Recent Diversification Rates in North American Tiger Beetles Estimated from a Dated mtDNA Phylogenetic Tree

Timothy G. Barraclough and Alfried P. Vogler*†*

*Department of Biological Sciences and NERC Centre for Population Biology, Imperial College at Silwood Park, Ascot, Berkshire; and †Department of Entomology, The Natural History Museum, Cromwell Road, London

Species-level phylogenies derived from DNA sequence data provide a tool for estimating diversification rates and how these rates change over time, but to date there have been few empirical studies, particularly on insect groups. We use a densely sampled phylogenetic tree based on mitochondrial DNA to investigate diversification rates in the North American tiger beetles (genus *Cicindela*). Using node ages estimated from sequence data and calibrated by biogeographical evidence, we estimate an average per-lineage diversification rate of at least 0.22 ± 0.08 species/Myr over the time interval since the most recent colonization that led to a radiation within the continent. In addition, we find evidence for a weak, recent increase in the net diversification rate. This is more consistent with a late Pleistocene increase in the speciation rate than with a constant rate of background extinction, but the results are sensitive to the dating method and taxon sampling. We discuss practical limitations to phylogenetic studies of diversification rates.

Introduction

A key goal for understanding the evolution of biological diversity is to explain variation in diversification rates. Phylogenetic studies have been critical for identifying shifts in net diversification rates and for correlating biological factors with species richness in sister clades (Cracraft 1985; Sanderson and Donoghue 1996; Barraclough, Vogler, and Harvey 1998). Yet, extant diversity is the result of a dynamic evolutionary process determined by the relative rates of speciation and extinction and changes in these rates over time. Therefore, a complete investigation of the factors determining diversity requires a more detailed consideration of the effects on speciation and extinction (Ribera, Barraclough, and Vogler 2001). Methods have been developed for estimating speciation and extinction rates separately, on the basis of the lineages-through-time approach (Hey 1992; Harvey, May, and Nee 1994; Nee et al. 1994; Avise 2000; Nee 2001), and potentially for quantifying how these rates have changed over time (Kubo and Iwasa 1995; Wollenberg, Arnold, and Avise 1996; Pybus and Harvey 2000). But these methods require an estimate of phylogeny for all species in a clade and information on the relative timing of cladogenesis on the basis of DNA sequences or other evidence (Barraclough and Nee 2001). To date, there have been few applications of the approach because phylogenetic trees of the required resolution are still rare.

One area that should benefit from this approach is the determination of how diversification rates have been affected by major climatic changes, in particular by climatic fluctuations during the Pleistocene. Many authors have argued that recurrent ice ages over the last 2.5 Myr increased speciation rates by promoting founder events (Hewitt 1999) and the divergence of populations in isolated glacial refugia (Haffer 1969). Key evidence for this hypothesis has been that pairwise DNA sequence divergences between closely related taxa often support a high frequency of species origins during the Pleistocene (Brower 1994; Hackett 1996; Roy 1997). But other authors have argued that speciation rates declined during the Pleistocene (Zink and Slowinski 1995), possibly because of increased mixing between populations (Coope 1979). Also, the DNA evidence remains controversial (Klicka and Zink 1997; Avise and Walker 1998; Klicka and Zink 1999). Demonstrating that there are large numbers of divergences consistent with Pleistocene origins does not prove that speciation rates increased at that time because we might expect most species to have had recent origins, even if speciation rates have remained constant (Avise and Walker 1998; Klicka and Zink 1999). To resolve this issue, we need to investigate the dynamics of diversification, comparing Pleistocene speciation rates with those observed before the onset of glacial cycles. The one previous study taking this approach concluded that speciation rates in 11 North American bird genera had declined rather than increased during the Pleistocene (Zink and Slowinski 1995).

Here, we investigate recent diversification rates in the North American tiger beetles of the genus *Cicindela*, using a phylogenetic tree reconstructed from mtDNA that is densely sampled at the species level (A. P. Vogler, A. C. Diogo, and T. G. Barraclough, unpublished data). The genus *Cicindela* represents a spectacular radiation of insects, with around 130 species in North America and over 1,000 species found on all continents worldwide (except Antarctica, Pearson 1988). All species are sleek, raptorial predators, relying on fast locomotion (both in flight and on foot) and large mandibles to actively chase down a variety of arthropod prey. They have been well studied, particularly in North America, where their taxonomy, ecology, and geographic distributions are better known than in most other insect groups. Authors have proposed scenarios for the origin of tiger beetle species since the early days of their study (Leng 1902) and invariably phrase the speciation history

Key words: diversification rates, speciation, extinction, tiger beetles, Pleistocene.

Address for correspondence and reprints: Timothy G. Barraclough, Department of Biological Sciences, Imperial College at Silwood Park, Ascot, Berkshire SL5 7PY, UK.
E-mail: t.barraclough@ic.ac.uk.

Mol. Biol. Evol. 19(10):1706–1716. 2002
© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

of subgroups in the context of glaciation events (Freitag 1965; Rumpp 1967; Willis 1967; Acorn 1992). Because of the extreme scarcity of fossil remains (less than 10 fossils are known from North America, none more than 20,000 years old; Nagano, Miller, and Morgan 1982), attempts to investigate speciation and extinction in the group rely almost exclusively on phylogenetic data (Vogler and DeSalle 1993; Vogler, Welch, and Barraclough 1998; Barraclough, Hogan, and Vogler 1999; Barraclough and Vogler 2000; Morgan, Knisley, and Vogler 2000).

We extend recent methods to estimate average diversification rates among the North American *Cicindela* and to investigate how these rates have changed over time, in particular during the Pleistocene. Our analyses use dates estimated from the mtDNA tree and calibrated using biogeographic evidence. Glacial cycles are thought to have started in North America around 2.5 MYA, with increased intensity over the last 0.7 Myr (Webb and Bartlein 1992). Therefore, if glaciation increased speciation rates, we expect to observe an increase in per-lineage speciation rates around that time. But other processes can lead to similar patterns, for example, a constant background extinction rate is expected to cause an apparent acceleration in diversification rate toward the present (Harvey, May, and Nee 1994; Nee et al. 1994). We use statistical models to distinguish these alternatives and demonstrate a weak increase in *Cicindela* speciation rates within the last million years.

Materials and Methods

Phylogeny

We reconstructed the phylogeny of the North American tiger beetles using an aligned matrix of 1,897 nucleotide positions from three mtDNA regions: cytochrome *b*, 16S rRNA, and cytochrome oxidase III. Details of laboratory procedures, tree reconstruction, and tree support are provided elsewhere (A. P. Vogler, A. C. Diogo, and T. G. Barraclough, unpublished data). Our sampling covers the entire continent of North America, including the United States, Canada, and Mexico. We did not include the Caribbean or Central America. The tree includes 110 of the 147 described *Cicindela* species from the most recently published, comprehensive checklist of the region (Boyd 1982, p. 6), plus out-groups and a few species found in neighboring regions. Our original tree included a handful of subspecific taxa as well, but these have been excluded for the present analyses. Missing taxa were those for which we were unable to obtain specimens, many of them being rare or endangered species. The effects of missing taxa will be discussed subsequently. From taxonomic treatments (Rivalier 1954) and preliminary analyses including species from South America and the western and eastern Palearctic (A. P. Vogler, unpublished data), it is apparent that the North American assemblage is not monophyletic but comprises several radiations within the continent following independent colonization events.

Estimating the Relative Ages of Nodes from Sequence Data

We used the sequence data to estimate relative node ages for the phylogeny of *Cicindela*. First, branch lengths were fitted to the maximum parsimony tree of the full analysis (Vogler, Diogo, and Barraclough 2001) using maximum likelihood (ML) implemented in PAUP* 4.0 (Swofford 2001). The HKY85 model (transition-transversion ratio estimated from the data) with gamma-distributed rate variation among sites was chosen as significantly better than simpler models based on log likelihood ratio tests (Goldman 1993). Fitting more complex models led to improved likelihoods, but the branch lengths were highly correlated with those obtained under the chosen model (results not shown).

Likelihood ratio tests between rate-constant and rate-variable models revealed significant deviation from a molecular clock (Felsenstein 1981). Character-based methods are available for correcting rate variation among lineages (Thorne, Kishino, and Painter 1998; Huelsenbeck, Larget, and Swofford 2000), but implementation is still difficult for large matrices. Instead, we used two computationally simpler methods. First, we used Sanderson's nonparametric rate smoothing algorithm (NPRS) to estimate relative node ages from the unconstrained ML branch lengths (Sanderson 1997). The algorithm does not assume a strict molecular clock, simply that neighboring branches on the tree tend to have similar rates. Second, we fitted ML branch lengths under the chosen substitution model but assuming a molecular clock. Although our data deviate from a strict clock, there is a strong linear relationship between unconstrained and clock branch lengths ($r^2 = 0.88$), suggesting that rate variation may not affect our estimates of node ages too greatly (Losos and Schluter 2000). We perform all analyses on NPRS and ML clock estimates of node ages in turn. Details of calibrating the relative ages in terms of millions of years are provided in the results section.

Estimates of relative node ages using both methods will have associated errors caused by our finite sample of nucleotide characters. To assess the level of error, we generated 20 resampled bootstrap data matrices using the SEQBOOT program in PHYLIP, Version 3.573 (Felsenstein 1995). Each matrix was imported into PAUP* 4.0 and the maximum parsimony tree found by heuristic search (100 random addition replicates, tree bisection-reconnection [TBR] swapping, maximum of five trees held at each stage). Node ages were then fitted to one of the shortest trees obtained for each bootstrap replicate, using the procedures outlined previously. Analyses outlined subsequently were repeated on each of the resulting ultrametric trees to assess the effects of errors in topology and branch lengths caused by having a finite sample of characters.

Missing Taxa

Subsequent analyses assume that all extant species within the assemblage have been sampled in the phylogeny. To examine the possible effects of missing spe-

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

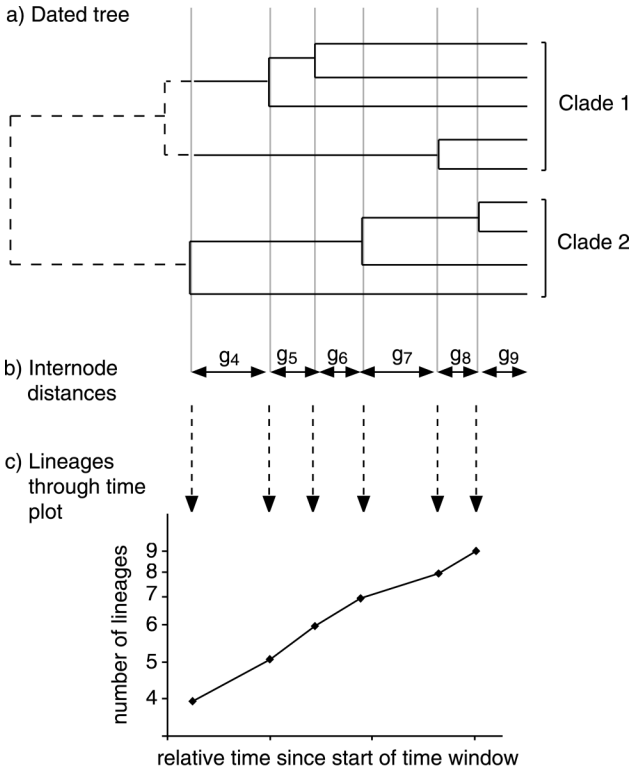

FIG. 1.—Hypothetical example of how information on the timing of diversification can be extracted from a dated phylogeny. The example shows a continental assemblage derived from two independent colonizations, clade 1 and clade 2. On the basis of the age of the first within-continent split in each clade, clade 2 radiated more recently within the area than did clade 1. Therefore, we consider diversification rates in the assemblage only since the root node within clade 2 (branches shown in bold). The internode distances, g_i, after this time can be used to estimate diversification rates within the continent (the subscript indicates the number of lineages present during that internode, for details see *Materials and Methods*). Internode distances before this time may depend upon events outside the continent and rates of immigration or emigration (or both) as well as diversification rates within the continent. The accumulation of species over time can be represented graphically as a plot of the number of lineages through time.

cies, we placed all the missing North American species described in Boyd (1982) in their most likely place on the tree on the basis of taxonomic accounts (Cazier 1954; Rivalier 1954; Boyd 1982). A further problem is that taxonomic effort has been greater in the United States and Canada than in Mexico. One unpublished checklist of Mexican species describes 16 additional species not included in the Boyd (1982) checklist (W. Sumlin, personal communication), mostly the result of upgrading existing subspecific taxa and new discoveries in previously unsampled areas. Hence, although the alpha taxonomy of the Mexican taxa remains unstable, we also added these presently ''undescribed'' species in their most likely place on the mtDNA tree. The result is a tree containing 146 species. For subsequent analyses we removed the four species that are members of predominantly South American clades from our original tree because these clades are very incompletely sampled in our mtDNA tree. All analyses were repeated with this tree as well as our original mtDNA tree. This treatment cannot tell us the relationships we would obtain if missing taxa had been included in our matrix, but we use it to give some indication of how missing taxa may affect our analyses. The modified tree is available online. Note that we have no information for branch lengths connecting the added species. Instead, we assigned nodes to be half way along the branch to which each species was added (Losos 1990). This would tend to be conservative with respect to detecting a recent increase in the apparent diversification rate.

Estimating Diversification Rates

We plot the log of the number of lineages against the branch length distance from the root node (for the ultrametric trees obtained by NPRS and ML clock). Because the North American *Cicindela* are not monophyletic, we consider diversification rates only over the time interval since the most recent, first within-continent split for a radiation confined to the continent (see fig. 1). Under a constant speciation rate model we expect a straight line with slope b, the speciation rate. If there has been a recent increase in the speciation rate caused by the onset of glacial cycles, then we expect an increase in slope at around 2.5–0.7 MYA (Klicka and Zink 1999). If there has been a constant background extinction rate, d, then we expect an apparent acceleration in diversification rate toward the present, with the slope changing from $b - d$ to b, starting at around $1/(b - d)$ Myr before the present (Harvey, May, and Nee 1994).

Standard statistical methods were used to estimate diversification rates and to test for significant departures from the constant speciation rate model. First, we estimated b assuming a constant speciation rate model, using the Kendall-Moran estimator (Baldwin and Sanderson 1998; Nee 2001). For a time window starting at time 0 and finishing at time t, the ML estimate of the per-lineage speciation rate is $b = (n - m)/B$, where m and n are the number of lineages at the start and the end of the time window, respectively, and B is the sum of the branch lengths falling within the time window. Following recommendations by Nee (2001), we use the Moran estimate of the variance, $\mathrm{var}(b) = b^2/(n - m)$, to calculate confidence intervals for our estimate on the basis of the error arising because our estimate is derived from a finite number of nodes.

To test for significant departures from the constant speciation rate model we follow Pybus and Harvey (2000). Their statistic, γ, compares the relative positions of nodes in a phylogeny to those expected under a constant speciation rate model (see also Zink and Slowinski 1995). This can be generalized as described subsequently. For a time window starting at time 0 with m species and finishing at time t with n species and where $g_m, g_{m+1}, \ldots, g_n$ are the internode distances during the time period as shown in figure 1, the statistic is equation (1)

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

$$\gamma = \frac{\left(\frac{1}{n-m}\sum_{i=m}^{n-1}\left(\sum_{k=m}^{i} kg_k\right)\right) - \left(\frac{T}{2}\right)}{T\sqrt{\frac{1}{12(n-m)}}},$$

$$T = \sum_{j=m}^{n} jg_j \qquad (1)$$

Under a constant speciation rate model, the statistic follows a standard normal distribution. Positive values signify that nodes are closer to the tips than what is expected under the constant speciation rate model, i.e., there has been an apparent increase in diversification rate toward the present. Negative values signify an apparent deceleration. Therefore, the null hypothesis of constant b cannot be rejected at 5% level in a two-tailed test if $-1.96 < \gamma < 1.96$. Hence, we calculated γ for the time window since the most recent invasion by a major group to test whether net diversification rates changed over time.

An increase in the speciation rate during the Pleistocene would lead to a significant positive value of γ. But a constant rate of background extinction could also cause a significant increase in the apparent diversification rate toward the present. To distinguish these alternatives we performed the following test (see also Paradis 1997; Emerson, Oromi, and Hewitt 2000*a*, 2000*b*). Under a constant speciation and extinction rate model of clade growth, the likelihood of each internode distance g_i is given by

$$i(b-d)e^{-i(b-d)g_i} \times \frac{\left(1 - \frac{b}{d}e^{-(x_i-g_i)(b-d)}\right)^{i-1}}{\left(1 - \frac{b}{d}e^{-x_i(b-d)}\right)^{i}} \qquad (2)$$

(eq. [17] in Nee, May, and Harvey 1994). Therefore, the log likelihood of the set of internode distances observed within a time window is simply the sum of the log likelihood of each internode distance. We used this basic formula to compare the likelihoods of three models for the diversification of tiger beetles during the time window: model (1) a constant speciation rate model with no extinction (i.e., d is held at 0), model (2) a constant speciation rate and extinction rate model (b and d are estimated), and model (3) a step-model in which a constant speciation rate b_1 shifts to a new constant speciation rate b_2 at time T, where T is a time within the time window that is optimized, and d is held at 0. The models have one, two, and three parameters, respectively. We used the Solver in Excel to find the ML estimates of the parameters of each model. If an increase in the apparent diversification rate was caused by a Pleistocene increase in speciation rate, we would expect model (3) to provide a significantly better fit to the data than would model (2) and that the ML estimate of T should be around 2.5–0.7 MYA. Because model (2) is not nested within model (3), we use the Akaike information criterion (AIC) to select the best model (Akaike 1974). The model with the highest AIC is chosen, where AIC = 2 Log$L$ − 2p, where L is the likelihood and p is the number of parameters in the model.

Our analyses consider the average diversification rates across at least four independent radiations within North America. To test for variation among them, we repeated our analyses on each of the major subgroups in turn. Apart from estimating b for each clade, we also tested for significant differences in b among groups. For each group we multiplied each internode distance by the number of lineages present during that internode, i.e., we calculated ig_i for all internodes. Under the constant speciation rate model, these transformed internode distances are expected to be constant and equal to 1/b (Purvis, Nee, and Harvey 1995; Nee 2001). Hence, we tested for significant differences among clades using an ANOVA with subgroup as the single factor and transformed internode distances as the data. For each clade, we also tested for significant departures from the constant speciation rate model, using the γ statistic.

Results

Figure 2 shows the phylogenetic tree of the North American tiger beetles, with ultrametric branch lengths fitted by NPRS of ML branch lengths. The tree with the ML clock branch lengths is shown in the electronic appendix. Relative node ages obtained in the two trees are broadly correlated ($r^2 = 0.94$). The main difference is that the clock method estimates relatively older ages for all *Cicindelidia* and Beach clade nodes and relatively younger ages for the all *Cicindela* s.s. and *Ellipsoptera* nodes than does NPRS. This is because rate variation in our data appears to be primarily between the major subclades, rather than within subclades, and the dating methods differ in how they deal with that variation. The tree with species added that are missing from our sample of sequenced taxa is provided in the electronic appendix.

In the absence of an appropriate fossil record, we used biogeographic information to calibrate the relative node ages in millions of years. First, we identified five splits associated with the Florida peninsula—two between a terrestrial species restricted to Florida and a sister clade found on the ''mainland'' in the United States and three between coastal sister clades separated onto the Gulf and Atlantic coasts by the peninsula (table 1 and fig. 2). Extensive work by Avise and coworkers (reviewed in Avise 1992, 1994, p. 242) has demonstrated a large number of vicariances in the region, apparently associated with falling but fluctuating sea levels since the late Pliocene. On the basis of reconstructed global sea levels (Lane 1994), Florida would have first emerged around 4 MYA; hence, we assume a maximum date for these splits of 4 Myr (i.e., any individual split could have occurred since that time, Avise 1992). Second, we identified South American groups (the South American *Cylindera* and *Brasiella*) that have crossed into North America followed by speciation leading to sister species in Mexico and the southern United States. Both groups inhabit dry terrestrial habitats (table 1 and fig. 2). If those species invaded North America via land,

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

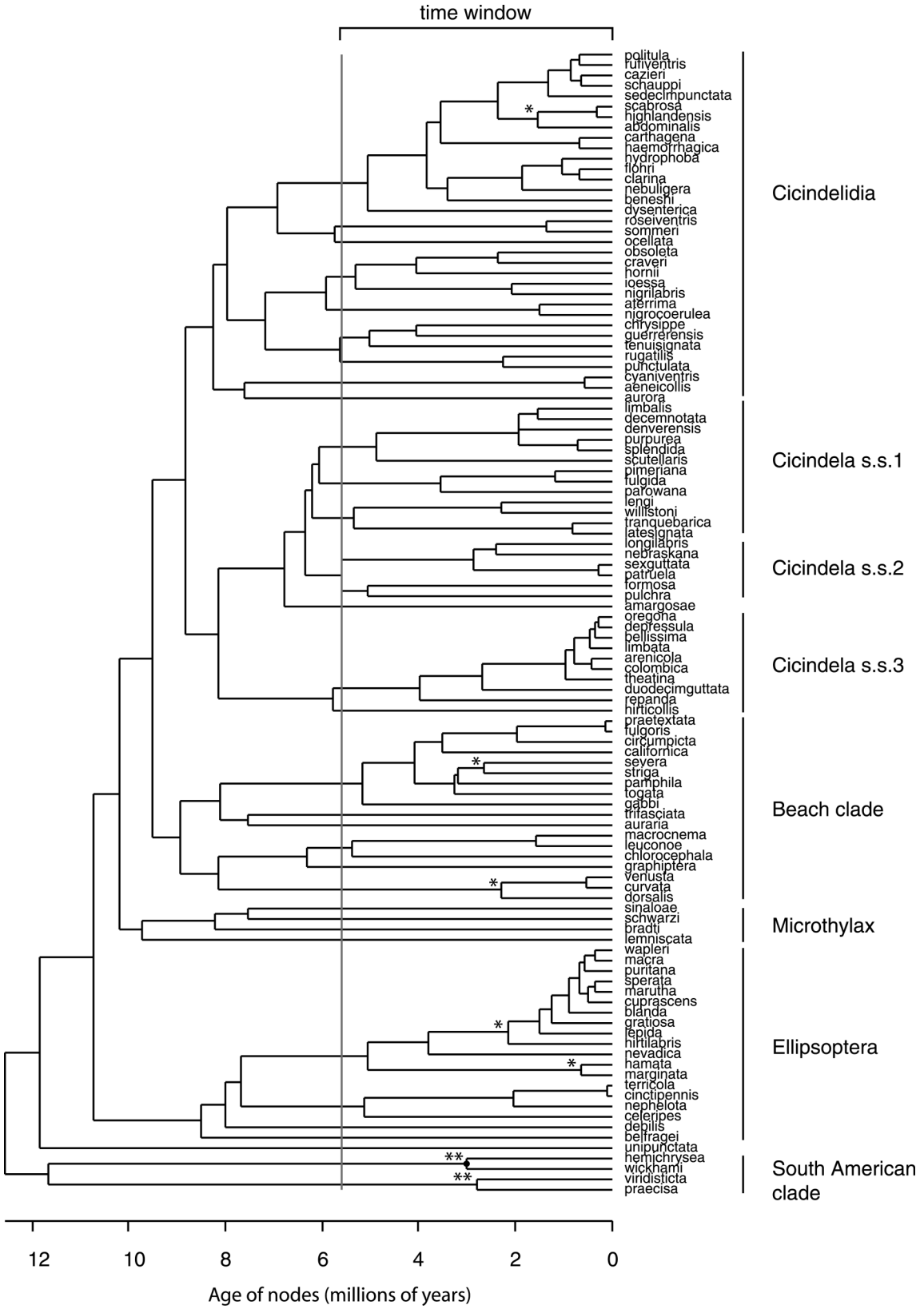

FIG. 2.—Phylogenetic tree of the North American *Cicindela* species showing relative node ages obtained by nonparametric rate smoothing of ML DNA branch lengths. The scale bar shows the calibration of ages in real time based on biogeographic information. Nodes with biogeographic maximum dates are indicated: * for splits associated with Florida, and ** for splits in South American lineages that have crossed into North America. The node used to calibrate the tree, satisfying all constraints, is indicated with a circle. Names of independent radiations within the continent are shown. *Cicindela* s.s. 2 is the most recent lineage to have undergone extensive radiation within the continent; hence, our analyses consider diversification over the time window since its root node. The gray line indicates the start of the time window.

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

Table 1
Maximum Biogeographical Dates and Pairwise Sequence Distances for Nodes Used to Calibrate the Tree

Node	Scenario[a]	Sister Clade 1	Sister Clade 2	Maximum Age from Biogeography (Myr)	Uncorrected Pairwise Distance (%)[b]	Age from mtDNA Clock (Myr)[c]
(a) . .	Florida	dorsalis	venusta + curvata	4	5.0	2.5
(b) . .	Florida	striga	severa	4	6.4	3.2
(c) . .	Florida	hirtilabris	lepida to wapleri	4	3.9	2.0
(d) . .	Florida	abdominalis	scabrosa + highlandensis	4	7.2	3.6
(e) . .	Florida	hamata	marginata	4	0.8	0.4
(f). . .	Panama	hemichrysea	wickhami	3	4.9	2.5
(g) . .	Panama	viridisticta	praecisa	3	3.6	1.8

[a] Nodes fall into one of the two biogeographical scenarios: speciation associated with sea level fluctuations since the emergence of Florida and speciation within Mexico and the United States of lineages that crossed from South America via Panama.

[b] The average of all uncorrected pairwise distances between species in sister clade 1 and species in sister clade 2.

[c] Age of each node assuming a mtDNA clock of 2% pairwise uncorrected sequence divergence per million years.

that would suggest a maximum age of around 3 Myr for those splits, the time of final emergence of the Isthmus of Panama (Lessios 1998).

To check the consistency of our estimates, we compare our dates with those expected using the widely cited insect mtDNA clock of around 2% pairwise divergence/Myr (Brower 1994), shown in table 1. The maximum biogeographical dates are consistent with the published mtDNA rate. For example, the oldest of the Florida splits has a pairwise sequence divergence of 7.2%, suggesting a date of 3.6 Myr (table 1), and the oldest of the Panama splits has a pairwise sequence divergence of 4.9%, suggesting a date of 2.5 Myr. Hence, to calibrate the tree we fixed the *B. hemichrysea-B. wickhami* node at 3 MYA, providing maximum dates that satisfy all our constraints. The timescale is shown in figure 2. Although the calibration is based on several uncertainties, it appears to give reasonable absolute dates for nodes on the tree. We used the same logic to calibrate each of the 20 bootstrap replicate trees. Despite differences in topology among replicates, at least five of the nodes were recovered in any single bootstrap, and so we assigned a maximum date that satisfied the constraints of those splits still found.

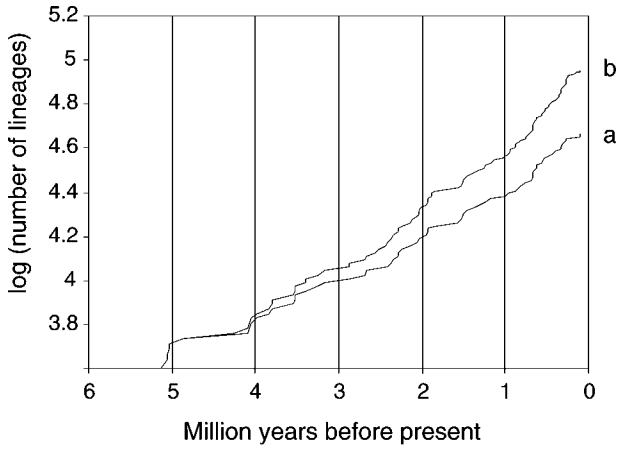

FIG. 3.—Lineages-through-time plot for the North American *Cicindela* over the last 5.6 Myr using NPRS dates. Line (*a*) is the plot for species sampled in our mtDNA survey, (*b*) the plot including missing species named in the checklist by Boyd (1982) and those described by W. Sumlin (personal communication).

Because the North American *Cicindela* are not monophyletic (see *Materials and Methods*), the average rates of diversification were only calculated across subclades representing radiations confined to this continent. These clades are shown in figure 2 and include four major subgroups: the subgenera *Cicindela* s.s., *Cicindelidia*, *Ellipsoptera*, and a clade confined to ocean beaches and salt flats comprising the subgenera *Habroscelimorpha*, *Opilidia*, *Eunota*, and the divergent *Cicindelidia trifasciata* (hereafter referred to as ''Beach clade,'' Vogler, Diogo, and Barraclough 2001). *Microthylax* was also included for the overall analysis of diversification rates but was not included in the separate analyses of subclades (see subsequently) because the number of species is too small to allow an accurate estimation. The recent South American invaders were not included in our analyses of diversification rates.

Cicindela s.s. probably comprises three independent radiations within North America (Vogler, Diogo, and Barraclough 2001) and includes the most recent colonization leading to a major in situ radiation in North America, the *Cicindela* s.s. group 2. The earliest within-continent split in this group is estimated at 5.6 MYA in the NPRS tree. Hence, we consider diversification rates from this date. Figure 3 shows the lineage-through-time plot for the North American *Cicindela* sampled in our tree over the last 5.6 Myr. The plot obtained when missing species were added to the phylogeny is superimposed. The pattern observed is that of a roughly linear increase on the semi-log plot but with some indication of acceleration in rate toward the present, which becomes more marked when missing species are added.

When only species from our mtDNA survey were included in the analyses, the Kendall-Moran estimate of the per-lineage net diversification rate from the NPRS tree is 0.22 ± 0.05 species Ma^{-1} (95% confidence interval, table 2). This represents a minimum estimate be-

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

Table 2
Estimates of the Diversification Rate and the Change in Diversification Rate Over Time in North American *Cicindela* During the Time Window

Sampling[a]	Dating Method	Number of Lineages[b]		Net Diversification Rate[c] *b* 95% CL			Changes Over Time[d]	
		Start	End	Myr	Lower	Upper	γ	*P*
(A) . . .	NPRS	32	106	0.22	0.18	0.27	0.74	0.46
	CLOCK	38	106	0.22	0.18	0.27	1.38	0.17
(B) . . .	NPRS	32	142	0.29	0.25	0.34	1.81	0.07
	CLOCK	38	142	0.29	0.25	0.35	2.30	0.02

[a] Sampling treatments were: (A) only including species sampled in our mtDNA survey, and (B) adding missing species named in the checklist by Boyd (1982) and those described by W. Sumlin (personal communication).

[b] The number of lineages at the start and end of the time window.

[c] The Kendall-Moran estimate of net diversification rate, *b*, with Moran 95% confidence limits.

[d] Change in diversification rate over time measured by γ: positive values indicate an apparent increase in diversification rates toward the present. *P* is the two-tailed probability of observing a more extreme value of γ under a constant speciation rate model.

cause our calibration of the tree was based on maximum ages for particular nodes. The quoted errors are caused by estimating diversification rates from a finite number of nodes. Another source of error comes from an uncertainty in the topology and branch length estimates, which we addressed by resampling the original data matrix. The distribution of the estimates of *b* obtained from the 20 bootstrap replicate trees is positively skewed but with median 0.23 ± 0.03 (95% confidence intervals, Zar 1984, p. 113). Even if we combine the two sources of error, the data provide a fairly narrow estimate of the net diversification rate. Very similar estimates were obtained using the ML clock dates (table 2, median across 20 bootstrap replicates, 0.20 ± 0.02). Table 2 also shows the estimates of the diversification rate when missing species were added to the tree. As expected, adding missing species increases the Kendall-Moran estimate of diversification rates.

There is no significant increase in diversification rates toward the present using the γ test when only species from our mtDNA survey are included (table 2), although the pattern is stronger using ML clock dates than when using NPRS. The value of γ is sensitive to bootstrap resampling in both cases: the 95% limits obtained from the bootstrap trees are wide but do not include the value obtained from the original matrix and tree (NPRS dates, $\gamma = -0.23 \pm 0.72$, ML clock dates, $\gamma = 1.82 \pm 0.36$). Curiously, resampling tends to reduce γ when using NPRS dates but tends to increase γ when using ML clock dates. When missing species are added to the tree, the increase in diversification rate toward the present becomes more significant but remains weak (table 2). Therefore, we have only weak evidence for rejecting the constant speciation rates model.

These findings were confirmed by comparing the likelihoods of the data under the three models of diversification rates (table 3). When only species sampled in our mtDNA survey were included, the best model for the ML clock dates was the constant speciation rate model and that for the NPRS dates was, marginally, a step-model with a decrease in the speciation rate shortly after the start of the time-window (at about 5 MYR). When missing species were added, both ML and NPRS dates were best explained by a step-model with an 80% increase in the speciation rate at 1 MYR, but this was only marginally better than the constant speciation and extinction rates models (with extinction fairly high relative to speciation, $d/b > 0.7$). The favored model is consistent with the predictions of the glacial speciation model that speciation rates increased between 2.5 and 0.7 MYR.

One possible bias affecting these results might be that we fitted missing species into our tree relatively too close to the tips. This might occur if taxonomic accounts tended to name a single close relative for a given species rather than naming a clade within which the species is

Table 3
Comparison of Three Models of Diversification During the Time Window

Sampling[a]	Dating Method	Model 1[b]		Model 2			Model 3			
		b	AIC	*b*	*d*	AIC	*b*1	*b*2	*T*	AIC
(A). . .	NPRS	0.23	247.4	0.29	0.17	246.9	0.47	0.21	5.04	247.6[c]
	CLOCK	0.22	233.6[c]	0.29	0.20	229.69	0.18	0.31	1.00	230.3
(B). . .	NPRS	0.30	474.8	0.46	0.34	479.4	0.24	0.43	1.00	479.7[c]
	CLOCK	0.29	453.7	0.42	0.36	457.2	0.23	0.42	0.95	457.8[c]

[a] Sampling treatments were: (A) only including species sampled in our mtDNA survey, and (B) adding missing species named in the checklist by Boyd (1982) and those described by W. Sumlin (personal communication).

[b] Model 1 is the constant speciation rate (*b*) model, Model 2 is the constant speciation (*b*) and extinction rate (*d*) model, and Model 3 is the model in which speciation rate jumps from one constant rate (*b*1) to a second constant rate (*b*2) at time *T*. *b* and *d* are in units of Myr and *T* is in units of MYA. AIC is the Akaike Information Criterion.

[c] The highest value of AIC in each row, signifying the best model for that treatment.

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

found. In addition, we do not really know where the missing species would go if we sampled their DNA: including them might change the topology of the existing tree. To address this, we performed the Monte Carlo constant-rates (MCCR) test of Pybus and Harvey (2000), a more conservative test of the effects of missing species that simulates expected values of γ, assuming incomplete sampling of a clade. This test does not assume knowledge of where missing species should go or even the topology of sampled species. Using a constant speciation rate model we simulated 1,000 phylogenies of 142 species (the number of species from the checklist by Boyd [1982] plus additional taxa from W. Sumlin, personal communication). For each trial, we deleted 36 species at random to obtain a phylogeny for 106 of the 142 species. We then calculated γ for each trial and generated the expected distribution across the 1,000 trials. The two-tailed probabilities for the observed values of γ were 0.25 using NPRS dates and 0.04 using ML clock dates. This suggests that the recent increase in diversification rates observed for ML clock dates is robust to the exact location of missing species but that the weaker trend for NPRS dates could be lost if missing species were added further from the tips.

Tables and figures comparing diversification among the four major North American subgroups, *Cicindela* s.s., *Cicindelidia*, the Beach clade, and *Ellipsoptera*, are provided in the electronic appendix. The ANOVA of transformed internode distances using NPRS dates revealed little difference in the net diversification rate among subgroups, becoming marginally significant only when missing species are added ($F = 3.8$, $P < 0.05$). The differences among groups were more significant using ML clock dates (sampled taxa, $F = 3.2$, $P < 0.05$; missing species added, $F = 6.8$, $P < 0.01$). The main trend is for a lower diversification rate in the Beach clade compared with the other three groups. The γ statistics for each group revealed little significant departure from the constant speciation rate model, except for an increase toward the present in *Cicindela* s.s. using the ML clock dates. Interestingly, *Cicindela* s.s. and *Ellipsoptera*, both of which have a northerly distribution in the continent, displayed more positive values of γ than did the other two clades. The Beach clade displayed a nonsignificant decline in diversification rates toward the present, and *Cicindelidia* displayed a decline that became a slight increase when missing species are added.

Discussion

From the sample of species in our mtDNA tree we calculated an average per-lineage diversification rate over roughly the last 5 Myr of at least 0.22 ± 0.08 species/Myr, rising to 0.29 species/Myr when missing species were added to the tree. The estimate is robust to resampling of the data matrix and to the methods used to establish an ultrametric tree from rate-variable data. But the estimate does rely critically on the accuracy of our biogeographic calibration. No suitable fossils are available to confirm our dates (Nagano, Miller, and Morgan 1982), but the dates are consistent with those expected assuming the widely cited mtDNA clock (Brower 1994). Future work will attempt to identify further biogeographic information for dating the *Cicindela* tree, for example, in other continents and at deeper levels using between-continent divergences.

How does this compare to the typical diversification rates in insects and other groups? Mayhew (2002) estimated rates in insect orders from fossil dates, ranging from 0.008 to 0.06 species/Myr, but the wider taxonomic scale limits the usefulness of direct comparison. McCune (1996) cites speciation rates in Hawaiian insects that range from 0.66 to 1.21 species/Myr on the basis of the numbers of extant species and the geological age of the oldest extant island in the Archipelago. But if the taxa first colonized older but now extinct islands, these estimates could fall by up to one-tenth (McCune 1996). A few studies have used the methods we followed to estimate diversification rates in other groups, for example, between 0.09 and 0.34 for a range of primate clades and 0.56 for Hawaiian silversword plants (Purvis, Nee, and Harvey 1995; Baldwin and Sanderson 1998). More studies are needed to establish typical diversification rates in different organisms, but compared with present studies the North American *Cicindela* provide no indication of unusually high diversification rates.

Our results provide some evidence for the effects of Pleistocene glaciations on tiger beetle diversification. When missing species were added to our mtDNA tree, we observed a weakly significant twofold increase in the net diversification rate within the last million years, consistent with a response to the increased intensity of glacial cycles in the late Pleistocene (Webb and Bartlein 1992). Previous work showed that the North American tiger beetles have experienced major range movements (Barraclough and Vogler 2000), and the effects of climate change on species ranges would be a likely cause for the increased speciation rate. But the alternative explanation that speciation rates have been constant over time but with a high background extinction rate was only slightly less favored. This shows that it can be difficult to distinguish alternative explanations for a given pattern of diversification even with a sample size of over 70 nodes, large potential effect sizes (i.e., doubling in speciation rate or $d > 0.7b$), and efficient statistical methods (Barraclough and Nee 2001).

The results on how diversification rates changed over time were sensitive to our reconstruction of node ages from the mtDNA data. The increase in diversification toward the present was strongest using ML clock dates: with NPRS dates there was only a weak evidence to reject the constant speciation rate model, and the trends were sensitive to resampling. This arose because NPRS stretched out the ages of clades with low sequence divergence among species, whereas the clock method stretched internal branches and left terminal clades with shorter branches. Comparisons with unrelated data suggest that the direction of bias varies among data sets (T. G. Barraclough, unpublished data). More work is needed to evaluate these methods, in particular how they affect the distribution of ages obtained for a given data and how well they perform when their as-

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

sumptions are not met. Although the ML clock method assumes an unjustified molecular clock, the assumptions of NPRS (i.e., rates change smoothly across the tree) may not be any more justified. Because both methods make unjustified assumptions, it is not clear which method will provide the best estimates even though NPRS nominally allows for rate variation. We believe that character-based approaches that incorporate rate variation among lineages will be the best solution (Thorne, Kishino, and Painter 1998; Huelsenbeck, Larget, and Swofford 2000) once methods that can be applied to large data sets become available.

Missing species had a sizeable effect on the change in diversification rates over time. The increase in rates was only found when missing species were added to our mtDNA tree. Although we know that at least 20 recognized North American *Cicindela* species are missing from our tree, we can only guess what the tree would have been if mtDNA from those species had been sampled. Simulations suggested that the ML clock results were robust to this uncertainty, but the weaker NPRS results were sensitive to the location of missing species.

More fundamentally, our study assumes that currently recognized taxonomic species correspond to evolutionary units. Tiger beetle species have been described mostly on the basis of evolutionarily labile traits, such as elytral coloration and body proportions, which may indicate substantial historical divergence in some cases but not others (Morgan, Knisley, and Vogler 2000). Although in the present study practical limitations of sampling all species in the region made it necessary to follow the taxonomy, future studies at the population-species boundary in representative *Cicindela* are needed to justify this approach. Our results show how sensitive the analyses can be to departures from full sampling of a lineage; therefore, reliably identifying and sampling all the lineages within a clade is vital for this kind of study (Avise 2000; Barraclough and Nee 2001).

Finally, note that we compared three simple diversification models for the North American tiger beetles, but other scenarios could lead to similar patterns. For example, if a mass extinction event occurred 1 MYA, perhaps triggered by changes in glacial cycles, the expected pattern is for a change in the slope on the lineages-through-time plot at the time of the event (Harvey, May, and Nee 1994; Kubo and Iwasa 1995). Alternatively, if newly formed species have a higher risk of extinction than do older species, perhaps because their geographic ranges tend to be smaller, then this could lead to a shorter lag between speciation and extinction than in the constant extinction rates model. Both scenarios could lead to similar quantitative predictions to those of the Pleistocene speciation model. But although it is impossible to pin down a single explanation for the observed patterns, clearly our results reject the hypothesis that speciation rates declined during the Pleistocene because of mixing of populations in a dynamically changing landscape. Similar conclusions have been obtained from population studies in other insect groups (Hewitt 1999; Knowles 2001).

In conclusion, our study estimated diversification rates within an insect group from phylogenetic data and found marginal evidence for a late Pleistocene increase in the net diversification rate. Application of similar methods in a range of groups could provide critical insights into the effects of Pleistocene climate, and environmental change more generally, on diversification rates. Successful applications in the future will depend on robust methods for dating phylogenetic trees and on establishing the evolutionary status of taxonomically recognized species, a key assumption of studies of this kind.

Supplementary Material

Figures showing the tree and lineage-through-time plots obtained using the ML clock method of date estimation, the tree with unsampled species added, and separate analyses for the major subgroups are provided online.

Acknowledgments

We thank J. Avise, S. Nee, O. Pybus, and two referees for helpful comments, and W. Sumlin for permission to use his unpublished checklist. This work was supported by the Natural Environment Research Council (NERC) (GR3/10632 and NERC/A/S/2000/00489). T.G.B. is a Royal Society University Research Fellow.

LITERATURE CITED

ACORN, J. H. 1992. The historical development of geographic color variation among dune Cicindela in Western Canada (Coleoptera: Cicindelidae). Pp. 217–233 *in* G. R. NOONAN, G. E. BALL, and N. E. STORK, eds. The biogeography of ground beetles of mountains and islands. Intercept, Andover, UK.

AKAIKE, H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. A-C **19**:716–723.

AVISE, J. C. 1992. Molecular population-structure and the biogeographic history of a regional fauna—a case-history with lessons for conservation biology. Oikos **63**:62–76.

———. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.

———. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, Mass.

AVISE, J. C., and D. WALKER. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc. R. Soc. Lond. B. **265**:457–463.

BALDWIN, B. G., and M. J. SANDERSON. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proc. Natl. Acad. Sci. USA **95**:9402–9406.

BARRACLOUGH, T. G., J. E. HOGAN, and A. P. VOGLER. 1999. Testing whether ecological factors promote cladogenesis in a group of tiger beetles (Coleoptera: Cicindelidae). Proc. R. Soc. Lond. B. **266**:1061–1067.

BARRACLOUGH, T. G., and S. NEE. 2001. Phylogenetics and speciation. Trends Ecol. Evol. **16**:391–399.

BARRACLOUGH, T. G., and A. P. VOGLER. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. Am. Nat. **155**:419–434.

BARRACLOUGH, T. G., A. P. VOGLER, and P. H. HARVEY. 1998. Revealing the factors that promote speciation. Philos. Trans. R. Soc. Lond. B. **353**:241–249.

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

BOYD, H. P. 1982. Checklist of Cicindelidae: the tiger beetles. Plexus, Marlton, NJ.

BROWER, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* Inferred from patterns of mitochondrial-DNA evolution. Proc. Natl. Acad. Sci. USA **91**:6491–6495.

CAZIER, M. 1954. A review of Mexican tiger beetles of the genus *Cicindela* (Coleoptera: Cicindelidae). Bull. Am. Mus. Nat. Hist. **103**:227–310.

COOPE, G. R. 1979. Late Cenozoic fossil Coleoptera: evolution, biogeography, and ecology. Annu. Rev. Ecol. Syst. **10**:249–267.

CRACRAFT, J. 1985. Biological diversification and its causes. Ann. Mo. Bot. Gard. **72**:794–822.

EMERSON, B. C., P. OROMI, and G. M. HEWITT. 2000*a*. Tracking colonization and diversification of insect lineages on islands: mitochondrial DNA phylogeography of *Tarphius canariensis* (Coleoptera: Colydiidae) on the Canary Islands. Proc. R. Soc. Lond. B. **267**:2199–2205.

———. 2000*b*. Colonization and diversification of the species *Brachyderes rugatus* (Coleoptera) on the Canary Islands: evidence from mitochondrial DNA COII gene sequences. Evolution **54**:911–923.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences—a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

———. 1995. PHYLIP (phylogeny inference package). University of Washington, Seattle.

FREITAG, R. 1965. A revision of the North American species of the *Cicindela maritima* group with a study of hybridization between *Cicindela duodecimguttata* and *oregona*. Quaest. Entomol. **1**:87–170.

GOLDMAN, N. 1993. Statistical tests of models of DNA substitution. J. Mol. Evol. **36**:182–198.

HACKETT, S. J. 1996. Molecular phylogenetics and biogeography of tanagers in the genus *Ramphocelus* (Aves). Mol. Phylogenet. Evol. **5**:368–382.

HAFFER, J. 1969. Speciation in Amazonian forest birds. Science **165**:131–137.

HARVEY, P. H., R. M. MAY, and S. NEE. 1994. Phylogenies without Fossils. Evolution **48**:523–529.

HEWITT, G. M. 1999. Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. **68**:87–112.

HEY, J. 1992. Using phylogenetic trees to study speciation and extinction. Evolution **46**:627–640.

HUELSENBECK, J. P., B. LARGET, and D. SWOFFORD. 2000. A compound Poisson process for relaxing the molecular clock. Genetics **154**:1879–1892.

KLICKA, J., and R. M. ZINK. 1997. The importance of recent ice ages in speciation: a failed paradigm. Science **277**:1666–1669.

———. 1999. Pleistocene effects on North American songbird evolution. Proc. R. Soc. Lond. B. **266**:695–700.

KNOWLES, L. L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. Mol. Ecol. **10**:691–701.

KUBO, T., and Y. IWASA. 1995. Inferring the rates of branching and extinction from molecular phylogenies. Evolution **49**: 694–704.

LANE, E. 1994. Florida's geological history and geological resources. Florida Geological Survey.

LENG, C. W. 1902. The geographical distribution of Cicindelidae in eastern North America. J. N Y Entomol. Soc. **20**: 1–17.

LESSIOS, H. A. 1998. The first stage of speciation as seen in organisms separated by the Isthmus of Panama. Pp. 186–201 *in* D. J. HOWARD and S. H. BERLOCHER, eds. Endless forms: species and speciation. Oxford University Press, New York.

LOSOS, J. B. 1990. Ecomorphology, performance capability and scaling of West Indian *Anolis* lizards: an evolutionary analysis. Ecol. Monogr. **60**:369–388.

LOSOS, J. B., and D. SCHLUTER. 2000. Analysis of an evolutionary species-area relationship. Nature **408**:847–850.

MAYHEW, P. J. 2002. Shifts in hexapod diversification and what Haldane could have said. Proc. R. Soc. Lond. B **269**:969–974.

MCCUNE, A. R. 1996. How fast is speciation? Molecular, geological and phylogenetic evidence from adaptive radiations of fishes. Pp. 585–610 *in* T. GIVNISH and K. SYSTMA, eds. Molecular evolution and adaptive radiation. Cambridge University Press, Cambridge.

MORGAN, M., C. B. KNISLEY, and A. P. VOGLER. 2000. New taxonomic status of the endangered tiger beetle *Cicindela limbata albissima* (Coleoptera: Cicindelidae): evidence from mtDNA. Ann. Entomol. Soc. Am. **93**:1108–1115.

NAGANO, C. D., S. E. MILLER, and A. V. MORGAN. 1982. Fossil tiger beetles (Coleoptera: Cicindelidae): review and new quaternary records. Psyche **89**:339–346.

NEE, S. 2001. On inferring speciation rates from phylogenies. Evolution. **55**:661–668.

NEE, S., E. C. HOLMES, R. M. MAY, and P. H. HARVEY. 1994. Extinction rates can be estimated from molecular phylogenies. Philos. Trans. R. Soc. Lond. B **344**:77–82.

NEE, S., R. M. MAY, and P. H. HARVEY. 1994. The reconstructed evolutionary process. Philos. Trans. R. Soc. Lond. B **344**:305–311.

PARADIS, E. 1997. Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. Proc. R. Soc. Lond. B **264**:1141–1147.

PEARSON, D. L. 1988. Biology of tiger beetles. Annu. Rev. Entomol. **33**:123–147.

PURVIS, A., S. NEE, and P. H. HARVEY. 1995. Macroevolutionary inferences from primate phylogeny. Proc. R. Soc. Lond. B **260**:329–333.

PYBUS, O. G., and P. H. HARVEY. 2000. Testing macroevolutionary models using incomplete molecular phylogenies. Proc. R. Soc. Lond. B **267**:2267–2272.

RIBERA, I., T. G. BARRACLOUGH, and A. P. VOGLER. 2001. The effect of habitat type on speciation rates and range movements in aquatic beetles: inferences from species-level phylogenies. Mol. Ecol. **10**:721–735.

RIVALIER, E. 1954. Démembrement du genre *Cicindela* Linne. II. Faune americaine. Rev. Fr. Entomol. **21**:249–268.

ROY, M. S. 1997. Recent diversification in African greenbuls (Pycnonotidae: *Andropadus*) supports a montane speciation model. Proc. R. Soc. Lond. B **264**:1337–1344.

RUMPP, N. L. 1967. A new species of *Cicindela* from Idaho (Coleoptera: Cicindelidae). Proc. Calif. Acad. Sci. **38**:129–140.

SANDERSON, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. **14**:1218–1231.

SANDERSON, M. J., and M. J. DONOGHUE. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol. Evol. **11**:15–20.

SWOFFORD, D. L. 2001. PAUP*: phylogenetic analysis using parsimony (*and Other Methods). Sinauer Associates, Sunderland, Mass.

THORNE, J. L., H. KISHINO, and I. S. PAINTER. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. **15**:1647–1657.

VOGLER, A. P., and R. DESALLE. 1993. Phylogeographic patterns in coastal North American Tiger Beetles, *Cicindela*

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024

dorsalis, inferred from mitochondrial DNA sequences. Evolution **47**:1192–1202.

VOGLER, A. P., A. WELSH, and T. G. BARRACLOUGH. 1998. Molecular phylogeny of the *Cicindela maritima* (Coleoptera: Cicindelidae) group indicates fast radiation in western North America. Ann. Entomol. Soc. Am. **91**:185–194.

WEBB, T. I., and P. J. BARTLEIN. 1992. Global changes during the last 3 Myr: climatic controls and biotic responses. Annu. Rev. Ecol. Syst. **23**:141–173.

WILLIS, H. L. 1967. Bionomics and zoogeography of tiger beetles of saline habitats in the central United States (Coleoptera: Cicindelidae). Univ. Kansas Sci. Bull. **47**:145–313.

WOLLENBERG, K., J. ARNOLD, and J. C. AVISE. 1996. Recognizing the forest for the trees: testing temporal patterns of cladogenesis using a null model of stochastic diversification. Mol. Biol. Evol. **13**:833–849.

ZAR, J. H. 1984. Biostatistical analysis. Prentice-Hall, Englewood Cliffs, NJ.

ZINK, R. M., and J. B. SLOWINSKI. 1995. Evidence from molecular systematics for decreased avian diversification in the Pleistocene epoch. Proc. Natl. Acad. Sci. USA **92**:5832–5835.

DIETHARD TAUTZ, reviewing editor

Accepted May 16, 2002

Downloaded from https://academic.oup.com/mbe/article/19/10/1706/1258929 by guest on 23 April 2024