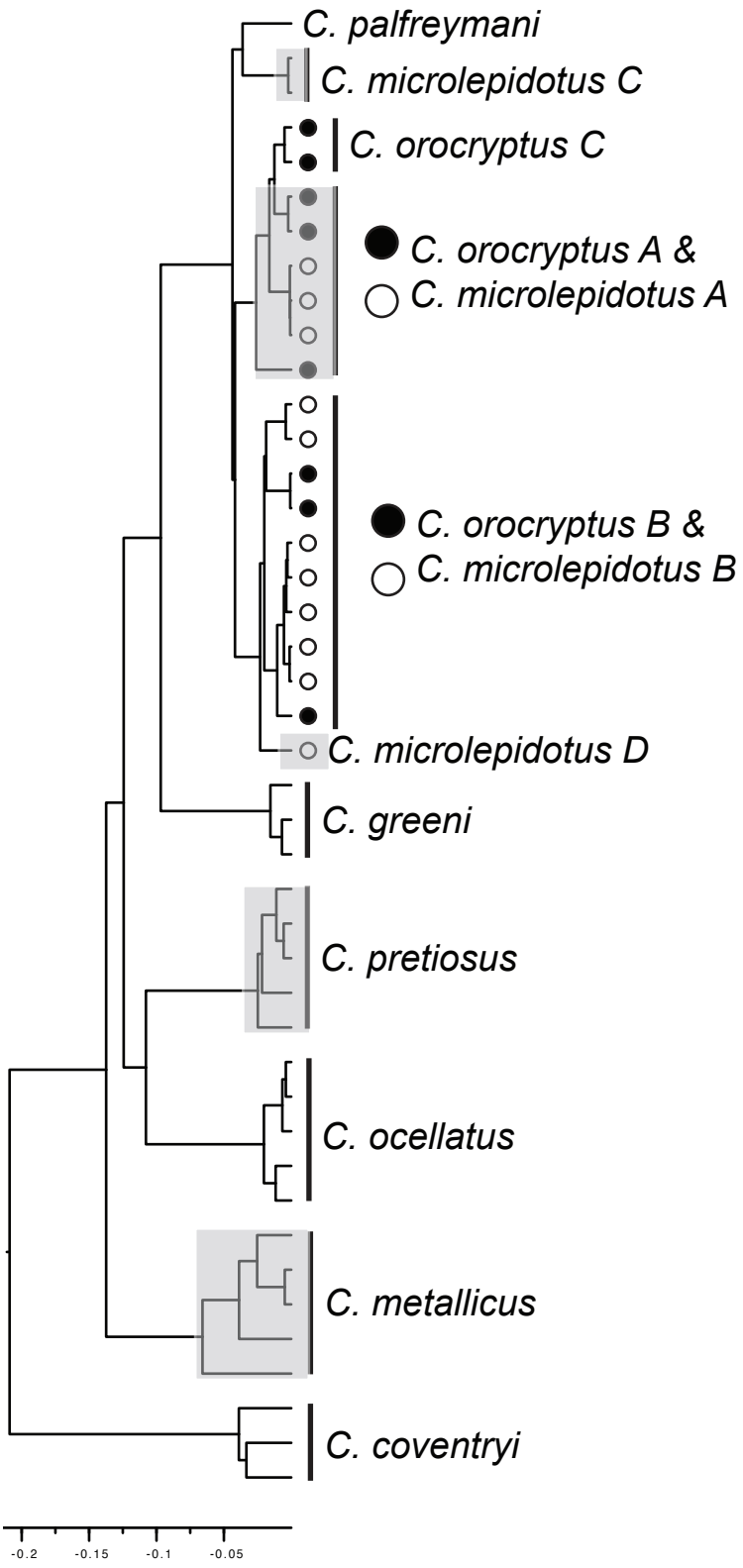
Potter, S. et al. Contrasting scales of local persistence between monsoonal and arid biomes in closely related, low-dispersal vertebrates. *J. Biogeog.* **46**, 2506-2519 (2019). https://doi.org/10.1111/jbi.13698

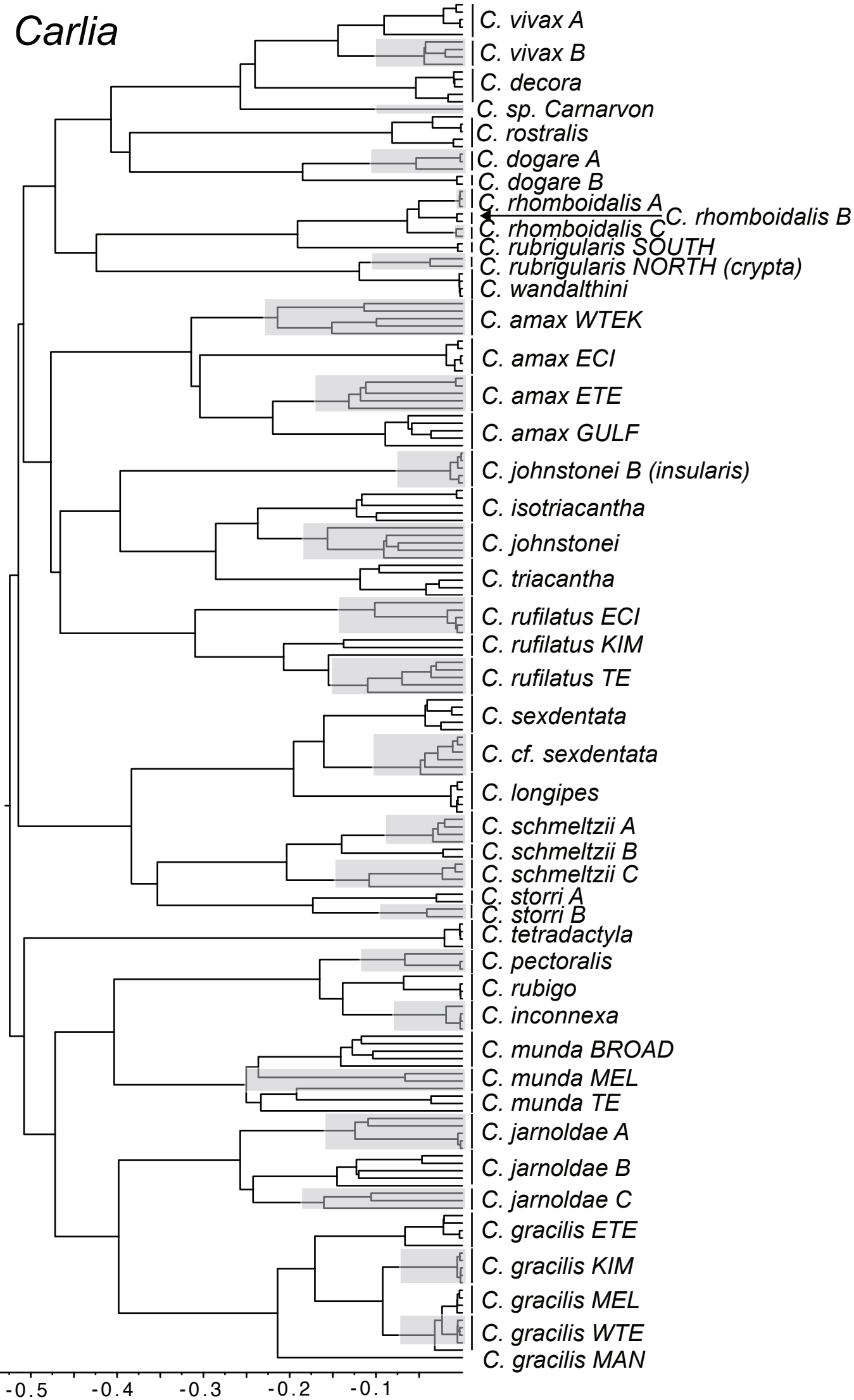
Singhal, S. & Moritz, C. Testing hypotheses for genealogical discordance in a rainforest lizard. *Mol. Ecol.* **21**, 5059-5072 (2012). https://doi.org/10.1111/j.1365-294X.2012.05747.x

Singhal, S. & Bi, K. History cleans up messes: the impact of time in driving divergence and introgression in a tropical suture zone. *Evolution* **71**, 1888-1899 (2017). https://doi.org/10.1111/evo.13278

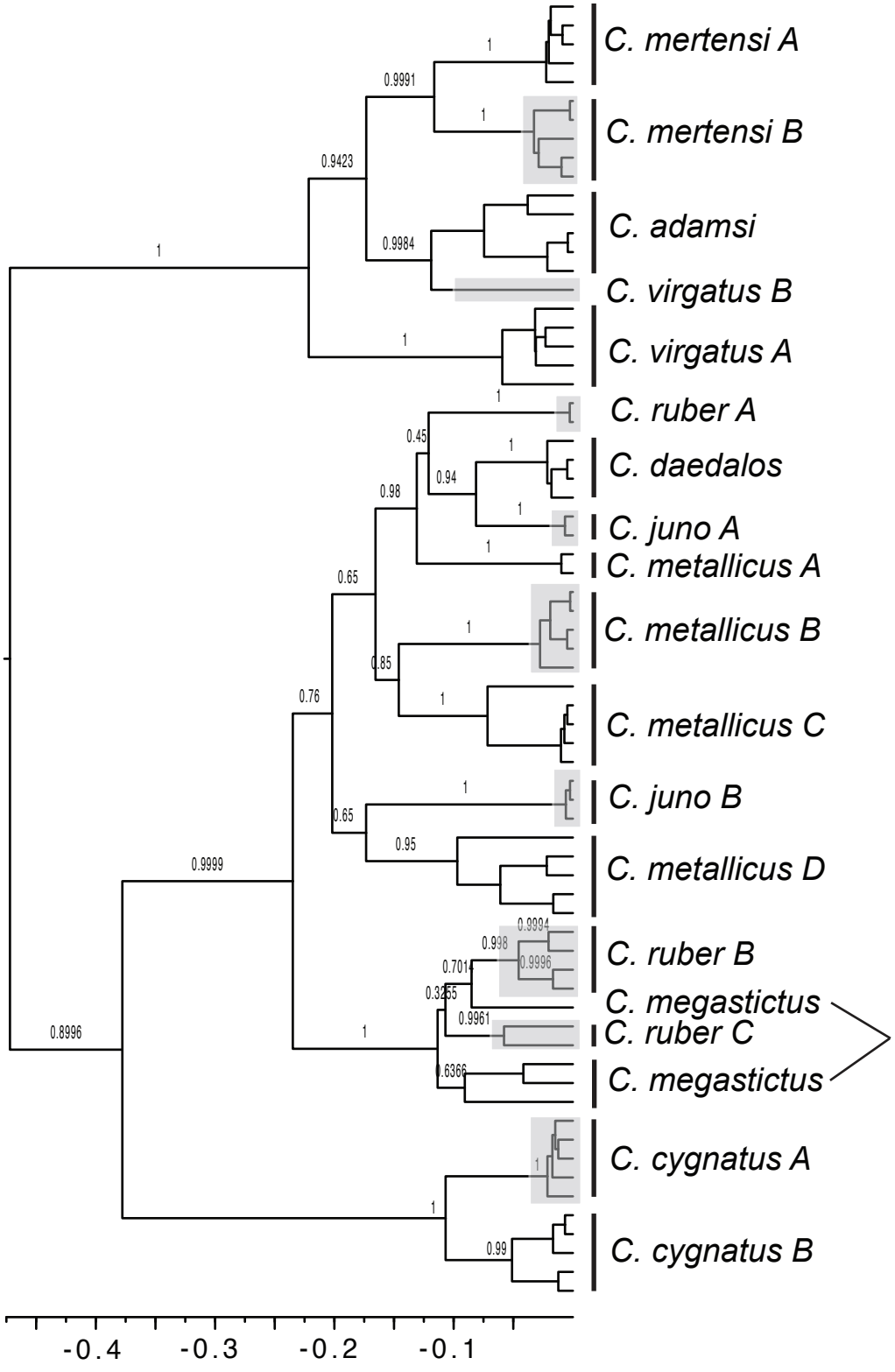
Singhal, S. et al. A framework for resolving cryptic species: a case study from the lizards of the Australian wet tropics. *Syst. Biol.* **67**, 1061-1075 (2018). https://doi.org/10.1093/sysbio/syy026

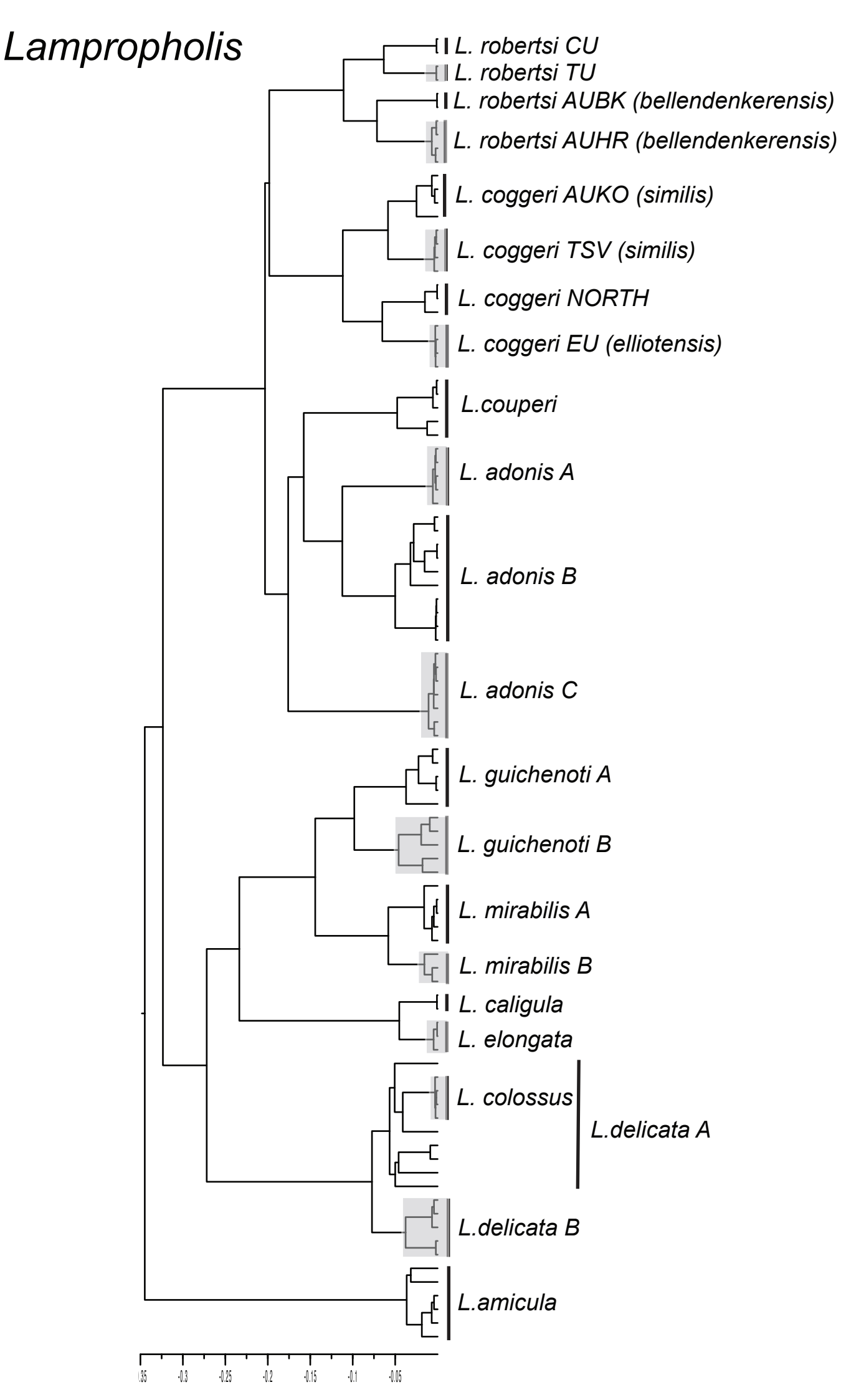
Carinascincus



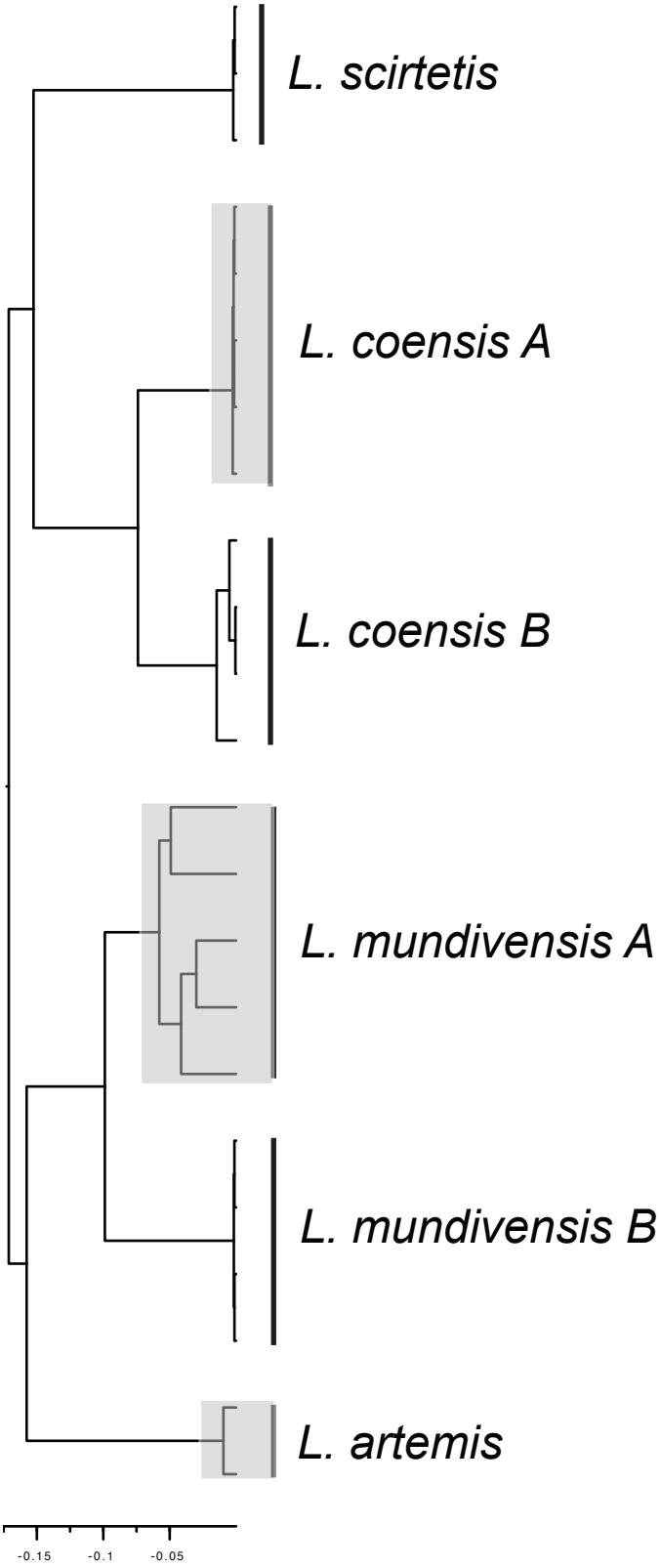

Carlia
C. vivax A
C. vivax B
C. decora
C. sp. Carnarvon
C. rostralis
C. dogare A
C. dogare B
C. rhomboidalis A
C. rhomboidalis B
C. rhomboidalis C
C. rubrigularis SOUTH
C. rubrigularis NORTH (crypta)
C. wandalthini
C. amax WTEK
C. amax ECI
C. amax ETE
C. amax GULF
C. johnstonei B (insularis)
C. isotriacantha
C. johnstonei
C. triacantha
C. rufilatus ECI
C. rufilatus KIM
C. rufilatus TE
C. sexdentata
C. cf. sexdentata
C. longipes
C. schmeltzii A
C. schmeltzii B
C. schmeltzii C
C. storri A
C. storri B
C. tetradactyla
C. pectoralis
C. rubigo
C. inconnexa
C. munda BROAD
C. munda MEL
C. munda TE
C. jarnoldae A
C. jarnoldae B
C. jarnoldae C
C. gracilis ETE
C. gracilis KIM
C. gracilis MEL
C. gracilis WTE
C. gracilis MAN
-0.5
-0.4
-0.3
-0.2
-0.1

Cryptoblepharus

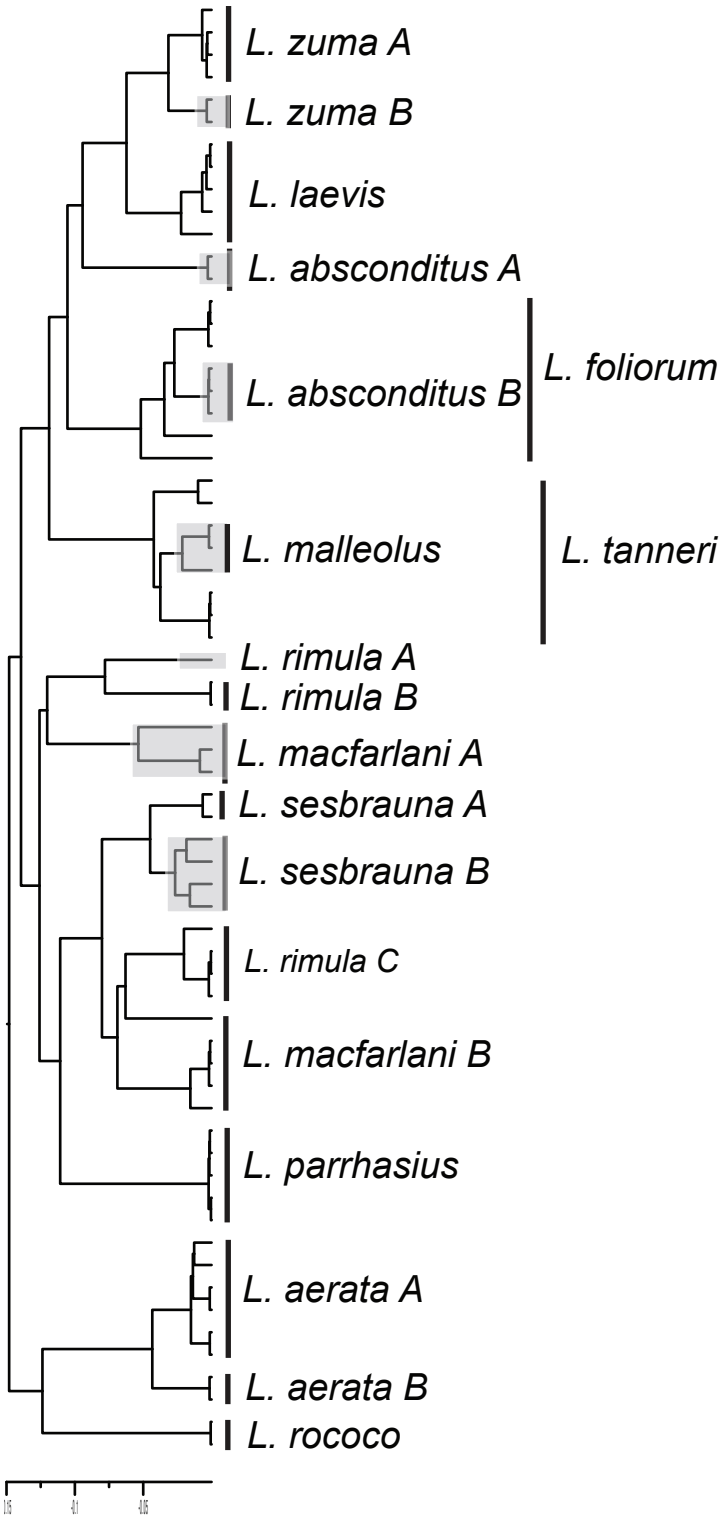



Lampropholis
L. robertsi CU
L. robertsi TU
L. robertsi AUBK (bellendenkerensis)
L. robertsi AUHR (bellendenkerensis)
L. coggeri AUKO (similis)
L. coggeri TSV (similis)
L. coggeri NORTH
L. coggeri EU (elliotensis)
L.couperi
L. adonis A
L. adonis B
L. adonis C
L. guichenoti A
L. guichenoti B
L. mirabilis A
L. mirabilis B
L. caligula
L. elongata
L. colossus
L.delicata A
L.delicata B
L.amicula
.35
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05

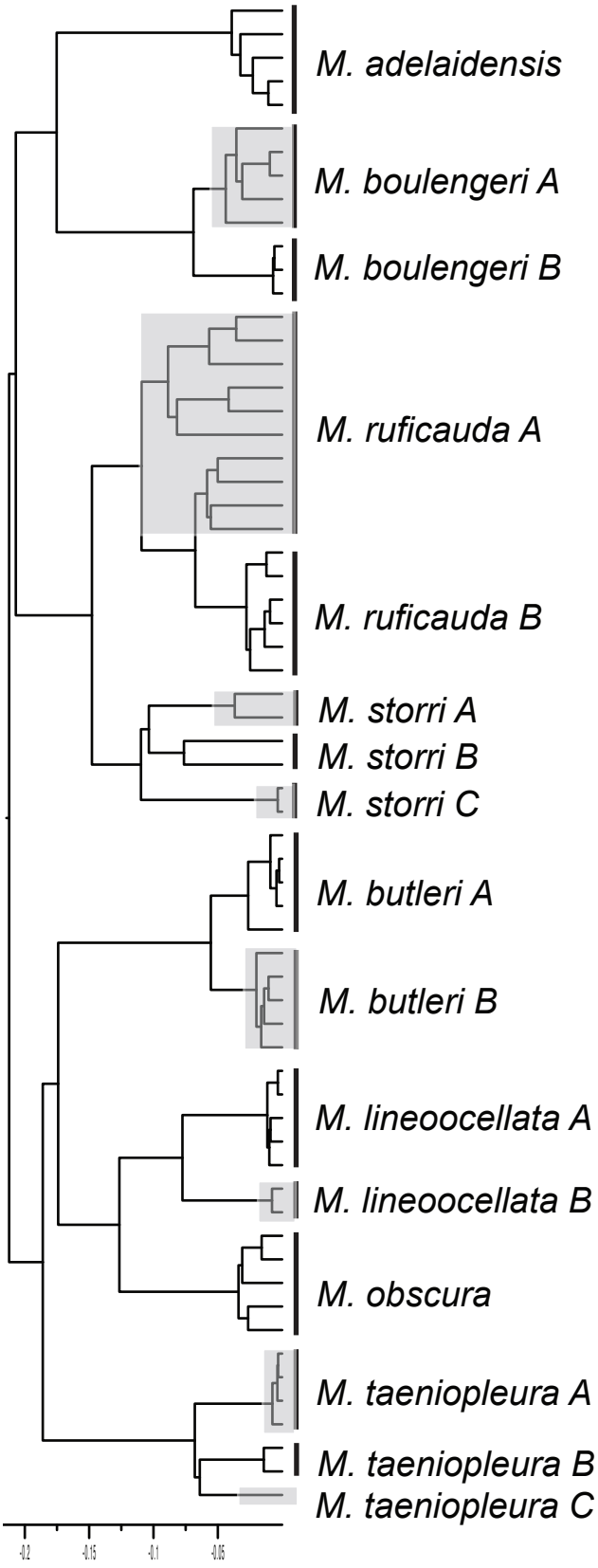
Liburnascincus

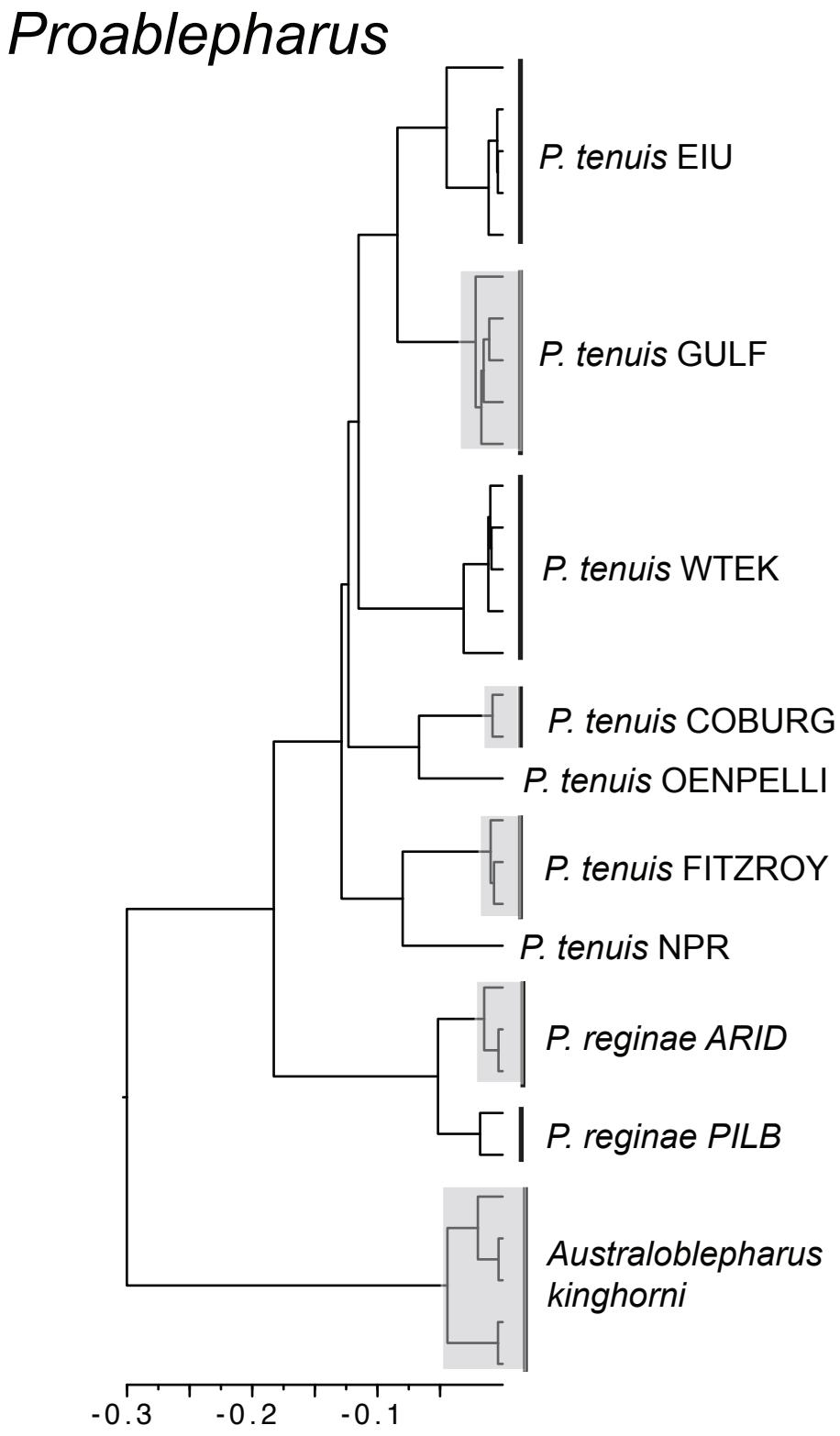


Lygisaurus

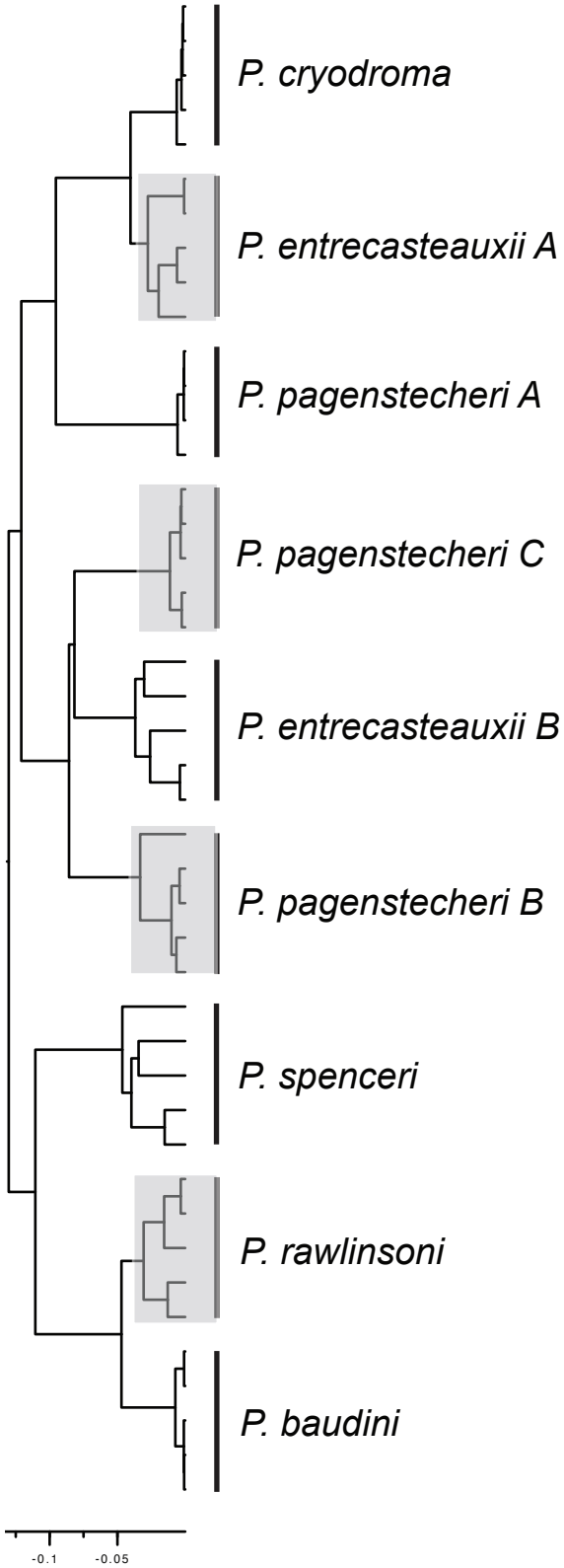

L. zuma A
L. zuma B
L. laevis
L. absconditus A
L. absconditus B
L. foliorum
L. malleolus
L. tanneri
L. rimula A
L. rimula B
L. macfarlani A
L. sesbrauna A
L. sesbrauna B
L. rimula C
L. macfarlani B
L. parrhasius
L. aerata A
L. aerata B
L. rococo

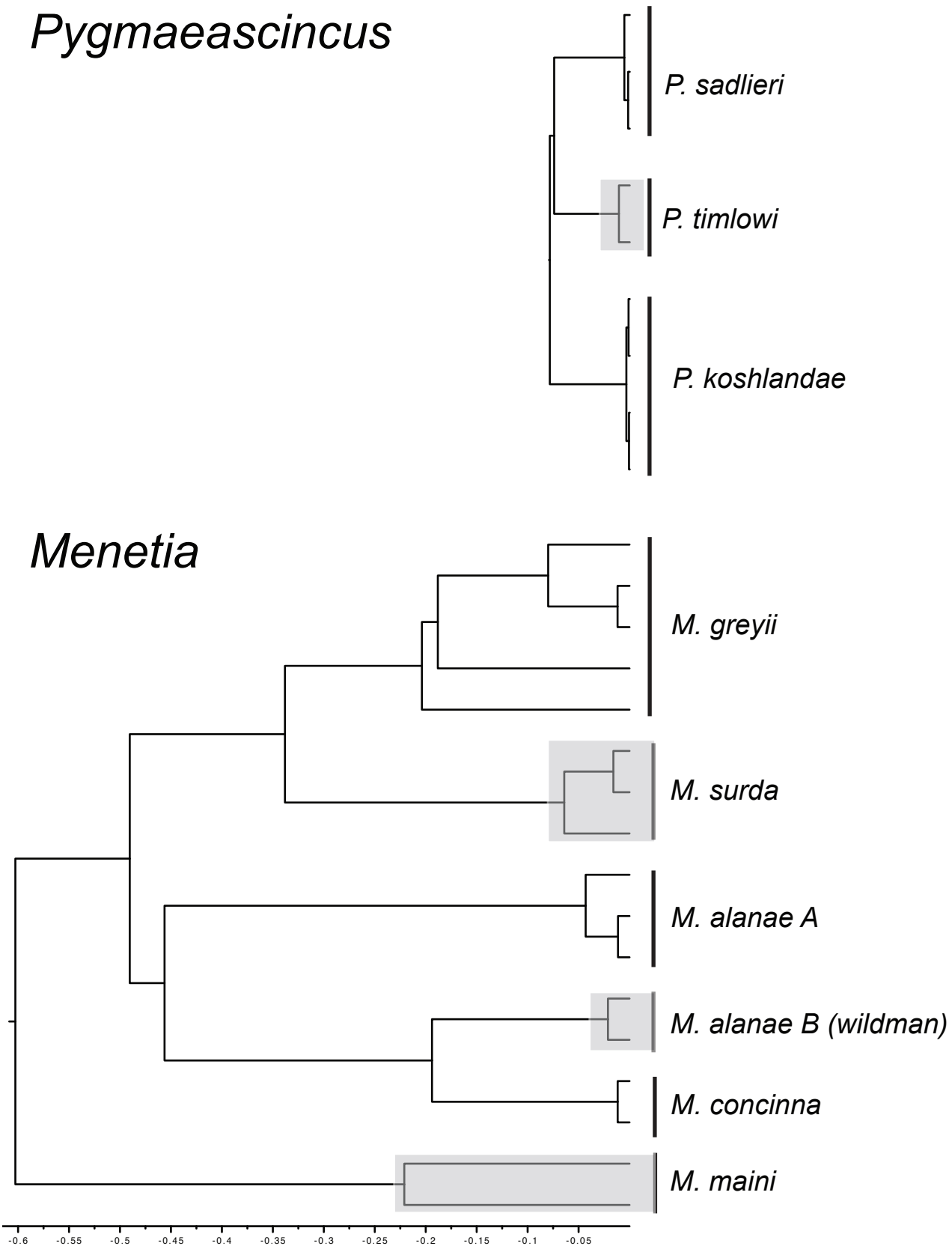
Morethia


M. adelaidensis
M. boulengeri A
M. boulengeri B
M. ruficauda A
M. ruficauda B
M. storri A
M. storri B
M. storri C
M. butleri A
M. butleri B
M. lineoocellata A
M. lineoocellata B
M. obscura
M. taeniopleura A
M. taeniopleura B
M. taeniopleura C
-0.2
-0.15
-0.1
-0.05

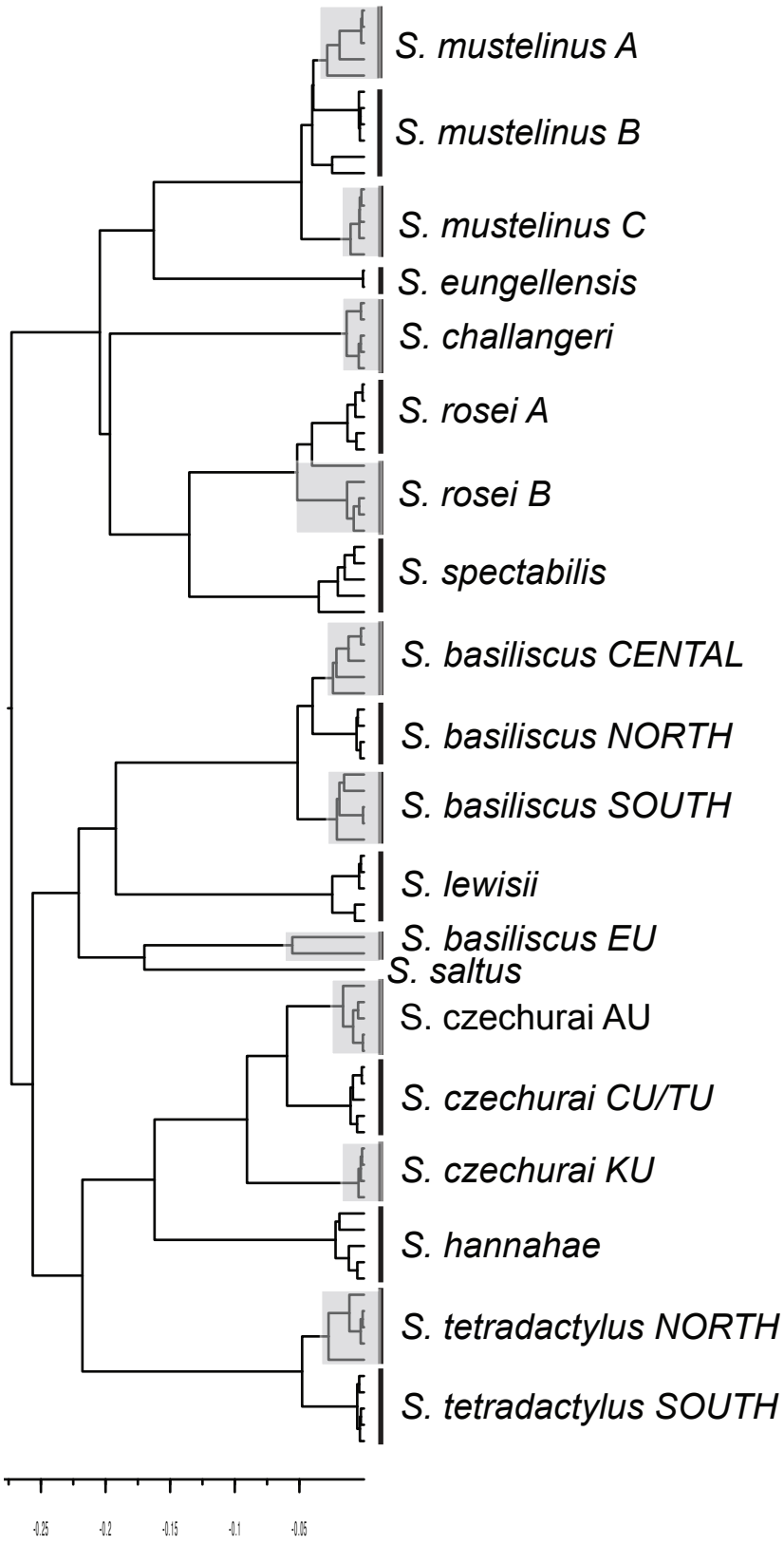

Proablepharus
P. tenuis EIU
P. tenuis GULF
P. tenuis WTEK
P. tenuis COBURG
P. tenuis OENPELLI
P. tenuis FITZROY
P. tenuis NPR
P. reginae ARID
P. reginae PILB
Australoblepharus kinghorni
-0.3
-0.2
-0.1

Pseudemoia


P. cryodroma
P. entrecasteauxii A
P. pagenstecheri A
P. pagenstecheri C
P. entrecasteauxii B
P. pagenstecheri B
P. spenceri
P. rawlinsoni
P. baudini
-0.1
-0.05


Pygmaeascincus
P. sadlieri
P. timlowi
P. koshlandae
Menetia
M. greyii
M. surda
M. alanae A
M. alanae B (wildman)
M. concinna
M. maini
-0.6
-0.55
-0.5
-0.45
-0.4
-0.35
-0.3
-0.25
-0.2
-0.15
-0.1
-0.05

Saproscincus



Supporting Text

Data generation: sampling, sequencing, and phylogenetic estimation

Sampling: We present genomic (exon sequence) data and analyses for 410 skink samples from 126 species, including two outgroups (*Emoia nigra*, *Oligosoma lichenigera*; Chapple et al. 2023; Supplementary file, 1.2). For many species, these data were derived from previous studies of phylogenetics or comparative phylogeography. A summary of prior and evidence for presence of intraspecific lineages, and sampling details are provided in Supplementary file, 1.1. Samples were collected from museums as mostly liver tissues stored frozen or in ethanol. Additional samples (liver or tail tips) were obtained from fieldwork by us or collaborators and stored at 4C in RNAlater. We sampled from unique sites where possible spanning the species distribution. Where evidence on intraspecific diversity was not already published, we screened geographically dispersed samples of species for mitochondrial diversity as below (results in Supplementary file, 1.1).

Mitochondrial DNA sequencing: DNA was extracted using a salting-out method (Sunnucks & Hales 1996). We sequenced the mitochondrial ND4 gene using the primers ND4 light (5'-3') CACCTATGACTACCAAAAGCTCATGTAGAAGC (Arèvalo et al. 1994) and Leu3 heavy (5'-3') GAATTAGCAGTTCTTT(AG)TG (Stuart-Fox et al. 2002). PCRs were performed following the methods of Potter et al. (2016) in 25uL reactions. In brief, ~100ng of genomic DNA was used, 2.5 uL PCR Buffer, 0.2mM dNTPs, 2.5 mM MgCl2, 10 pmol each of primers, and 0.5 U Taq DNA Polymerase (Invitrogen). In general, PCR conditions included initial denaturation (95 °C for 3 min), and touchdown cycling conditions (95 °C for 30 sec, 53-50 °C for 30 sec, 72 °C for 45 sec), followed by a final extension of 72 °C for 10 min. PCR products were visualised on a 1.5% agarose gel and then purified using 5 uL PCR product, 3uL double-distilled water, together with 0.4 uL Exonuclease 1 (New England BioLabs) and 1.6 uL of Shrimp Alkaline Phosphatase (In Vitro Technologies Pty. Ltd), incubated at 37 °C for 45 min followed by 80 °C for 15 min. These samples were then sequenced in 20 uL sequencing reactions which included 13.5 uL double-distilled water, 0.8 uL BigDye Terminator v3.1 (Applied Biosystems), 4.5 uL 5x sequencing buffer, 3.2 pmol of primer together with 1 uL of the purified PCR product. Sequencing reactions were run for 25 cycles (94 °C for 5 sec, 50 °C for 10 sec, 60 °C for 4 min) and then purified using sodium acetate (see Pepper et al. 2006). Samples were eluted in 20 uL of HiDi formamide and sequenced directly on an ABI 3100 DNA Analyzer (Applied Biosystems) at the Australian National University.

Sequences were visualised, edited and aligned using Geneious v6.0.5 (Kearse et al. 2012). Initial analyses involved identifying candidate lineages based on the mtDNA. Here we generated mitochondrial trees in BEAST v2.6.6 (Bouckaert et al. 2019), to show time-dependent phylogenies to enable assessment of divergence between genera. We calculated Dxy between mitochondrial lineages using the PopGenome R package (Pfeifer et al. 2014) and included missing data. Average Dxy was estimated by dividing the total number of variable sites by the number of valid sites in the alignment. Subsamples of individuals were used in the exon capture experiments based on divergent and distinct evolutionary lineages from the mtDNA results, evidence of monophyly from initial exon-based phylogenies and geography and prior phylogeographic analyses (Supplementary file, 1.1).

Exon sequencing: We generated a phylogenomic dataset using exon-capture sequencing. This approach uses oligonucleotide probes to hybridize and capture a set of nominated sequences in a DNA sample, resulting in a library that is enriched with the target DNA. The

probe set used here targets 3320 exons and has been described in detail previously (Bragg et al. 2015, 2018). We used the protocol of Meyer and Kircher (2010) to prepare sequencing libraries, with modifications described in Bi et al. (2012). The resulting DNA libraries were pooled at equimolar concentrations, usually in batches of 56 at a time. The pooled libraries were hybridized with the probe kit (SeqCap EZ Developer Library; NimbleGen) following the protocol supplied by the manufacturer. We note that the hybridization mix was modified, and contained 1.2 μg of pooled sample DNA, 5 μg of skink Cot-1 DNA, and a set of barcode-specific blocking oligonucleotides (1000 pmol). The Cot-1 DNA was made by isolating Cot-1 DNA from *Lampropholis coggeri* (Singhal & Bi 2017), following the method described by Trifonov et al. (2009). Polymerase Chain Reactions (PCRs) were performed to enrich the libraries following hybridization 'capture' (17 cycles). The resulting DNA libraries were sequenced using an Illumina HiSeq instrument (Biomolecular Resource Facility, Australian National University). A qPCR was conducted to evaluate enrichment of targets and de-enrichment of non-target DNA as described in Potter et al. (2016) and following methods of Bi et al. (2012). Sequencing reads are available in the Short Read Archive (NCBI BioProject PRJNA289283).

Bioinformatics: Raw reads were cleaned using a workflow described by Singhal (2013), which removes duplicates, merges overlapping reads, and trims adaptors and poor-quality bases. We assembled the reads for each sample using a workflow described in detail by Bragg et al (2015). This workflow finds reads homologous to each exon (blastall 2.2.26, expectation= 1E-9, program = blastx, Altshul et al. 1990) and assembles these using velvet (version 1.2.08, Zerbino & Birney 2008; assemblies with $K = 31$, 41, 51, 61 and 71). For each exon, resulting contigs were merged with cap3 (version 08/06/13, with parameters: -o 20 –p 99; Huang & Madan 1999) and flanking introns were removed using exonerate (2.2.0, Slater & Birney 2005). Where multiple contigs assembled, we identified the one that was putatively orthologous to the target using a best reciprocal BLAST hit (blastall 2.2.26). We then mapped the clean reads to the assembled exons for each sample (bowtie2, version 2.2.4, Langmead & Salzberg 2012), identified heterozygous sites, and phased them using overlapping sequence reads (Genome Analysis Toolkit, version 3.3–0-g37228af, McKenna et al. 2010). Sequence processing and alignment closely followed approaches used in Bragg et al. 2018. Sites where the genotype quality score (GQ) was less than 20 were replaced with an 'N'. Sample identity and quality was assessed in preliminary phylogenetic analyses and via checks for abnormally high observed heterozygosity. We then selected samples (typically 2) per species, where possible from geographically remote localities. For each locus, we took a single haplotype sequence for each sample (‘h0’) and aligned the sequences from different samples using MACSE (v1.01b, Ranwez et al. 2011). Codons were removed from alignments if they contained a site with greater than 20% missing data (trimAl v1.4.rev15, Capella-Gutiérrez et al. 2009).

***Phylogenetic estimation*:** We used StarBEAST2 (Ogilvie et al. 2017) to estimate the phylogeny of the eugongylus skinks. This package infers a species tree under the multispecies coalescent model in a Bayesian framework. To compare phylogenies based on distinguishing between gene and species-trees versus concatenation methods, and inferences derived from them, we analysed the same loci but concatenated using BEAST (v2.5.0, Bouckaert et al. 2019). Compared with species-tree approaches, concatenation can substantially overestimate tip lengths relative to true divergence times (Ogilvie et al. 2016), affecting downstream macroevolutionary inferences. This would be of concern in our study which compares properties of trees with and without intraspecific lineages. However, the fully Bayesian species tree method implemented in StarBEAST2 is also highly computationally intensive,

and so could not feasibly be run on a dataset consisting of hundreds of samples and >1000 loci. We therefore followed a strategy of choosing subsets of loci. To do this, we selected two independent and disjoint subsets of 100 randomly selected loci across sites and individuals from the larger data set (e.g. 1268 loci in Ivan et al. 2019).

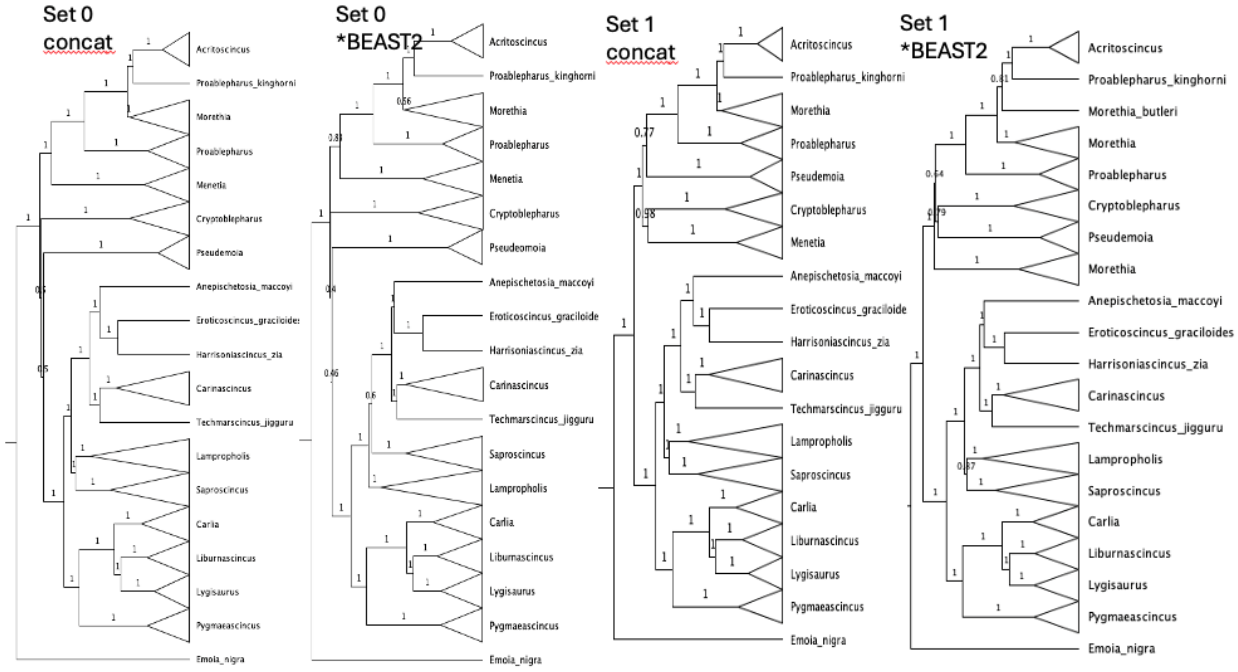
Within each locus, sequences were treated as outliers and removed when it improved the linear regression R^2 correlation coefficient between the distance matrix for that locus and a representative distance matrix by at least 0.05. The representative distance matrix was constructed by taking the median distance across all loci for a given specimen pair. Loci from the larger dataset were only considered for inclusion in the abovementioned subsets if their correlation coefficient was at least 0.33 after outlier removal.

For each subset of 100 loci, it was still not possible to estimate the species tree for the full set of 400 samples. We therefore adopted a divide-and-conquer approach (see also Alvarez et al. 2023). We began by using a set of summary coalescent approaches (ASTRAL, NJ, SVD quartets) with bootstrapping across loci to identify well supported clades and divergent taxa within them. This resulted in the assignment of samples and taxa to 13 clades for which species trees were then estimated separately using StarBEAST2 under a strict clock and an HKY substitution model (with N = 4 to 31 taxa). A "backbone" tree with N = 40 taxa was similarly analysed. For each clade, and for the backbone tree, six independent chains were run for 1.6 billion iterations, with the first 30% of each chain discarded as burn-in. The post-burn-in chains were combined and thinned to 900 posterior samples.

Posterior distributions from each of the 13 clades and the backbone tree were then combined into a single set of trees. For a given clade, the posterior sample with the youngest root node was linearly rescaled to have the same height as the matching node in the posterior sample with the youngest matching node. The matching node was then replaced with the clade tree. This process was repeated for the second youngest node and so on, and for every clade until the full set of posterior samples were merged into a sample of supertrees. Concatenation analyses were almost identical to StarBEAST2 in all the above respects, except that molecular sequences were assumed to all evolve along the same tree.

These analyses included several non-Australian species of *Cryptoblepharus* (see Blom et al. 2019) which were removed from phylogenies prior to analyses reported below. Key results are described in the main text and summary phylogenies across loci and methods, which are very similar, are shown in Figure 1 below. The BEAST xmls files and sequence data, and the resulting posterior distributions of trees can be retrieved from a Zenodo archive

Figure 1. *Comparison of genus level relationships inferred across two independent sets of 100 loci (labelled Set 0 and Set 1) and using two methods; concatenation with BEAST (concat) and the StarBEAST2 species-tree approach (*BEAST2). Details can be seen in Supplementary file, 2.2.*



Evolutionary inference: supplementary methods and results

Associations between cryptic lineage diversity, distribution and environment: We wanted to know if the number of cryptic lineages within species was associated with properties of those species or the environments in which they live. We therefore tested associations among species between the number of intra-specific lineages observed and potential predictor variables including the size of their geographic range, the latitude at the centre of the range, and mean annual temperature. These species attributes were estimated via locality records and associated georeference and climatic data harvested (with obvious outliers removed) from the Atlas of Living Australia in June 2020 (see Supplementary file, 1.4). To test these associations, we used phylogenetic glm to account for the phylogenetic relatedness among the species. We implemented these analyses in the R package 'phylolm' (function 'phyloglm'; Ho & Ane 2014). We used Poisson models, because the number of lineages is expressed in integer values. In the main text, we report relationships for the StarBEAST2 phylogeny from one set of loci – results were the same qualitatively for the species phylogenies from the second set of loci. We also checked that these associations were robust to error in the estimation of the species trees, by performing the tests of association for 100 samples from the posterior distribution. Results are illustrated in Figure 2 and significant associations are reported in the main text.

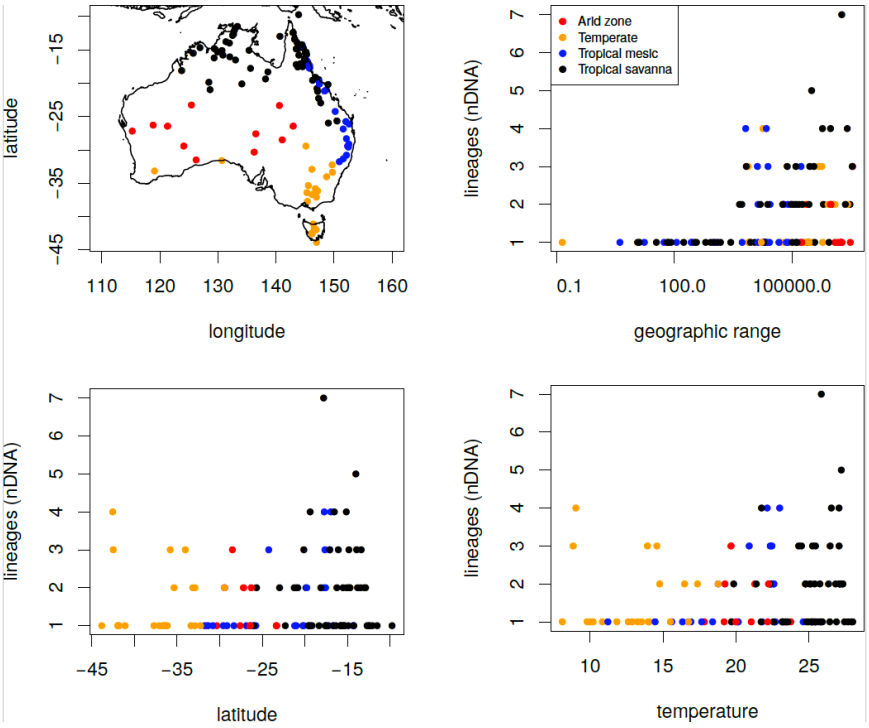


Figure 2. Location of species geographic centroids coloured by biome type and associations between number of lineages per taxonomic species and independent variables.

Diversification: We wanted to know whether there were differences in inferred diversification dynamics between trees at the lineage and species level. To test this, we fit simple diversification models to trees, and estimated parameters using maximum likelihood. these analyses were implemented in R package RPANDA (using function 'fit_bd'; Morlon et al. 2016). We fit models in which the rate of lineage birth (speciation) and lineage death (extinction) could be constant or could vary as an exponential function of time. This meant we estimated the parameters of several different models for each tree, and these models varied in the number of inferred parameters, so we compared the fit of the data to the different models using a corrected AICc (Table 1).

Table 1. *Summary of AICc values for three different models, applied to four different phylogenetic trees (StarBEAST2 for locus sets 0 and 1). The models that were estimated had:*

(i) constant (C) birth and death, (ii) exponential (exp) birth and constant death, and (iii) constant birth and exponential death.

Tree	birth, death = C	birth = exp death = C	birth = C death = exp
lineage, set 0	1023.517	1025.517	1025.579
species, set 0	715.3073	712.0693	717.4082
lineage, set 1	1040.112	1041.992	1041.966
species, set 1	725.4467	723.0097	727.5474

At the lineage level, models with constant birth and death had smaller values of AICc than models in which rates of birth and death varied exponentially. The constant birth rates were 0.204 and 0.196 (for trees based on locus sets 0 and 1, respectively), and both models had lineage death rates approximately equal to 0. At the species level, models with varying birth rates and constant death rates provided a better fit to the data than models with constant birth and death rates. In these models, lineage birth rates were relatively small in the present, 0.117 and 0.114 (species trees based on locus sets 0 and 1, respectively), increasing towards the past with coefficients of 0.0388 and 0.0346 (locus sets 0 and 1, respectively). This would result in birth rates at the crown of the tree (depth 23.21) of 0.288 and 0.254 (locus sets 0 and 1, respectively). We note that we also fit models with birth and death rates that varied exponentially, but the estimated models appeared numerically unstable, and we attributed this to overfitting, and did not consider this combination further.

It has been observed that different combinations of lineage birth and death rates can result in the same 'lineages through time' curves (Louca & Pennell 2020). To remedy this, an approach of estimating 'pulled' speciation rates has been suggested, which infers the slope of the lineages through time curve at different times since the present. We estimated pulled speciation rates for both lineage and species trees, each over seven time intervals (Fig. 3). We implemented these analyses in R package castor (Louca & Doebeli 2017). Pulled speciation rates for lineage and species trees were very similar for deeper intervals in the trees. Nearer the present (ancestor depths < 5), the trees pruned to a single lineage representing each species exhibited a reduced pulled speciation rate, compared to the trees that included all intraspecific lineages (Fig. 3).

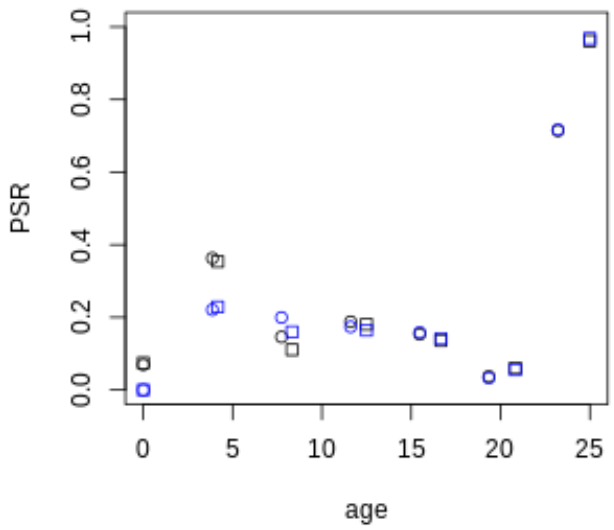


Figure 3. Pulled speciation rates estimated for trees at the lineage (black) and species (blue) levels. This was performed for StarBEAST2 phylogenies estimated using two locus sets (set 0, circles; set 1, squares).

Branch specific rates of diversification:
As a complement to the above analyses of diversification dynamics, we compared rates of diversification at the tips using branch-specific models between species-level and lineage-level phylogenies using the CLaDS2 software (Maliet et al. 2019; Maliet & Morlon 2022). This method allows detection of shifts in rates with high sensitivity and estimates the following parameters: alpha – the deterministic change in speciation rate; sigma – the stochasticity in inheritance of rates from parent to daughter lineages, m – the ratio of daughter/parent rates (m = alpha X exp(sigma^2/2); and epsilon – the rate of turnover. These analyses were conducted across trees including all lineages (lins), trees including all lineages predicted from divergence times to substantial reproductive isolation (RI) and trees including only taxonomically recognised species (SP). To account for uncertainty, each of these analyses where repeated across posterior distributions of phylogenies estimated for both sets of 100 loci and for concatenated (BEAST) and species-tree (StarBEAST2) methods. Fig. 4 illustrates how inferred rates vary across taxa and between phylogenies including all lineages versus species alone. The Julia code to run the models, as well as the code used to compare the results, together with the full results are in Supplementary file, 2.3.

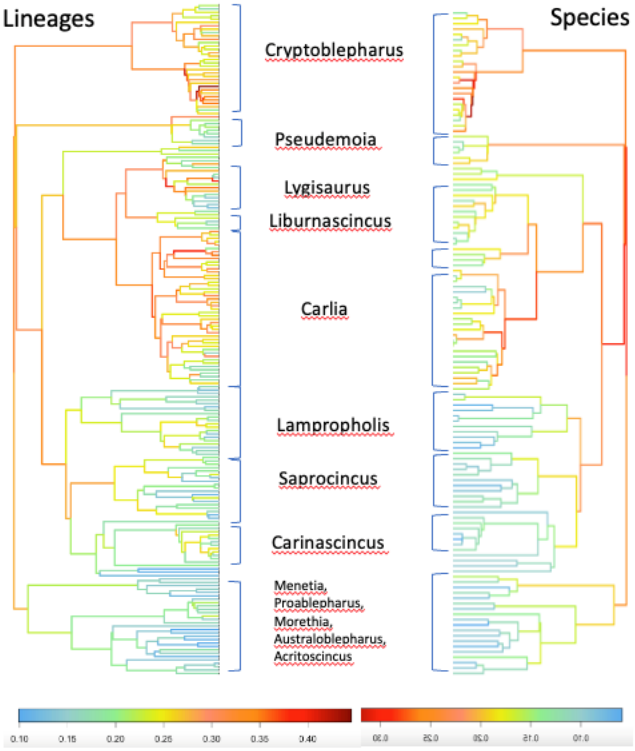


Figure 4. Visualisation of estimates diversification rates on StarBEAST (locus set 1) trees from CLaDS2. Note differences in scaling.

Simulations of cryptic lineage diversity: We performed simulations of lineage and phenotypic evolution and used these to study the evolution of intraspecific lineages. The key link here is that lineages can be evolutionarily diverged in the simulations, but can only be recognised as different species when they exceed a threshold level of phenotypic divergence.

Our model of lineage divergence is based on the 'protracted birth death' (PBD) model, which was simulated for different parameterizations using the Gillespie algorithm (Gillespie 1976, following Cutter and Gray 2016). In classical models of lineage diversification, new lineages originate by speciation, and are lost through extinction. These processes occur at parameterized rates, denoted birth (b) and death (d). If these rates are constant, the net rate of increase in species over time is equal to $b - d$. However, many observed phylogenies are not consistent with this simple model of diversification, and instead show a decline in the rate $b - d$ over time. This seems to imply that the speciation rate has declined over time, or the extinction rate increased, or both. The protracted birth-death (PBD) model explicitly acknowledges that speciation is not immediate. Here, new lineages are formed in a 'nascent' state, and are converted to the state of being 'good' lineages at a parameterized rate (at a rate here denoted as c). This model is capable of describing the commonly observed decline in net speciation rates towards the present, while recognizing that, at any given time, there are lineages on a reversible path towards speciation.

If we want to understand how intraspecific lineages are distributed across phylogenies, or how (genetically) diverged lineages might remain unrecognized and phenotypically 'cryptic,' the protracted birth-death model seems an appropriate starting point. However, there is an incompatibility between the PBD model and a model of the formation of cryptic lineages. That is, we do not expect that 'nascent' species (in the PBD) would always be morphologically cryptic, and it is possible for good species (in the PBD) to remain morphologically cryptic. We therefore developed an approach to model the occurrence of cryptic lineages in a diversifying clade, based on the PBD. In this model (Fig. 5), two lineages are deemed 'cryptic' if they are separated by less than a threshold distance (m_t) in a 'morphospace.' This morpho-space is generated by letting a number ($n_t = 5$) of traits evolve randomly on the tree. That is, at each time step (size 0.0001, between 0 and 1), each trait of each lineage jumps by a value (δ) that is sampled from a normal distribution (mean=0, variance =1; a 'random walk'). Newly formed species inherit the trait values of their parent species, and their traits begin changing independently. This means that, in general, two lineages that share an ancestor recently tend to be separated by a smaller distance in morphospace (and have a greater chance of remaining 'cryptic') than lineages that share an ancestor long ago. Here, we make 'nascent' species an important exception. Part of the motivation for representing 'nascent' species was the expectation that incompletely diverged lineages might remain connected by gene flow. We assume this gene flow could cause the trait values of a nascent species to covary with those of its parent over time. To represent this, we assign a probability (p) that, at a given time step, a nascent species will receive the same trait value jump (δ) as its parent species, as opposed to receiving its own randomly generated value. Note that there is a complicated question here of how to deal with trait values when a parent of multiple nascent species has gone extinct. In the current implementation, we calculate trait jumps for extinct species, and continue to allow them to be inherited by their descendent nascent lineages. This is done to reflect the possibility that two species that were nascent with respect to a common parent continue to covary with each other due to continuing gene flow between each other.

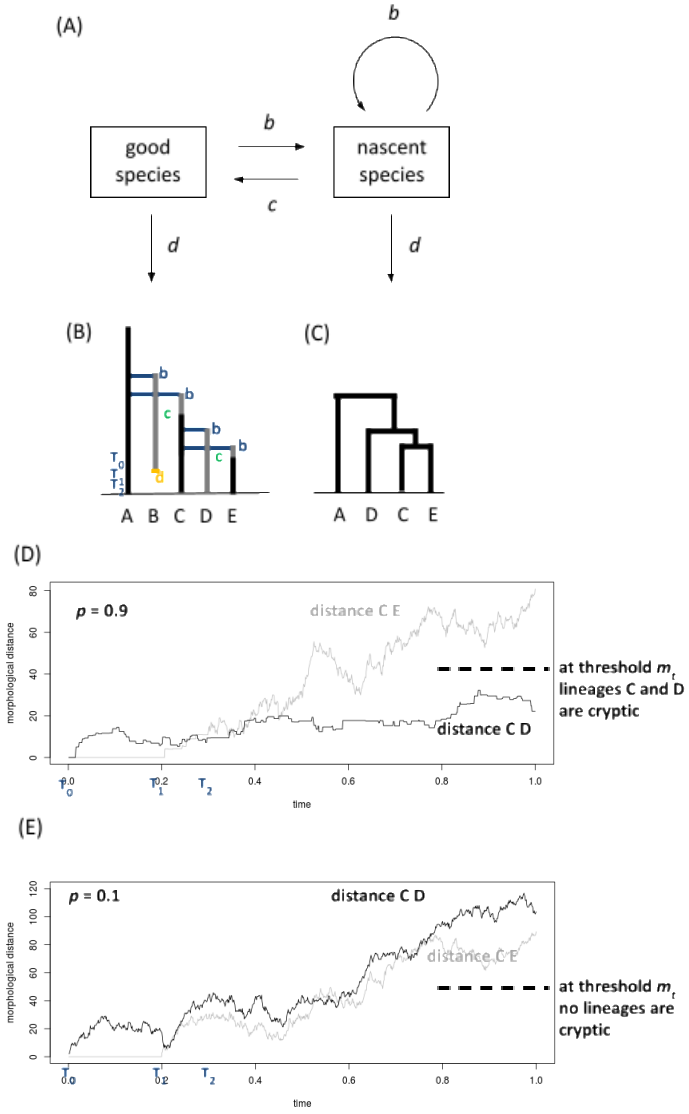


Figure 5. Diagram illustrating simulation approaches. Radiations of lineages were simulated using a modified protracted birth death model (panel A). Nascent species were produced (by good and nascent species) according to birth parameter, *b*. All lineages die according to death parameter, *d*. Nascent species convert to good species according to parameter *c*. An illustrative example is shown (B-E). In panel B, we that a good species, A, was initialized. Nascent species B was born from A, and subsequently became extinct. Another nascent lineage, C, was born from A, and by chance, rapidly completed speciation. Two lineages were born from C: D at time T_0, and C at time T_1. Lineage C completed speciation soon after birth (time T_2). Morphological trait variation was simulated, and used to identify cryptic lineages. The examples (panels D and E) show how morphological crypsis is more likely if lineages are nascent, and when *p* is large.

In sum then, we have implemented a model that can be used to simulate a tree, along with a list of extant pairs of lineages with morphological distances that remain smaller than some threshold value. This represents a set of groups of intraspecific lineages within a morphologically determined taxonomy. The simulations therefore provide opportunities to ask questions about how we might expect such cryptic diversity to be distributed across trees and clades, and what kinds of processes might promote the formation of unrecognized or cryptic species.

More specifically, we performed 100,000 simulations while varying lineage birth rate (*b*; 4.2-6.0, as per RPANDA analyses), lineage death rate (*d*; 0-1), the rate of completion of speciation (*c*; 0-12), the probability that a nascent lineage inherits changes in trait values from its parent lineage (*p*; 0.5-1), and the morphospace threshold (m_t; 5-80). All these model parameter values were sampled from uniform distributions over the nominated ranges. These distributions were determined by some preliminary simulations over broader ranges, as well as the analyses presented above (Table 1) in which a likelihood approach was used to infer constant rates of speciation (here, *b*) and extinction (here, *d*). We identified a subset of simulations with the most similar outcomes to the eugongylus data using Approximate Bayesian Computation implemented in R package abc (Csilléry et al. 2012). To do this, we

used four summary statistics. One summary statistic was simply the number of extant lineages in the trees. A second summary statistic described the number of intraspecific lineage pairs in the trees, defined in simulated trees as the number of pairs that did not have a morphospace distance exceeding the threshold value. The remaining two summary statistics were inspired by the normalised Lineages Through Time (nLTT) approach of Janzen et al. (2015), which estimates the difference between two trees in the shape of the area under the curves of their (normalized) lineages through time plots. We calculated another statistic that was analogous to the nLTT, but estimating the difference between two trees in curves for their accumulation of intraspecific (cryptic) lineage pairs. In sum, then, we fit the simulations to data using four summary statistics: two that described the magnitude of the extant numbers of lineages and of cryptic lineage pairs, and two that described the normalized shapes for the accumulation of lineages through time, and for intraspecific lineage pairs through time. Formally, we identified the simulations that best fit the empirical tree (the posterior distribution of samples) using Approximate Bayesian Computation, implemented in R package abc, with rejection sampling. For most analyses we set the ABC tolerance value to 0.01, and considered the parameter values in the 1000 trees in the posterior distribution (Fig. 6). In particular, we note that the simulations that best fit the empirical Eugongylinae tree tended to have large values for *p*. This parameter was the probability that a nascent species receives the same trait value jump as its parent species, at each time step. The values of this parameter for simulations in the posterior distribution (Fig. 6D) tended to be large relative to the prior distribution, which was distributed uniformly between 0.5 and 1. To further examine the consequences of allowing *p* to vary, we performed a set of simulations similar to the above, but with *p* fixed at 0, to represent a diversification process in which nascent lineages began immediately to diverge from their parent lineage in phenotype values. We analysed these simulations analogously, with Approximate Bayesian Computation. We then took the trees from the posterior distributions of these two analyses (with variable *p*, and $p = 0$, respectively), to consider the shape of the accumulation of intraspecific lineage diversity in the two different models. Here we used an abc tolerance value of 0.001, to permit the plotting of tractable (100) number of trees for each model. We see that the models with $p = 0$ that best fit the data tend not to account well for the recent and rapid accumulation of intraspecific lineages (main text, Fig. 6C).

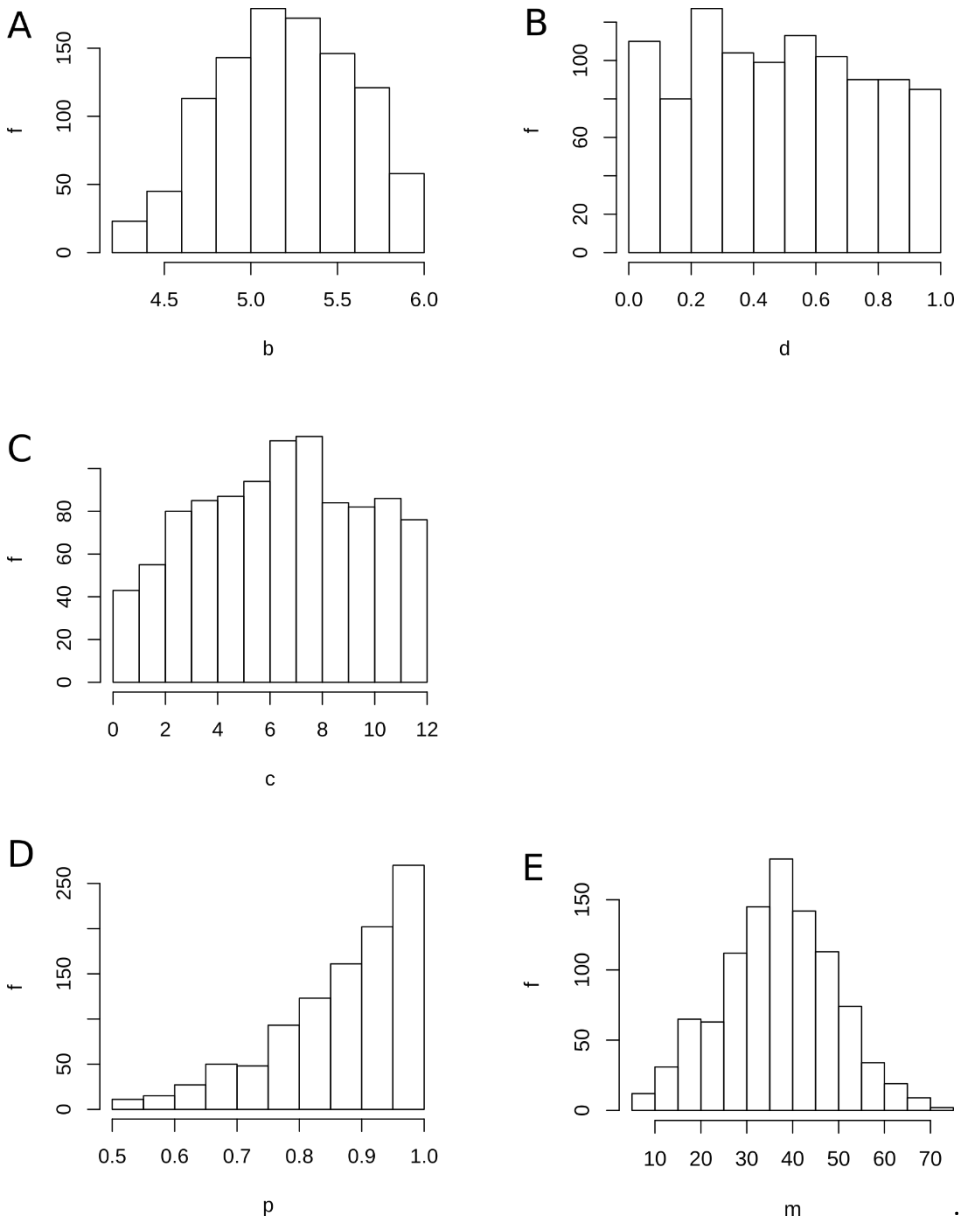
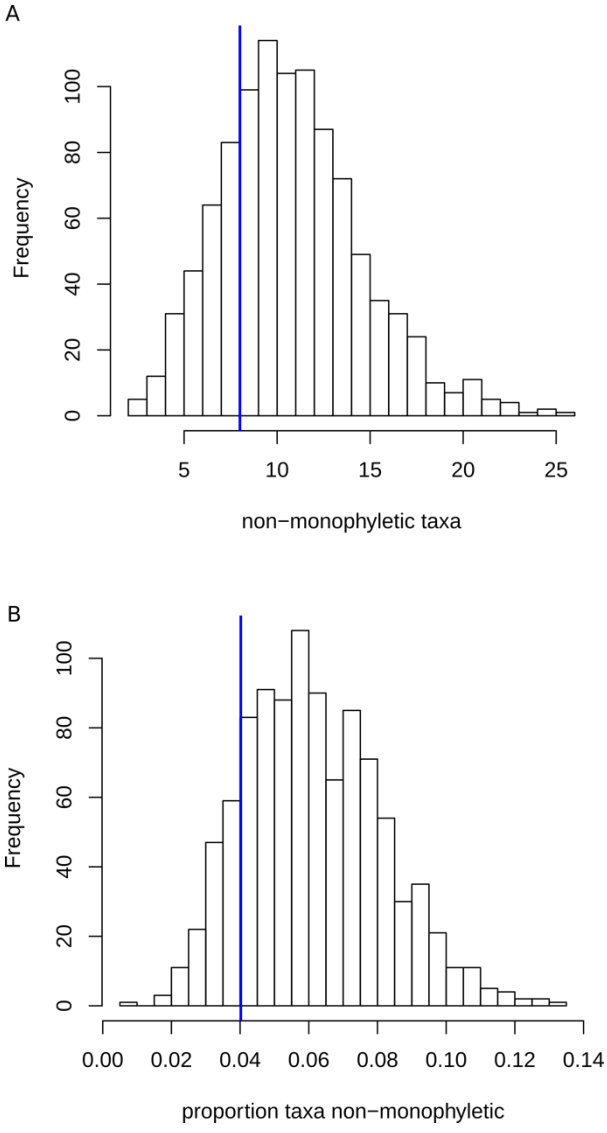


Figure 6. Parameter values for 1000 simulations in the posterior distribution when our evolutionary diversification model was fit to data from the Eugongylinae tree. Frequency distributions are shown for five parameter values.

In both the empirical tree and for simulated trees, it was possible to identify species that were not monophyletic. This is where a species contained multiple lineages, and one of those lineages was more closely related to a lineage in another species than to a lineage in the same species. There were eight such cases in the Eugongylinae tree. In the simulated trees, 'species' were defined based on morphology as sets of lineages that were indistinguishable from each other according to a threshold difference in morphospace. Similar to the empirical tree, the simulated 'species' were non-monophyletic where they contained a lineage that was more recently diverged from a lineage belonging to another species than a conspecific lineage. In Fig. 7, we see that simulated trees did not contain substantially different numbers of non-monophyletic species to the empirical tree, either when expressed as (A) the number, or (B) the proportion, of species that were non-monophyletic.



C)

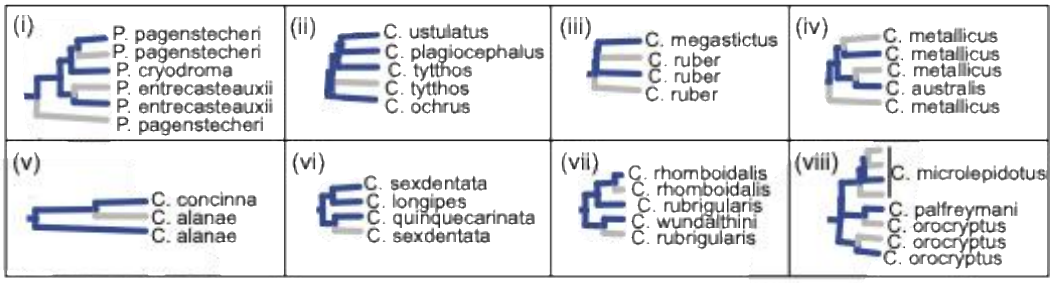


Figure 7. The frequency distributions of the (A) number and (B) proportion of species in simulated trees that were not monophyletic. The 'species' were defined in simulated trees based on differences among lineages in morphological traits. The 1000 trees represented here were from the posterior distribution of an approximate bayesian computation analysis. The blue lines are the corresponding values for the Eugongylinae empirical tree. (C) Cases of non-monophyletic species within the Eugongylus phylogeny for Pseudemoia (i), Cryptoblepharus (ii – iv), Menetia (v), Carlia (vi – vii), Carinascincus (viii).

References

Altshul, S.F. et al. Basic local alignment search tool. *J. Mol. Biol*. **215**, 403–1 (1990). https://doi.org/10.1016/S0022-2836(05)80360-2

Álvarez-Carretero, S. et al. A species-level timeline of mammal evolution integrating phylogenomic data. *Nature* **602**, 263–267 (2022). https://doi.org/10.1038/s41586-021-04341-1

Arèvalo E, Davis SK, Sites JW Jr (1994) Mitochondrial-DNA sequence divergence and phylogenetic-relationships among eight chromosome races of the *Sceloporus Grammicus* complex (Phrynosomatidae) in central Mexico. *Syst. Biol.* **43**, 387–418.

Bi, K. et al. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics* **13**, 403 (2012). https://doi.org/10.1186/1471-2164-13-403

Blom, M.P.K. et al. Habitat preference modulates trans-oceanic dispersal in a terrestrial vertebrate. *Proc. R. Soc. B.* **286**, 20182575 (2019). http://doi.org/10.1098/rspb.2018.2575

Bouckaert, R. et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comp. Biol*. **15**, e1006650 (2019). https://doi.org/10.1371/journal.pcbi.1006650

Bragg, J.G. et al. Exon capture phylogenomics: efficacy across scales of divergence. *Mol. Ecol. Res*. **16**, 1059-1068 (2016). https://doi.org/10.1111/1755-0998.12449

Bragg, J.G. et al. Phylogenomics of a rapid radiation: the Australian rainbow skinks. *BMC Evol. Biol*. **18**, 1-12 (2018). https://doi.org/10.1186/s12862-018-1130-4

Capella-Gutiérrez, S. et al. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–3 (2009). https://doi.org/10.1093/bioinformatics/btp348

Chapple, D.G. et al. Phylogenetic relationships in the Eugongylini (Squamata: Scincidae): generic limits and biogeography. *Aust. J. Zool*. **70**, 165-203 (2023). https://doi.org/10.1071/ZO23007

Csilléry, K. et al. abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecol. Evol.* **3**, 475-479 (2012). https://doi.org/10.1111/j.2041-210X.2011.00179.x.

Cutter, A.D. & Gray, J.C. Ephemeral ecological speciation and the latitudinal biodiversity gradient. *Evolution* **70**, 2171-2185 (2016). https://doi.org/10.1111/evo.13030

Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys*. **22**, 403-434 (1976). https://doi.org/10.1016/0021-9991(76)90041-3

Ho, L.S.T. & Ane, C. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Syst. Biol*. **63**, 397-408 (2014). https://doi.org/10.1093/sysbio/syu005

Huang, X. & Madan, A. CAP3: a DNA sequence assembly program. *Genome Res*. **9**, 868–77 (1999). https://doi.org/ 10.1101/gr.9.9.868

Ivan, J. et al. Temperature predicts the rate of molecular evolution in Australian Eugongylinae skinks. *Evolution* **76**, 252-261 (2022). https://doi.org/10.1111/evo.14342

Janzen, T. et al. Approximate Bayesian Computation of diversification rates from molecular phylogenies: introducing a new efficient summary statistic, the nLTT. *Methods Ecol. Evol.* **6**, 566-575 (2015). https://doi.org/10.1111/2041-210X.12350

Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012). https://doi.org/10.1093/bioinformatics/bts199

Langmead, B. & Salzberg, S. Fast gapped-read alignment with bowtie 2. *Nat. Methods* **9**, 357–9 (2012). https://doi.org/10.1038/nmeth.1923

Louca, S., & Doebeli, M. Efficient comparative phylogenetics on large trees. *Bioinformatics* **34**, 1053-1055 (2018). https://doi.org/10.1093/bioinformatics/btx701

Louca, S. & Pennell, M.W. Extant timetrees are consistent with a myriad of diversification histories. *Nature* **580**, 502-505 (2020). https://doi.org/10.1038/s41586-020-2176-1

Maliet, O. et al. A model with many small shifts for estimating species-specific diversification rates. *Nat. Ecol. Evol*. **3**, 1086-1092 (2019). https://doi.org/10.1038/s41559-019-0908-0

Maliet, O., & Morlon, H. Fast and accurate estimation of species-specific diversification rates using data augmentation. *Syst. Biol*. **71**, 353-366 (2022). https://doi.org/10.1093/sysbio/syab055

McKenna, A.H. et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. **20**, 1297–303 (2010). https://doi.org/10.1101/gr.107524.110

Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*. 2010.6 (2010). https://doi.org/10.1101/pdb.prot5448

Morlon, H. et al. RPANDA: an R package for macroevolutionary analyses on phylogenetic trees. *Methods Ecol. Evol*. **7**, 589-597 (2016). R package version 1.4, https://CRAN.R-project.org/package=RPANDA. https://doi.org/10.1111/2041-210X.12526

Ogilvie, H.A. et al. Computational performance and statistical accuracy of *BEAST and comparisons with other methods. *Syst. Biol*. **65**, 381-396 (2016). https://doi.org/10.1093/sysbio/syv118

Ogilvie, H.A. et al. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. *Mol. Biol. Evol*. **34**, 2101-2114 (2017). https://doi.org/10.1093/molbev/msx126

Pepper, M. et al. Molecular phylogeny and phylogeography of the Australian *Diplodactylus stenodactylus* (Gekkota; Reptilia) species-group based on mitochondrial and nuclear genes reveals an ancient split between Pilbara and non-Pilbara *D. stenodactylus*. *Mol. Phylogenet. Evol.* **41**, 539–555 (2006). https://doi.org/10.1016/j.ympev.2006.05.028

Pfeifer B. et al. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol*. **31**, 1929-1936 (2014). https://doi.org/10.1093/molbev/msu136

Potter, S. et al. Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. *Mol. Ecol*. **25**, 1367-1380 (2016). https://doi.org/10.1111/mec.13546

Ranwez, V. et al. MACSE: multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS One* **6**, e22594 (2011). https://doi.org/10.1371/journal.pone.0022594

Singhal S. *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. *Mol. Ecol. Resour*. **13**, 403–16 (2013). https://doi.org/10.1111/1755-0998.12077

Singhal, S. & Bi, K. History cleans up messes: the impact of time in driving divergence and introgression in a tropical suture zone. *Evolution* **71**, 1888-1899 (2017). https://doi.org/10.1111/evo.13278

Slater, G.S.C. & Birney, E. Automated Generation of heuristics for biological sequence comparison. *BMC Bioinform*. **6,** 31 (2005). https://doi.org/10.1186/1471-2105-6-31

Stuart-Fox, D.M. et al. A molecular phylogeny of rainbow skinks (Scincidae: *Carlia*): taxonomic and biogeographic implications. *Aust. J. Zool.* **50**, 39–51 (2002). https://doi.org/10.1071/ZO01051

Sunnucks, P. & Hales, D.F. Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Mol. Biol. Evol.* **13**, 510–24 (1996). https://doi.org/10.1093/oxfordjournals.molbev.a025612

Trifonov, V.A. et al. FISH with and without COT1 DNA. In: Liehr T, editor. Fluorescence In Situ Hybridization (FISH): Application Guide. p. 99–109 (Berlin: Springer; 2009).

Zerbino, D.R. & Birney, E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–9 (2008). https://doi.org/ 10.1101/gr.074492.107