# ORIGINAL ARTICLE

# Stasis of functionally versatile specialists

Miriam L. Zelditch,[1,2] Jingchun Li,[3] and Donald L. Swiderski[4]

[1]Museum of Paleontology, University of Michigan, Ann Arbor, Michigan 48109

[2]E-mail: zelditch@umich.edu

[3]Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado 80309

[4]Kresge Hearing Research Institute and Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109

A classic hypothesis posits that lineages exhibiting long-term stasis are broadly adapted generalists that remain well-adapted despite environmental change. However, lacking constraints that steepen adaptive peaks and stabilize the optimum, generalists' phenotypes might drift around a broad adaptive plateau. We propose that stasis would be likely for morphological specialists that behave as ecological generalists much of the time because specialists' functional constraints stabilize the optimum, but those with a broad niche, such as generalists, can persist despite environmental change. Tree squirrels (Callosciurinae and Sciurini) exemplify ecologically versatile specialists, being extreme in adaptations for forceful biting that expand rather than limit niche breadth. Here, we examine the structure of disparity and the evolutionary dynamics of their trophic morphology (mandible size and shape) to determine if they exhibit stasis. In both lineages, a few dietary specialists disproportionately account for disparity; excluding them, we find compelling evidence for stasis of jaw shape but not size. The primary optima of these lineages diverge little, if at all over approximately 30 million years. Once their trophic apparatus was assembled, their morphological specialization steepened the slopes of their adaptive peak and constrained the position of the optima without limiting niche breadth.

**KEY WORDS:** Complex morphologies, disparity, Sciuridae, stasis.

Long-term morphological stasis is commonly viewed as paradoxical, contradicting what we regularly observe over short time scales and therefore expect over long ones: organisms adapt to their changing environments. It is therefore paradoxical when environments change, even dramatically, but organisms do not (Wake et al. 1983). Stabilizing selection may seem to be an obvious cause of stasis, but that process does not explain the most perplexing feature of stasis: persistent constraints on the positions of adaptive peaks (Hansen and Houle 2004; Estes and Arnold 2007). Perhaps the constraints are not on the positions of adaptive peaks but rather are intrinsic to the organism; genetic constraints may even be a universal feature of complex phenotypes when more than one trait is under selection and traits are genetically correlated (Blows and Hoffmann 2005; Walsh and Blows 2009; Hine et al. 2014). Nevertheless, recent studies argue that intrinsic constraints do not explain stasis; standing genetic and mutational variation (Houle et al. 2017) and evolvability (Bolstad et al. 2014) cannot account for evolutionary rates as low as observed. Moreover, intrinsic constraints would only make stasis even more paradoxical because extinction is the likely fate of populations that cannot adapt to their environments. Nonetheless, some lineages singled out as static persist for millions of years, diversifying and colonizing novel environments and occasionally giving rise to strikingly divergent forms (Wake et al. 1983; Emry and Thorington 1984; Meyer 1984).

Stasis has been attributed to multiple causes including stable, persistent environments, generalist ecological strategies, evolutionarily limiting biotic interactions, and low diversification rates, as well as functional constraints that steepen the slopes of adaptive peaks and stabilize the adaptive landscape. Intuitively, stasis might seem most likely in constant environments, such as slowly shifting, long enduring forest belts, especially in the nearly permanent climatic zones of the subtropics and tropics (Simpson 1953, p. 334). Stasis could, however, also occur in changing environments owing to ecological conditions that favor persistence within the ancestral adaptive zone, especially broad niches of generalists (Simpson 1944, 1953). Biotic interactions could also explain stasis, either by affording a shelter from competition or predation by spatial segregation (Darwin 1859; Lindholm 2014), or by persistent competition that locks species

into their ecological roles (Boucot 1983, 1990; Morris et al. 1995; Kozak et al. 2005). The same ecological conditions have been invoked to explain exceptionally low rates of speciation and/or extinction, and stasis has been argued to be a corollary of those low rates: generalists are expected to have lower rates of extinction (Sheldon 1996) or speciation (Eldredge and Cracraft 1980), although specialists, too, could resist extinction if their favored resource persists (Vrba 1987).

What none of these hypotheses adequately explain is precisely what a hypothesis of stasis must explain: constancy of the optimum. In particular, although generalists are often singled out as most likely to be static, because they remain well-adapted within their broad adaptive zone, they need not be constrained to a narrow phenotypic range on that broad adaptive plateau. Yet, without such constraints, their phenotypes might drift across that plateau without being repeatedly pulled back to the ancestral optimum. Perhaps generalists might be constrained to an intermediate phenotype, between specialists' extremes, by a balance among conflicting functional demands; but, in that case, the problem is to explain why one particular balance-point persists as environments change. In the absence of functional constraints, an optimum that persists for millions of years remains difficult to explain. One form of constraints that would be independent of the external environment could account for that stability: internal stabilizing selection that maintains the coherence and functionality of a system of interdependent parts (Wagner and Schwenk 2000; Schwenk and Wagner 2001). But those constraints, such as intrinsic genetic constraints, may be universal and stasis clearly is not.

Another form of constraint, which is neither universal nor characteristic of generalists, is a set of consistent functional demands that limits the array of well-adapted forms such that even modest deviations from the optimum substantially reduce performance and thereby restricts divergence once the optimum is reached (Collar et al. 2009; de Alencar et al. 2017). Such strong functional constraints that steepen the slope of adaptive peaks seem difficult to reconcile with the expectation that generalists are more likely to be static than specialists because they imply specialization. Yet, one class that could be considered ecological generalists might be subject to strong functional constraints: functionally versatile specialists, which have morphologies specialized for a particular function, but behave as ecological generalists much of the time (Liem 1984, 1990). That contrast between morphological specialization and ecological breadth seems as paradoxical as stasis itself because broadening the range of usable resources is expected to prevent specializing on individual ones (Liem 1980). Despite that expectation, some morphological specializations do expand niche breadth, as demonstrated by smasher mantis shrimp that are specialized to consume hard prey but still feed on soft prey (deVries et al. 2016). Specializa-

tions that expand niche breadth increase the range of available resources, resulting in an ecologically broad adaptive zone but steeply sloped adaptive peaks. Functionally versatile specialists, whose specializations expand niche breadth, can remain well-adapted despite environmental change, at an optimum stabilized by functional constraints.

Our primary objective is to determine whether functionally versatile specialists with broad dietary niches have a static trophic morphology. Tree squirrels provide a useful model system in that they are morphologically specialized, being extreme among rodents in their adaptations for forceful biting/gnawing (Druzinsky 2010; Cox et al. 2012; Ercoli et al. 2019), but despite that specialization, they are not restricted to hard foods—their specialization enables them to eat hard foods without compromising their ability to consume soft foods. Most tree squirrels eat not only hard nuts and hard-shelled fruits but also small seeds, pulpy fruits, nectar, insects, and other soft foods (Moller 1983; Emmons and Feer 1997; Roth and Mercer 2008; Thorington et al. 2012). This could explain the famously conservative trophic morphology of tree squirrels. One lineage (*Sciurus*) has even been singled out as an extreme case of stasis, a "living fossil" because of the similarity between living *Sciurus* and the late Eocene *Douglassciurus jeffersoni* in mandibular and ankle morphology (Emry and Thorington 1984). Additionally, some genera of a distantly related lineage (Callosciurinae) are also considered exceptionally conservative, being similar to *Sciurus* in both ecology and trophic morphology, if not in ankle morphology (Emry and Thorington 1982; Emry and Thorington 1984). Another lineage, flying squirrels (Pteromyini), have never been regarded as either conservative or functionally versatile, and they are not hard-nut specialists, but their adaptations for gliding (and nocturnality) are so distinctive and divergent from primitive squirrels that they would not be considered living fossils even if they were, in fact, static once those adaptations arose.

Documenting stasis of complex traits, such as trophic morphology, is not straightforward because stasis is difficult to document convincingly even for simple (one-dimensional) traits. One complication is that lineages evolving at exceptionally low rates occasionally give rise to divergent, often specialized forms (Simpson 1944, 1953), which potentially mask stasis of what Simpson termed the "core lineage." Another is that two models can account for very low evolutionary rates, exceptionally low (Brownian) rates, and a stable adaptive peak, and the more complex model (a stable adaptive peak) might be favored only because conventional criteria for model selection are biased in favor of complex models (e.g., Boettiger et al. 2012; Ho and Ane 2014; Cooper et al. 2016). Likelihood-based methods perform especially poorly when modeling high-dimensional data (Adams and Collyer 2018) but complex morphologies are always high dimensional. Furthermore, no multivariate method can fit a

model containing a mixture of modes, such as a core lineage evolving at a low Brownian rate, but with occasional divergences to other adaptive peaks. Also, even the best-fitting candidate models might fit poorly, yielding unconvincing evidence for stasis. Finally, a complication more specific to our analysis is that a "late burst" of divergence could mimic stasis, yielding a covariance matrix identical to that of a single stationary peak (Uyeda et al. 2015). Diversification rates of *Sciurus* appear to have increased on colonization of the Neotropics (Pečnerová and Martínková 2012; Pečnerová et al. 2015; Zelditch et al. 2015); if that is paralleled by accelerating rates of divergence, it could provide misleading evidence of stasis.

We first reexamine that late burst of diversification, adding recently sequenced Neotropical species and their mandibular morphologies to those previously analyzed (Zelditch et al. 2015; Zelditch et al. 2017). We then examine the structure of disparity in all three lineages to determine if a few distinctive species contribute disproportionately to disparity, as expected if the core lineage is static except for a few extreme phenotypes. We then fit a series of models, first to isolate the potentially static core lineages, and then to determine if those are static rather than evolving at low (Brownian) rates. Finally, we ask whether the core lineages have diverged from each other (and from the late Eocene *D. jeffersoni*). Given the methodological challenge of modeling complex morphologies using likelihood-based methods, we also estimate model misspecification rates.

## *Materials and Methods*
### PHYLOGENY RECONSTRUCTION

A time-calibrated molecular phylogeny was generated for the available tree squirrels. Most of our data came from the previous analysis of diversification rates in Sciuridae (Zelditch et al. 2015) based on five mitochondrial genes (16S, 12S, COII, COIII, and Cyt-b) and three nuclear genes (C-myc, IRBP, and RAG1). We added newly deposited genes of eight species of Neotropical Sciurini (Pečnerová et al. 2015) and eight species of Callosciurinae (Hawkins et al. 2016), including *Sciurus colliaei*, *Sciurus deppei*, *Sciurus flammifer*, *Sciurus gilvigularis*, *Sciurus igniventris*, *Sciurus oculatus*, *Sciurus pyrrhinus*, *Sciurus spadecius*, *Dremomys everetti*, *Callosciurus adamsi*, *Callosciurus inornatus*, *Callosciurus notatus*, *Callosciurus orestes*, *Callosciurus phayrei*, *Sundasciurus altitudinus*, and *Sundasciurus tahan* (see Table S1 for detailed information).

New sequences were aligned with Zelditch et al. (2015) data using CLUSTAL W (Thompson et al. 1994) implemented in Geneious version 11.1.5 (https://www.geneious.com) and corrected by eye. Molecular substitution models of all genes were selected using PartitionFinder (Lanfear et al. 2012) based on Bayesian information criterion. The SYM + I + G model was selected for

Cyt-b, IRBP, and 16S; the GTR + I + G model was selected for other genes.

The phylogeny was reconstructed using BEAST version 1.8.0 (Drummond and Rambaut 2007) on the Cipres Science Gateway (http://www.phylo.org). A relaxed molecular clock with an uncorrelated lognormal distribution was used for each gene partition. A Yule process was used for the speciation model. Two calibration points were used following Mercer and Roth (2003): Sciuridae 36 million years ago (Ma) and *Sciurus* 9.8 Ma. Lognormal priors with mean = 1 and SD = 1 were applied to both calibration points. The third calibration point used in the previous study of Sciuridae was not applied here, because it was based on the age of *Atlantoxerus getulus*, a ground squirrel that is not used in this study.

Three independent Markov chain Monte Carlo (MCMC) analyses were run for 100 million iterations each, and sampled every 10,000 iterations. MCMC performance was examined using *Tracer* version 1.5 (http://beast.bio.ed.ac.uk/Trace) to ensure convergence and reliable effective sampling sizes (>200). Posterior trees from the three runs were combined after burn-in (20% for run1 and 10% for run2, 3) and resampled (i.e., thinning) to 13,000 trees in Log-Combiner (http://beast.bio.ed.ac.uk/LogCombiner). A maximum credibility consensus tree was generated in *TreeAnnotater* (http://beast.bio.ed.ac.uk/TreeAnnotator) and was used for all further analyses.

### DIVERSIFICATION RATES

Dynamics of speciation rate changes implied by the topography of the consensus tree were modeled using Reversible jump MCMC, using *BAMM* 2.5.0 (Rabosky 2014). All commonly recognized extant species were included in this analysis; previous taxonomic and biogeographical studies were used to assign those species that could not be included in the phylogenetic analysis to the smallest possible clade (Supporting Information). Two sets of four chains were run for $10^7$ generations and sampled every $10^4$ generations. Functions in the R (R_Development_Core 2019) package *coda* (Plummer et al. 2006) were used to test for MCMC convergence and effective sample size >200. Functions in *BAMMtools* (version 2.1.6; Rabosky 2014) were used to confirm that rate shifts were found on the same branches of the tree as in Zelditch et al. (2015) and had comparable marginal shift probabilities (previously called branch-specific Bayes Factors).

### MORPHOLOGICAL DATA

The morphological analyses are based on measurements of lower jaws from 822 adults representing 102 of the 148 recognized living species, photographed in lateral view, and one published image of the late Eocene *Douglassciurus jeffersoni*, from the Chadronian White River Formation in the Flagstaff Rim area, Natrona County, Wyoming (USNM 214936; Emry and Korth

**Figure 1.** Landmarks (in black) and semilandmarks (in gray) on the mandible of a representative tree squirrel *Callosciurus notatus*.

1996, p. 778). We used the dataset of Zelditch et al. (2015) (http://datadryad.org/) consisting of 14 landmarks plus 84 semilandmarks included to capture the complex curvature of the jaw (Fig. 1), supplemented by measurements of 15 additional species, including eight Neotropical Sciurini, six Callosciurinae, and one Pteromyini. Our morphological sample of Sciurini is nearly complete (Table 1), but our sample of Callosciurinae is less so and our sample of Pteromyini is sparse (for the full list of species in all analyses, see Table S1; sample sizes for all species are in Table S2).

Landmarks were superimposed by Generalized Procrustes analysis (GPA); semilandmarks were slid to minimize bending energy (Green 1996; Bookstein 1997; Zelditch et al. 2012). Size was measured as (ln-transformed) centroid size (LCS), which is highly correlated with body size (Zelditch et al. 2015). Following superimposition, the mean shape and mean size were computed for each species. GPA was done in *geomorph*, version 3.1.1 (Adams et al. 2019).

## ANALYZING THE STRUCTURE OF DISPARITY

Shape disparity is measured by the average squared Procrustes distance of each species' shape to the mean shape for its clade, equivalent to the sum of variances over all superimposed coordinates except that the denominator is $N - 1$ (Zelditch et al. 2003; Zelditch et al. 2012). Size disparity is measured by the variance of LCS. To determine whether higher disparity is due to most species being far from the mean or to a few highly divergent species, we examine partial disparities, which are the squared deviations of each subgroup (here, each species) from the mean of the group, weighted by the sample size of the subgroup relative to the total group sample size minus one (Foote 1993). Because partial disparities are additive, it is possible to measure the contribution that each species makes to the disparity of its lineage.

## ANALYZING DYNAMICS OF PHENOTYPIC EVOLUTION

To analyze the dynamics of phenotypic evolution, we first used the sole method properly implemented for high-dimensional data, a comparison of Brownian rates of shape and size (Adams 2014). These comparisons use the ratio of the maximum to minimum rate as the test statistic; significance testing was done by phylogenetic simulation using compare.evol.rates in *geomorph*.

We also used maximum likelihood to evaluate a series of models (see MODELS, below), as implemented in the R package *mvMORPH* (Clavel et al. 2015). Models fit to shape require reducing the dimensionality of the data because the number of parameters for complex (multivariate) models can exceed the number of species. We used the first six principal components (PCs), which explain 90.9%, 88%, and 85.6%, of the variance of Callosciurinae, Pteromyini, and Sciurini, respectively. Subsequent PCs explain so little variation that to reach 99% would take 17 PCs for Callosciurinae and Pteromyini and 16 for Sciurini.

We assessed relative support by the Akaike information criterion (AICc) adjusted for small sample size. All models for

**Table 1.** Species richness of each lineage (total in lineage, cf. Thorington et al. 2012), the number of species included in the morphological data (total in sample), the number of species in the phylogenetic analysis (total in tree), and the disparity of shape and size (ln CS) for the entire sample (disparity of sample) and for the species in the tree (disparity of species in tree).

| | Total | | Disparity of sample | | Total | Disparity of species in tree | |
|---|---|---|---|---|---|---|---|
| | Lineage | Sample | Shape | Size | Tree | Shape | Size |
| Lineage | $N$ | | Disparity | | $N$ | Disparity | |
| Callosciurinae | 67 | 44 | 0.00391 | 0.08003 | 42 | 0.00383 | 0.07659 |
| Pteromyini | 44 | 22 | 0.00338 | 0.12359 | 22 | 0.00338 | 0.12359 |
| Sciurini | 37 | 36 | 0.00172 | 0.04983 | 27 | 0.00180 | 0.051045 |

shape were fit with and without constraints; in the case of the Ornstein-Uhlenbeck (OU) models, the PCs are constrained to be adaptively independent (using the decomp = "diagonal" constraint) and in the case of the BM (Brownian Motion) models, the evolutionary rate matrix is a diagonal matrix. The constrained models have the advantage that the optimizer usually converged, whereas it rarely did for the unconstrained models. However, the evolutionary rate matrix is not constrained when fitting OU models and in that sense, the unconstrained BM models are most comparable to the OU models. *mvMORPH* does not report values for $\alpha$ and $\sigma^2$ for multidimensional data, instead providing values for each individual dimension, but the values for multidimensional data can be calculated from the sum of the diagonals of the $\sigma^2$ and stationary variance ($v_y$) matrices because $v_y = \sigma^2/2\alpha$. We can thus calculate the value for $\alpha$, interpretable as the strength of the pull to the optimum, and $\sigma^2$, interpretable as the stochastic component of the evolutionary process due to random genetic drift or the effects of factors not in the model.

Because information-theoretic criteria often favor over-parameterized models over simpler ones, we also assessed support by parametric bootstrapping (Boettiger et al. 2012), simulating data under each model (using the maximum likelihood (ML) parameters estimated by *mvMORPH*), then fitting each model to every simulated dataset, using the difference in log likelihoods ($\delta$) as the test statistic: $\delta = -2(\text{Log } L_0 - \text{Log } L_1)$ where $L_0$ and $L_1$ are the likelihoods of the data simulated under the simpler and the more complex model, respectively. To determine whether the simpler model is better than the complex model, we test the hypothesis that the data came from that simpler model by comparing the difference in log likelihoods ($\delta$) for the original data to the distribution under the simpler model ($L_0$). The proportion of the simulated values under the simpler model that are larger than that observed $\delta$ provides an approximation to the *P*-value for the test, the probability that a difference at least as large would be seen under model $L_0$.

These methods of model selection identify the relatively best-supported of the candidate models but do not assess how well the models fit the data. To assess model adequacy, we used posterior predictive simulation to compare the observed estimate of disparity and those obtained by simulating the data under the models. Because all models of evolutionary dynamics predict disparity, the ability of models to predict observed disparity accurately is evidence of model adequacy. Of special relevance for this analysis is the adequacy of BM as a model for the evolution of shape, given that this is the sole model properly implemented for shape data (Adams and Collyer 2018). Even if it is not the best-supported of the candidate models, BM may be adequate if it predicts disparity nearly as well as the best-supported complex model. Additionally, for the selected models, we compared the estimates of the parameter values ($\sigma^2$ and $\alpha$) obtained from the

data to those obtained by simulating the data under the selected model, then fitted to that model.

## ESTIMATING MODEL MISSPECIFICATION RATES

To determine whether likelihood-based methods yield exceptionally high model misspecification rates given the characteristics of these shape data, reduced in dimensionality, we used the same simulations as those used for parametric bootstrapping to estimate misspecification rates for our data. More specifically, we fit the data to a series of candidate models, used the parameters of those models to simulate 1000 datasets, and then fit each model to each simulated data set (see Supporting Materials, R.script, for details on simulations). The models were first evaluated by selecting the one with the lowest AICc, and then by excluding the equivocal cases, in which the best-supported model differed in AICc by four or fewer units. We restricted these analyses to a subset of models, BM1, OU1, and a sample of the multi-peak OU models.

## MODELS

We fit the three simple models commonly analyzed in studies of phenotypic evolution to shape, constant-rate Brownian motion (BM1), a single stationary peak (OU1), and an Early Burst (EB). We then fit models based on dietary ecology, to the extent that it is known for these groups (for more information on dietary ecology and its relationship to mandibular size and shape in Sciuridae, see Zelditch et al. 2017). Because dietary ecology is not well-understood for the Asian Callosciurinae and Pteromyini, and because theory-based models may fail to find the best model, we also used a heuristic (lasso-based) search for shifts in adaptive optima, implemented in the R package *l1ou* (Khabbazian et al. 2016). To reduce the risk of selecting an over-parameterized model, we compared the fit of the best-supported model to simpler alternatives nested within it. In addition to the multiple-peak OU models, we fit the corresponding multi-rate BM models.

We fit two *a priori* models to Callosciurinae, one derived from a description of the callosciurine ecomorphs found in many Southeast Asian communities (Ellerman 1949; Musser et al. 2010). This was generalized to species other than the ones explicitly named (Fig. 2, Ecomorph). The Ecomorph model differs from all others in that most of the ground-foraging species occupy a single adaptive peak distinct from the others. There are two versions of this model, which differ in one expectation, that *Glyphotes simus* occupies a unique peak. The most complex model fit to Callosciurinae is the one obtained by the heuristic search (Fig. 2, l1OU: OU7). The simpler models nested within it include (1) two two-peak models, each positing a single peak for all species except those in one specialized diet class, either the bark-gouging miniatures (OU2.Mini) or the specialized insectivore, *Rhinosciurus laticaudatus* (OU2.Rhino); (2)
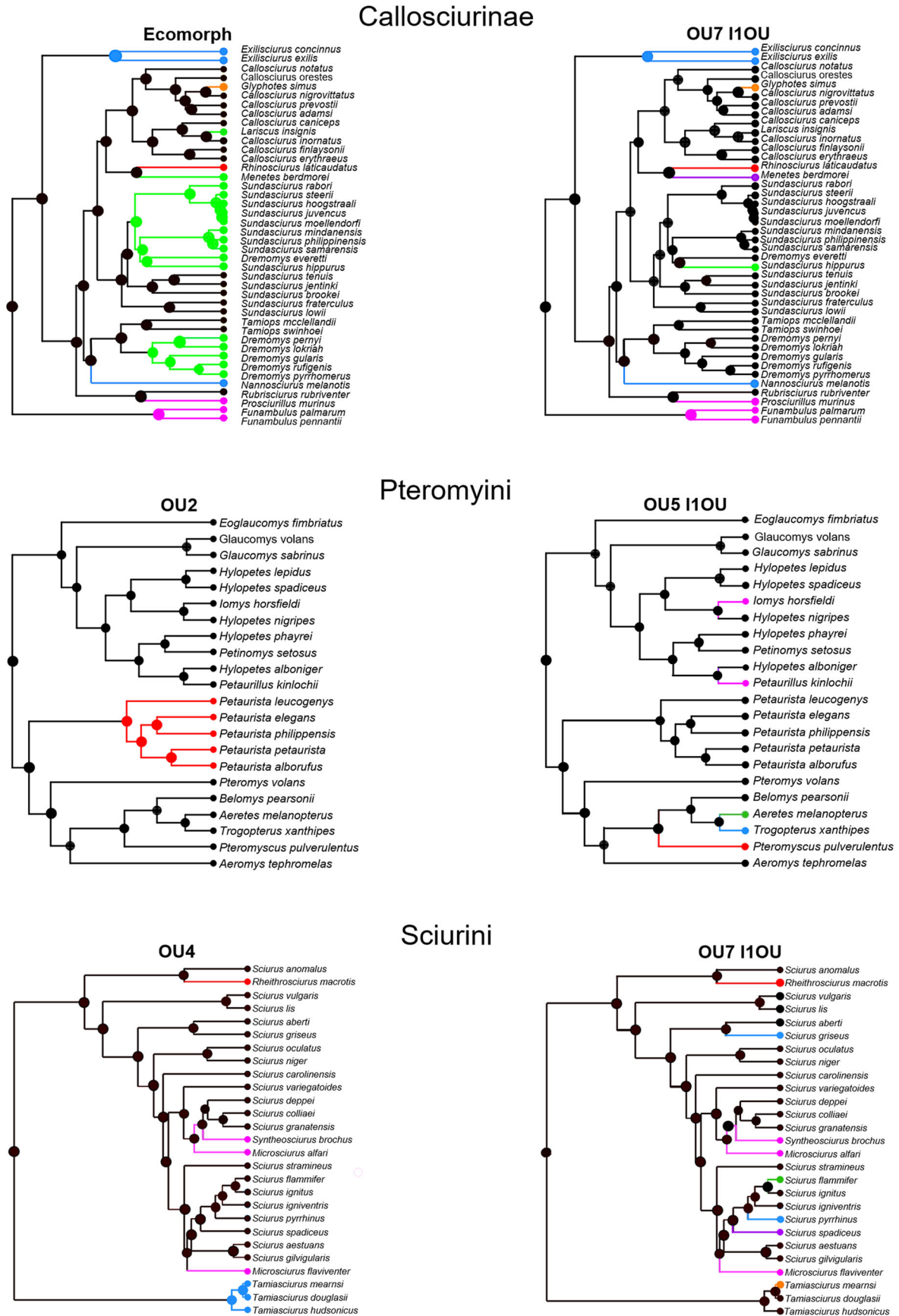
**Figure 2.** Adaptive regime models for tree squirrel jaw shape.

a three-peak model containing both those distinct peaks; (3) a four-peak model that adds a peak for *G. simus* separate from the other bark-gougers' optimum; (4) a five-peak model that adds a distinct optimum for *Sundasciurus hippurus*, which previous analysis found to converge on New World miniatures rather than on more closely related bark-gouging miniatures (Zelditch et al. 2017); and (5) a six-peak model, adding a distinct optimum shared by *Prosciurillus* and *Funambulus*, which feed primarily on soft foods (e.g., fruit, flowers, buds, nectar, caterpillars, and colonial insects; Thorington et al. 2012). The model obtained by the heuristic search, a seven-peak model (OU7.l1OU), adds a distinct peak for *Menetes berdmorei*. The two six-peak models (OU6.1 and OU6.2) both simplify the seven-peak models; one includes *M. berdmorei* within the core lineage (OU6.1) and the other places it on the same peak as *R. laticaudatus* (OU6.2).

Two models were fit to the Pteromyini. One was an ecological model that posits the giant folivores occupy a distinct adaptive peak (Fig. 2, OU2), as suggested by their specialized digestive anatomy (Muul and Lim 1978). The other model was obtained from the heuristic search (Fig. 2, OU5.l1OU).

We fit two models based on dietary ecology to Sciurini, a four-peak model with most species on one peak and separate peaks for the bark-gouging miniatures, the hard endocarp specialist (*Rheithrosciurus macrotis*), and the small conifer-seed and fungi-eating *Tamiasciurus* (Fig. 2, OU4), and a three-peak model that lacks the peak unique to *Tamiasciurus*. A more complex model was obtained by the heuristic search (Fig. 2, OU7.l1OU). In addition, we fit a model that proposes shifts in evolutionary mode on entry into the Neotropics: a late burst of divergence at 5 Ma.

The same models were fit to size, except an additional size model for Callosciurinae was derived by classifying species as "miniature," "small," "intermediate," or "giant," and an additional model for Sciurini was derived by placing *Tamiasciurus* with the miniatures (OU3.MiniTam) rather than with *Sciurus*.

After excluding species occupying peaks of specialists (or those with extreme shapes), we assessed support for a model of stasis for the core lineages of Sciurini and Callosciurinae, then combined the two lineages (recalculating the PCs and reducing the data to seven PCs) to determine whether both lineages occupy the same peak or each lineage occupies a unique peak.

## Results
### PHYLOGENETIC RECONSTRUCTION AND DIVERSIFICATION RATES
The topology and divergence timing of our phylogeny (Supporting Information and Fig. 1) largely agree with those of Zelditch et al. (2015). The addition of new species in *Sciurus*, *Dremomys*, *Callosciurus*, and S*undasciurus* did not affect the main topology,

with Sciurini and Pteromyini as sister groups, and together forming a clade with Callosciurinae. Divergence timing of our major clades is generally 2–3 million years younger than was reported in Zelditch et al. (2015), but the 95% confidence intervals are mostly overlapping. The differences between trees mainly concern weakly supported relationships. Notably several genera that were nested within *Sundasciurus* no longer hold that position. *Nannosciurus melanotis* is now placed as a sister lineage to *Dremomys* and *Tamiops*, although the node support value is still low (Posterior probability [PP] = 0.32). *Menetes* and *Rhinosciurus* are now placed as a sister group to *Callosciurus* (PP = 0.47). The Sulawesi genera *Prosciurillus* and *Rubrisciurus* are also moved out of *Sundasciurus*, now being the next branch after *Exilisciurus* (PP = 1). Most newly added species fall within their respective clades, except for *Dremomys everetti*, which is grouped with *Sundasciurus*, as found by Hawkins et al. (2016). Overall, relationships within Sciurinae are much better resolved with strong support compared to Callosciurinae, where the phylogenetic positions of some genera are still unstable.

The model of speciation rates that best fits the consensus tree has a relatively stable rate through most of the evolutionary history of tree squirrels, and two rate increases in the last 5–10 Ma (Fig. 3). The more strongly supported increase is in the New World branch of Sciurini. It is not clear whether this increase precedes the divergence of western North American *Sciurus* from the rest of the lineage; however, the highest speciation rates are inferred for a narrow window of time after that divergence when the main Neotropical lineages appeared (∼4–6 Ma). The much less strongly supported increase is in a *Sundasciurus* lineage (*Sundasciurus steerii* and relatives) that diversified on Palawan and adjacent islands very recently (<2 Ma).

### THE STRUCTURE OF DISPARITY
#### Shape
In the plane depicting the main dimensions of shape disparity (Fig. 4), the range of Callosciurinae is by far the largest owing to the extreme morphology of the specialized insectivore, *R. laticaudatus*. Over all dimensions, disparities of Callosciurinae and Pteromyini are nearly equal and about twice that of Sciurini (Table 1, "Disparity of Sample"). The distribution of distances to the mean (Fig. 5A) and the proportional contribution of each species to the total of its group (Fig. 5B) are highly skewed. A few outliers make large contributions to disparity of Callosciurinae: the specialized insectivore contributes 22% and the bark-gouging miniatures contribute another 15.6%, increasing to 44% with *G. simus* included. Those five species account for 11.4% of diversity but nearly half the disparity. The most distinctive shapes in Callosciurinae are both more extreme and more numerous than those of the other lineages. There are no extreme morphologies in Pteromyini, which has the highest median
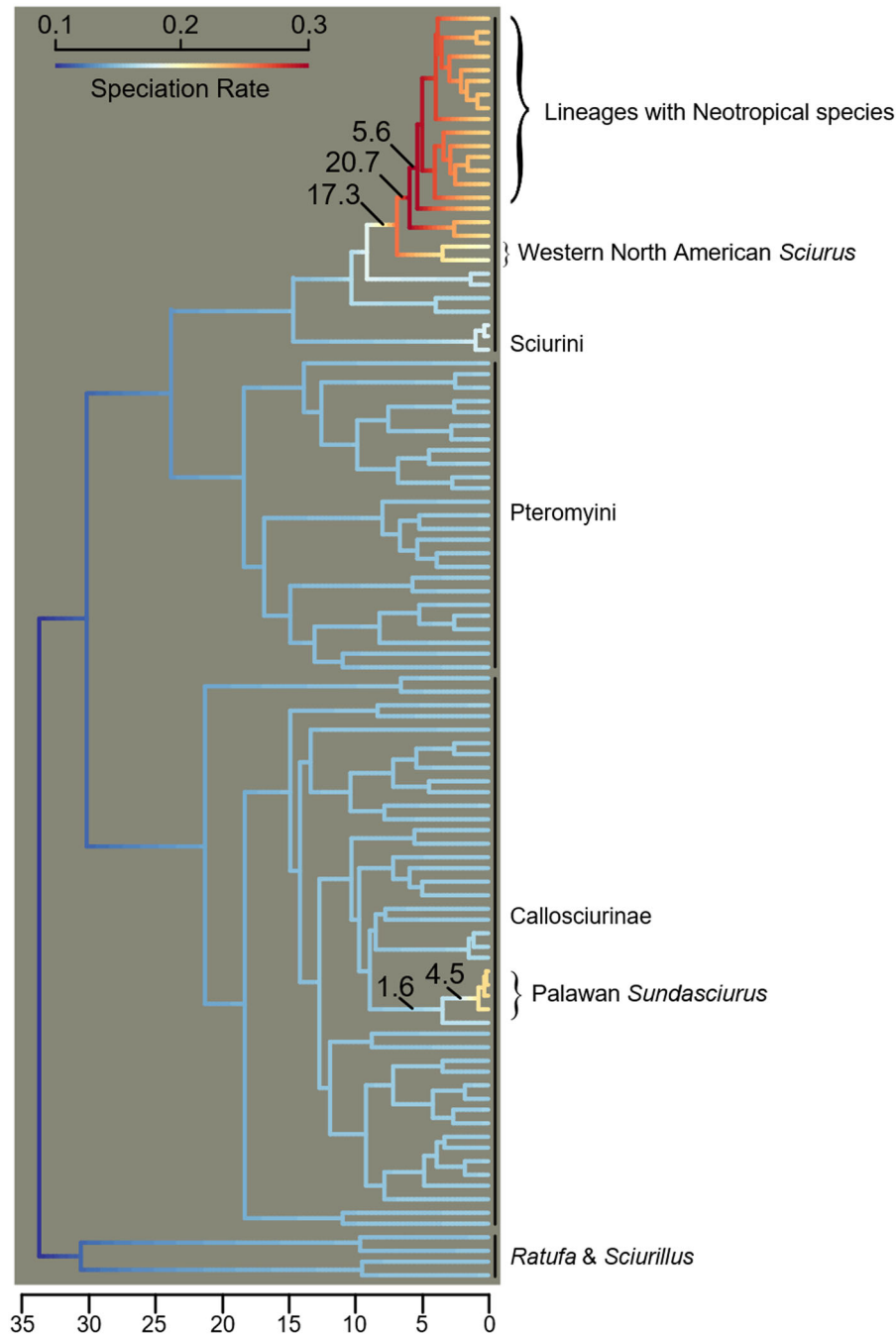
**Figure 3.** Consensus tree with inferred rate shifts (color bar) and marginal shift probabilities.
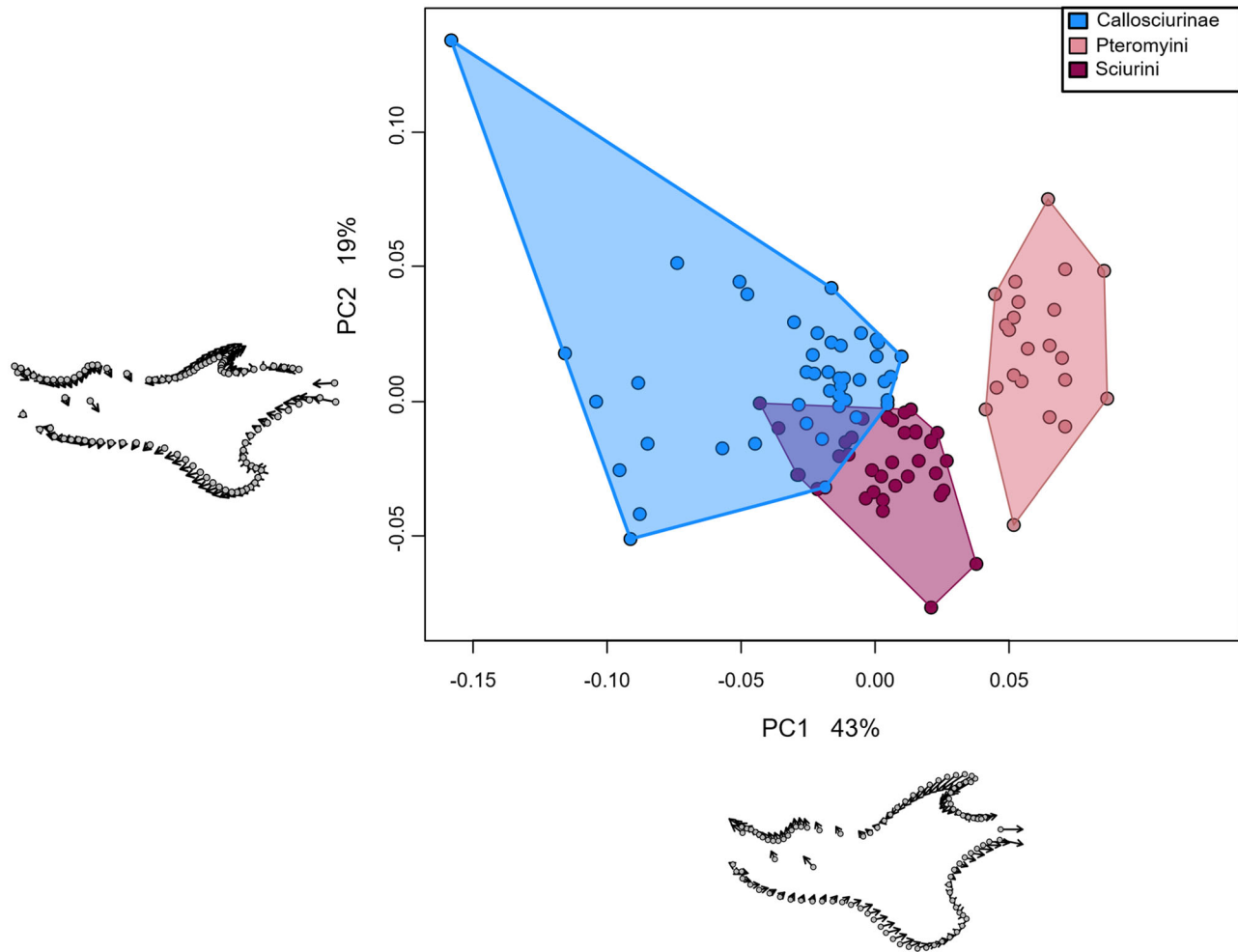
disparity. In Sciurini, the hard-endocarp specialist (*R. macrotis*) is the sole outlier, contributing 13% of the total disparity of that lineage. The four miniatures account for another 20% of that total. Thus, in Sciurini, five dietary specialists contribute 14% of the diversity but 33% of the disparity.

*Size*

Disparity of size for Callosciurinae is intermediate between that of Sciurini and Pteromyini (Table 1). The distribution of distances

to the mean (Fig. 5C) and the proportional contribution of each species to the disparity of its group (Fig. 5D) are skewed, but there are fewer outliers in size than in shape. Even so, miniatures contribute disproportionately to size disparity of Callosciurinae (22.6%) and one large-bodied species (*Rubrisciurus rubriventer*) contributes as much as some miniatures (3.3%). There are no outliers in Pteromyini, and, with one exception (*Petaurillus kinlochii*), no species contributes even as much as 3% to the total. In Sciurini, size disparity, such as shape disparity, is due primarily

**Figure 4.** The principal components of shape disparity of the three tree squirrel clades.

to the distinctive giant, hard-endocarp specialist (*R. macrotis*) and the miniatures, which are less extreme than miniature callosciurines and jointly contribute only 11.1% to sciurine size disparity.

## EVOLUTIONARY DYNAMICS
### Comparing Brownian evolutionary rates
Based on the ratios of rates for the one evolutionary model properly implemented for shape, Brownian rates for shape and size do not differ significantly among lineages ($P = 0.302$ and $0.262$ for shape and size, respectively).

### Evolutionary dynamics for shape
Models that constrain dimensions to be adaptively independent (in the case of the OU models), or to evolve independently (in the case of the BM models), rarely lose information relative to the unconstrained models (Table S3). The notable exception is the EB model; the unconstrained EB model invariably improves upon the constrained one. Both constrained and unconstrained versions of the models fit to Callosciurinae yield the same conclusions:

the three simple models all fit poorly, as do all Brownian models, and the seven-peak l1ou model is unnecessarily complex, fitting no better than the simpler six-peak variants of it, all of which improve substantially on the Ecomorph model (Table 2). In striking contrast, the best-supported models for Pteromyini are Brownian, either the four-rate BM model (if the rate matrix is constrained to be diagonal) or the constant rate BM model (if that matrix is unconstrained), although the constrained four-rate model is the best-supported. For Sciurini, the most complex model OU7.1lOU does not improve upon the simpler *a priori* ecological OU3 model, which is clearly the best-fitting unconstrained model.

Parametric bootstrapping largely supports the conclusions drawn for Callosciurinae and Sciurini, the two candidate static lineages. For Callosciurinae, the exception is that parametric bootstrapping favors OU7 rather than OU6; however, the distributions of δ for those models overlap (Fig. 6). The results are more equivocal for Sciurini because the three-peak model is strongly favored over only one of the two-peak models (OU2.Mini). Even so, only 7.5% of the values for δ under the simpler OU2.Rheithro
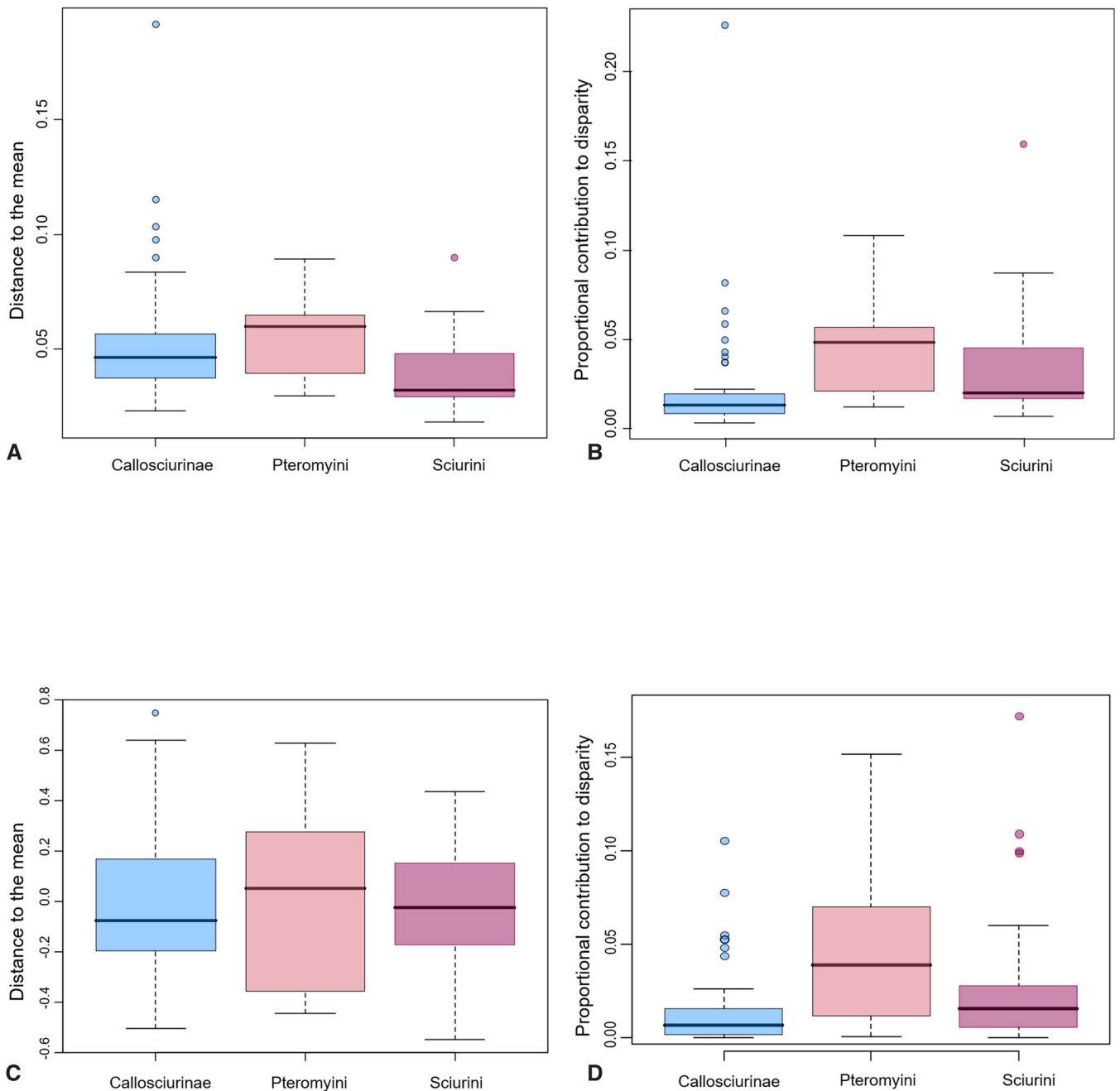
**Figure 5.** The distribution of distances to the means (for size and shape) and the proportional contribution that each species makes to the disparity of its lineage. (A) Distribution of Procrustes distances from each species to the mean shape. (B) Proportional contribution that each species makes to the total shape disparity. (C) Distribution of deviations from each species to the mean size (LCS). (D) Proportional contribution that each species makes to the total size disparity.

model exceed the observed δ (corresponding to a P-value of 0.075). Posterior predictive simulations find that a single rate Brownian motion model substantially overestimates shape disparity, predicting values up to twice the observed (Table 3). In contrast, all OU models predict disparities closer to the observed value, although only in two cases does the observed value lie within the confidence interval of the simulated values (OU3 fit

to both lineages). The parameter estimates for the selected OU models (Table 4) indicate a moderate to strong pull to the optimum in both Callosciurinae and Sciurini, although the estimates from the data are outside the confidence intervals of the simulated values. The parameter estimates for the four-rate Brownian model fit to Pteromyini are much further from the observed values, sometimes an order of magnitude lower (Table S4).
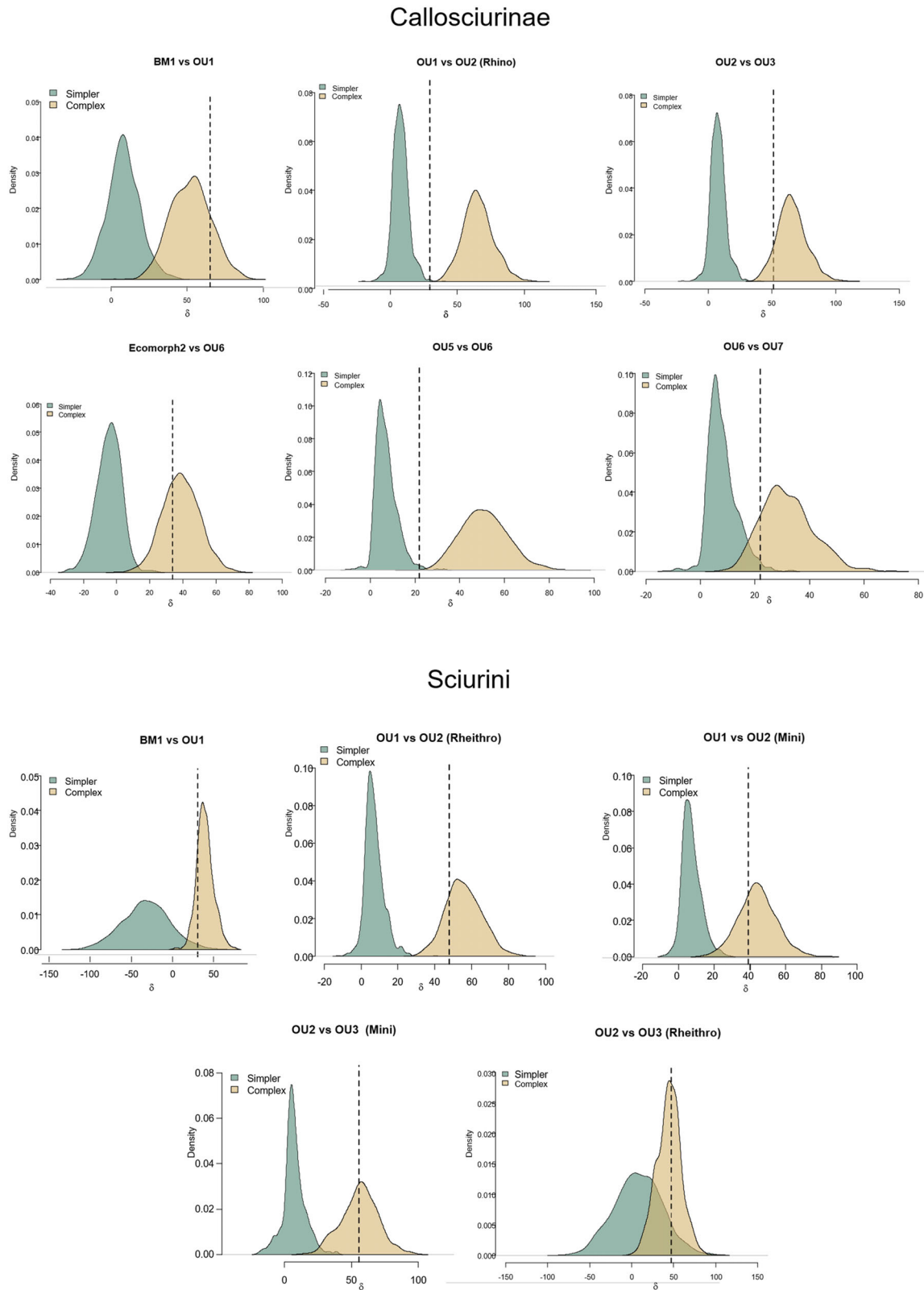
**Figure 6.** The distribution of δ under the simple and complex models for shape in the two lineages that are candidates for stasis. Shape data are simulated using the maximum likelihood estimate (MLE) parameters for the simpler model (M0), then fit to both that model and the more complex model (M1), and then simulated using the MLE parameters for the more complex model and fit to both the simpler and that more complex model. The distribution of δ (2 (log L0 − log L1)) from the simulations is compared to the δ obtained from the data (vertical line).

**Table 2.** Relative support for models fit to mandibular shape. Except for BM1 (uncon), the models incorporate the constraint decomp = "diagonal." Shown are the *a priori* models, the model discovered by l1ou and simplified versions of that l1OU model within $\Delta$AICc < 50 of the best constrained model and the corresponding BM models.

Callosciurinae

| Model | Pars | logLik | AICc | $\Delta$AICc |
|---|---|---|---|---|
| BM1 (uncon) | 27 | 668.64 | −1276.53 | 173.40 |
| BM1 | 12 | 654.76 | −1284.22 | 165.71 |
| OU1 | 33 | 687.29 | −1298.29 | 151.65 |
| EB | 28 | 654.76 | −1246.24 | 203.69 |
| OU2.R | 39 | 702.04 | −1311.35 | 136.75 |
| OU3.mR | 45 | 727.594 | −1345.09 | 103.01 |
| OU5.Ecomorph3 | 63 | 791.65 | −1414.40 | 35.54 |
| OU5.mRGS | 57 | 787.903 | −1427.72 | 22.21 |
| OU6.1 | 63 | 808.50 | −1448.10 | 1.84 |
| OU6.2 | 63 | 809.41 | −1449.93 | 0.00 |
| OU7.l1ou | 69 | 819.44 | −1447.80 | 2.14 |
| BM2.R | 18 | 663.66 | −1288.39 | 159.71 |
| BM3.mR | 24 | 669.31 | −1285.34 | 162.76 |
| BM5.Ecomorph3 | 42 | 715.40 | −1329.52 | 120.41 |
| BM5.mRGS | 36 | 713.03 | −1341.67 | 108.26 |
| BM6.1 | 42 | 718.06 | −1334.83 | 113.27 |
| BM6.2 | 42 | 712.86 | −1324.44 | 123.66 |
| BM7.l1ou | 48 | 714.24 | −1309.30 | 140.64 |

Pteromyini

| Model | pars | logLik | AICc | $\Delta$AICc |
|---|---|---|---|---|
| BM1 (uncon) | 27 | 336.86 | −605.18 | 47.53 |
| BM1 | 12 | 323.16 | −619.71 | 5.72 |
| OU1 | 33 | 350.62 | −612.34 | 44.37 |
| EB | 28 | 323.16 | −574.56 | 82.14 |
| OU2.foli | 39 | 356.89 | −601.86 | 54.85 |
| OU4 | 51 | 391.58 | −614.85 | 38.02 |
| OU5.l1ou | 57 | 398.21 | −593.073 | 59.80 |
| BM2.foli | 18 | 333.74 | −625.427 | 31.28 |
| BM4 | 30 | 365.56 | −652.71 | 0.00 |
| BMM5.l1ou | 36 | 367.57 | −635.10 | 17.78 |

Sciurini

| Model | pars | logLik | AICc | $\Delta$AICc |
|---|---|---|---|---|
| BM1 (uncon) | 27 | 446.28 | −827.27 | 65.49 |
| BM1 | 12 | 429.15 | −832.21 | 60.54 |
| OU1 | 33 | 461.58 | −839.63 | 53.13 |
| EB | 28 | 429.15 | −790.09 | 102.66 |
| OU2.Reithro | 39 | 485.56 | −867.55 | 25.20 |
| OU2.Mini | 39 | 481.12 | −858.67 | 34.09 |

*Continued*

**Table 2.** Continued.

Sciurini

| Model | pars | logLik | AICc | $\Delta$AICc |
|---|---|---|---|---|
| OU3 | 45 | 509.01 | −892.32 | 0.43 |
| OU4.Tam | 51 | 510.55 | −870.88 | 21.87 |
| OU6.l1ou | 69 | 567.88 | −892.76 | 0.00 |
| ER | 33 | 456.89 | −830.26 | 62.5 |
| BM2.Rheithro | 18 | 435.907 | −831.03 | 61.73 |
| BM2.Mini | 18 | 433.30 | −825.81 | 66.94 |
| BM3 | 24 | 442.67 | −828.58 | 64.18 |
| BM4.Tam | 30 | 445.08 | −815.96 | 76.80 |
| BM6.l1ou | 48 | 496.27 | −854.91 | 37.85 |

## Evolutionary dynamics for size

For Callosciurinae, one model, derived by classifying species by size, fits far better than the others, including the models that fit shape well (Table 5). For Pteromyini, the best-fitting model is the two-rate Brownian model, with a dramatic reduction in the rate of size evolution of giant folivores; nonfolivores evolve at 16 times the rate of folivores. This model only slightly improves upon BM1 ($\Delta$AICc = 4.29) but the parametric bootstrap (Fig. 7) shows that they are distinguishable. Also, the observed value for size disparity (0.124) is relatively far outside the confidence interval for the data simulated under BM1 (0.111–0.118) but within the confidence interval for the data simulated under BM2 (0.119–0.127). For Sciurini, four models fit equally well: two models with three peaks (one with separate peaks for giant *R. macrotis*, the miniatures, and *Sciurus* plus the smaller *Tamiasciurus* (OU3) and one that differs in placing *Tamiasciurus* on the same peak as miniatures (OU3.TamMini). The four-peak model has a peak unique to *Tamiasciurus*. The most complex model is the OU7.l1ou model (for shape). As evident from the distribution of $\delta$ (Fig. 7), the OU3 and OU4 models substantially overlap. The observed value of size disparity, 0.051, lies within the confidence intervals for data simulated under OU3 (0.049–0.051) and OU4 (0.050–0.051), indicating that either is adequate. The intermediate size of *Tamiasciurus* is not sufficiently different from either the larger *Sciurus* or the smaller miniatures to warrant a peak unique to it, but the results are ambiguous regarding the position of *Tamiasciurus* in a three-peak model. The estimates of the two Brownian rates from the data for Pteromyini, $\sigma_1^2 = 0.168$ and $\sigma_2^2 = 0.0104$, are close to the means obtained from the simulations: $\sigma_1^2 = 0.162$ (CI: 0.158–0.165) and $\sigma_2^2 = 0.010$ (CI: 0.0095–0.0104) even if slightly outside the confidence interval in the case of $\sigma_1^2$. The values of $\alpha$ estimated from the data are high, and in the case of Callosciurinae, $\alpha$ lies within the confidence interval of the simulated values, although the value for Sciurini is again outside the confidence interval (Table 6).

**Table 3.** Posterior predictive simulations of disparity. The observed values are obtained from the variance of the six principal components used in modeling the dynamics of shape. The values for the models are the mean and 95% confidence intervals (CI) of the 1000 samples simulated under each model. All values are multiplied by $10^2$.

| Callosciurinae | Disparity | | Sciurini | Disparity | |
|---|---|---|---|---|---|
| Observed | 0.348 | | Observed | 0.150 | |
| Model | Simulated | CI | Model | Simulated | CI |
| BM1 | 0.496 | 0.489–0.502 | BM1 | 0.320 | 0.316–0.338 |
| OU1 | 0.373 | 0.369–0.377 | OU1 | 0.158 | 0.156–0.160 |
| OU2.R | 0.382 | 0. 378–0.386 | OU2.Rheithro | 0.160 | 0.159–0.161 |
| OU3.mR | 0.349 | 0.346–0.351 | OU2.Mini | 0.147 | 0.149–0.151 |
| OU6.1 | 0.332 | 0.333–0.334 | OU3 | 0.156 | 0.155–0.158 |
| OU7.11ou | 0.344 | 0.342–0.346 | BM3 | 0.290 | 0.285–0.295 |
| OU.Ecomorph3 | 0.333 | 0.331–0.335 | | | |

**Table 4.** Estimates of the parameters of the selected OU models, $\sigma^2$ and $\alpha$ for shape. Given are the values obtained from the data and the mean or median (indicated by *) and confidence intervals obtained from simulations on a tree scaled to unit height.

| | Data | | Simulations | | | |
|---|---|---|---|---|---|---|
| | $\sigma^2$ | $\alpha$ | Mean $\sigma^2$ | CI | Mean $\alpha$ | CI |
| Callosciurinae | 0.011 | 5.08 | 0.010 | 0.00976–0.0097 | 5.26 | 5.14–5.38 |
| Sciurini | 0.021 | 11.08 | 0.025 | 0.0237–0.0257 | 14.26* | 13.88–14.76 |

## ARE THE CORE LINEAGES STATIC?

### Shape

When extreme morphologies are excluded, a constrained single peak (OU1) model fits shape substantially better than other models, including a constrained single rate Brownian (BM1) model (Table 7). Parametric bootstrapping again supports those conclusions for both core callosciurines and core sciurines (Fig. 8). None of the values for $\delta$ under the simulated simple unconstrained BM1 model equal or exceed the observed value. Estimates of disparity from the simulations show that BM1 grossly overestimates disparity for Sciurini (0.00311), yielding a mean nearly three times the observed value (0.00311 vs. 0.00108). In contrast, simulations under OU1 underestimate the observed value (0.00099 vs. 0.00108), which is only slightly outside the confidence interval (0.00098–0.00100). Similarly, BM1 grossly overestimates disparity of core callosciurines, yielding a mean over the simulations of more than twice the observed value (0.00247 vs. 0.00107), whereas the mean of the simulations under OU1 is very close to the observed value (0.00109 vs. 0.00107), which is just slightly outside the confidence interval (0.00108–0.00110). The parameter values estimated from the data on a tree scaled to unit height for both the core callosciurines ($\sigma^2 = 0.006$, $\alpha = 5.6$) and sciurines ($\sigma^2 = 0.016$, $\alpha = 7.4$) also clearly support an OU model over BM, although they are outside the confidence intervals estimated from the simulated data (for the core callosci-

urines, $\sigma^2 = 0.00948$–0.00967 and $\alpha = 4.66$–4.68; for the core sciurines, $\sigma^2 = 0.0106$–0.011 and $\alpha = 5.68$–5.69).

### Size

Size exhibits more complex dynamics. The single-rate Brownian model fits as well as a single peak model ($\Delta AIC = 2.44$ in favor of BM1), and, under the single peak model, $\alpha = 3.2 \times 10^{-9}$ on a tree scaled to unit height. For the core sciurines, the best-fitting model (found by 11ou) is neither a single peak nor a single rate model but rather it has two peak shifts, one at the common ancestor of *S. gilvigularis* and *Sciurus aestuans*, another along the branch to *Sciurus ignitus*. However, the estimate for $\alpha$ for that model also indicates a very weak pull to the optimum, $\alpha = 1.31$ and the phylogenetic half-life, $t_{1/2} = 7.8$ Ma, greater than half the age of the crown group. Moreover, the range of jaw size also is wide in the core lineages (60.84–111.66 mm and 84.67–141.34 mm in core callosciurines and sciurines, respectively).

## ONE OR TWO PRIMARY OPTIMA FOR SHAPE FOR CORE LINEAGE OF TREE SQUIRRELS?

The hypotheses of one and two primary shape optima receive nearly equal support (Table 8). The parametric bootstrap clearly favors the simpler model (Fig. 9), but that support for a single-optimum is countered by evidence that the lineages do not overlap within the plane of the first two PCs (Fig. 10A) and only
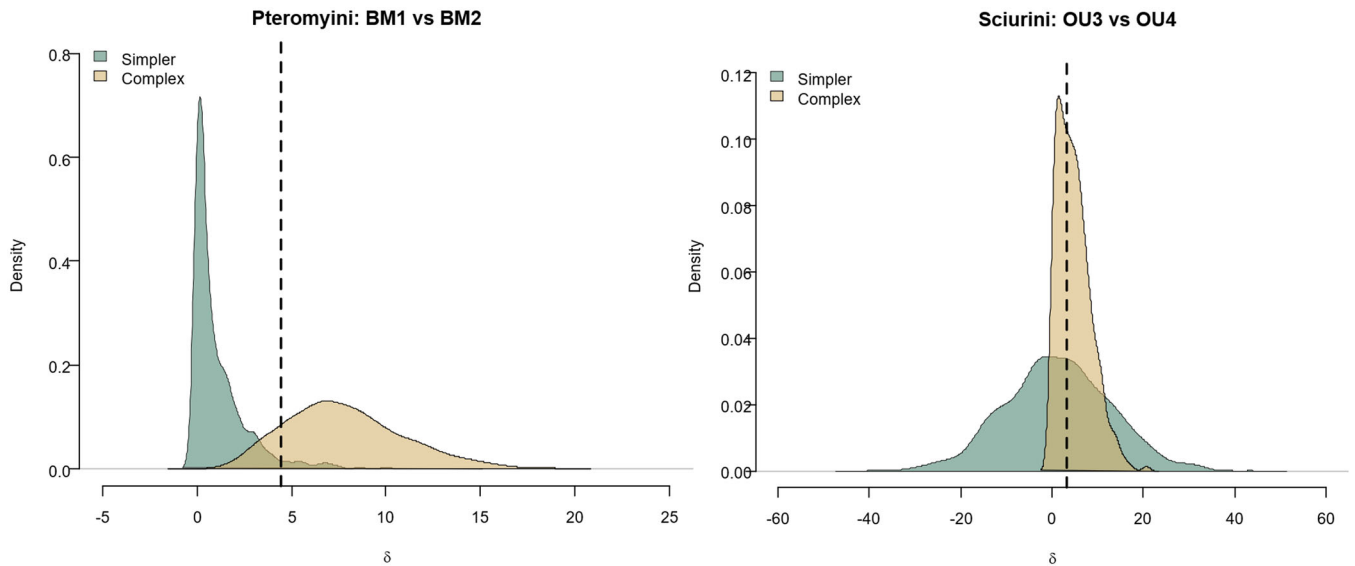
**Figure 7.** The distribution of $\delta$ under the simple and complex models for shape in the two lineages that are candidates for stasis. Shape data are simulated using the maximum likelihood estimate (MLE) parameters for the simpler model (M0), then fit to both that model and the more complex model (M1), and then simulated using the MLE parameters for the more complex model and fit to both the simpler and that more complex model. The distribution of $\delta$ (2 (log L0 − log L1)) from the simulations is compared to the $\delta$ obtained from the data (vertical line).
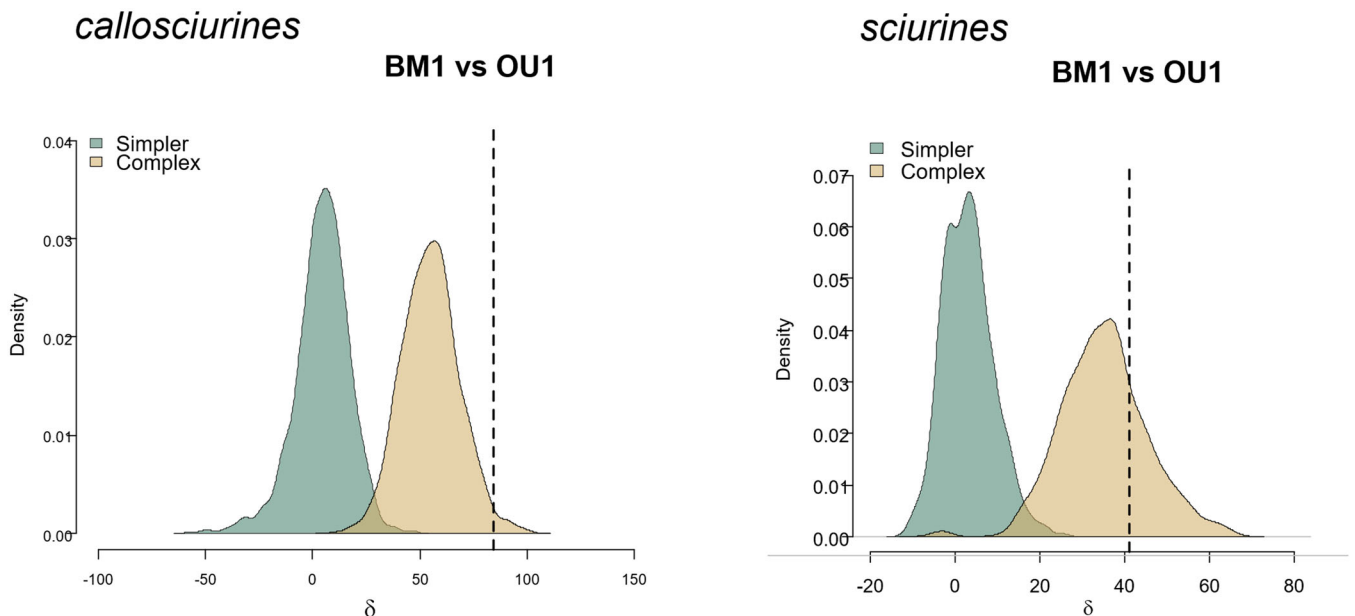


**Figure 8.** The distribution of $\delta$ under the simple and complex models for low evolutionary rates, a single rate Brownian motion (BM1) and a single stationary peak (OU1), for shape in the core lineages of callosciurines and sciurines after excluding the few divergent species.

slightly on the phylogenetic PCs (extracted given a tree rescaled by the estimated value for $\alpha$ on the one-peak tree) (Fig. 10B). Under both models, the observed disparity of the two core lineages (0.00173) is very close to, but slightly outside the confidence interval for the data simulated under both a single optimum (0.00167–0.00170) and two optima (0.00175–0.0.00178). Under

the one peak model, $\alpha = 6.59$; under the two-peak model, $\alpha = 15.78$. These values also are outside the confidence intervals for the mean of the simulated values, more so for the single peak model (3.92–4.01) than the two peak model (14.49–15.13). The two models, however, differ little in their estimates of the duration of stasis in that the time it would take to move halfway from

**Table 5.** Relative support for models fit to mandibular size. Shown are the a priori models, the model discovered by 1lou and simplified versions of that model that are within ΔAICc < 50 of the best model.

Callosciurinae

| Model | pars | logLik | AICc | ΔAICc |
|---|---|---|---|---|
| BM1 | 2 | 0.72 | 2.87 | 52.63 |
| OU1 | 3 | 1.32 | 4.00 | 53.75 |
| EB | 3 | 0.72 | 5.20 | 54.95 |
| OU5.Ecomorph3 | 8 | 23.62 | −26.87 | 22.88 |
| OU7.l1ou | 9 | 23.48 | −23.33 | 26.42 |
| OU6.1 | 8 | 23.48 | −26.59 | 23.16 |
| OU6.2 | 8 | 23.35 | −26.34 | 23.41 |
| OU.Size.Class | 6 | 32.08 | −49.75 | 0.00 |
| BM5.Ecomorph3 | 6 | 13.38 | −26.87 | 40.29 |
| BM7.l1ou | 7 | 12.20 | −7.11 | 42.64 |
| BM6.1 | 7 | 12.20 | −7.11 | 42.64 |
| BM6.2 | 7 | 12.63 | −7.96 | 41.79 |
| BM.Size.Class | 5 | 13.86 | −16.07 | 33.69 |
| Pteromyini | | | | |
| BM1 | 2 | −2.43 | 9.50 | 4.26 |
| OU1 | 3 | −2.12 | 11.56 | 6.33 |
| EB | 3 | −2.43 | 12.20 | 6.96 |
| OU2.foli | 4 | −0.21 | 10.77 | 5.54 |
| OU4 | 6 | −0.72 | 14.51 | 9.27 |
| OU5.l1ou | 8 | 4.90 | 17.27 | 12.03 |
| BM2.foli. | 3 | 1.05 | 5.24 | 0.00 |
| BM5.l1ou | 7 | 0.19 | 21.63 | 16.39 |
| BM4 | 5 | −0.38 | 14.51 | 9.27 |
| BM5.l1ou | 7 | 0.19 | 21.63 | 16.39 |
| Sciurini | | | | |
| Model | pars | LogLik | AICc | ΔAIC |
| BM1 | 2 | −0.32 | 5.13 | 22.75 |
| OU1 | 3 | 3.12 | 0.81 | 18.43 |
| EB | 3 | −0.32 | 7.68 | 25.30 |
| OU2.Rheithro | 4 | 6.34 | −2.86 | 14.75 |
| OU2. Mini | 4 | 9.24 | −8.66 | 8.96 |
| OUM.3 | 5 | 13.67 | −14.46 | 3.16 |
| OU3.MiniTam | 5 | 15.24 | −17.62 | 0.00 |
| OU4 | 6 | 15.28 | −14.36 | 3.26 |
| OU6.l1ou | 7 | 17.36 | −14.83 | 2.79 |
| BMEB | 3 | 2.79 | 3.36 | 19.09 |
| BM2.Rheithro | 3 | 0.57 | 5.91 | 23.53 |
| BM2.Mini | 3 | 0.99 | 5.06 | 22.68 |
| BM3 | 4 | 2.36 | 5.10 | 22.72 |
| BM3.MiniTam | 4 | 1.28 | 7.26 | 24.88 |
| BM4.RMiniTam | 5 | 2.76 | 7.34 | 24.95 |
| BMM6.l1ou | 6 | 19.97 | −11.35 | 6.26 |

## OU1 vs OU2
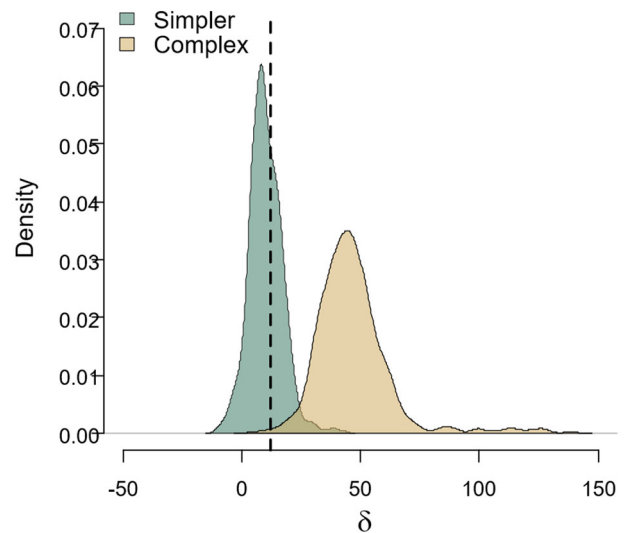


**Figure 9.** The distribution of δ under the simpler model of a single peak occupied by both lineages and the more complex model of two peaks.

the ancestral state to the optimum $t_{1/2}$ (Hansen 1997) is merely 1.2 Ma on a tree of ca 30 Ma.

The two models are so difficult to distinguish because, if there are two optima, they are very close to each other (Procrustes distance, $D_p = 0.066$). That distance, estimated from the values for $\theta$ in the space of seven PCs, is very close to the distance between the lineage means within the shape space of full dimensionality ($D_p = 0.054$). The slight difference between optimal shapes for the two lineages (Fig. 10C) is in robustness of the horizontal ramus and the robustness, orientation, and curvature of the coronoid process. The core lineage(s) differ more strikingly from *D. jeffersoni*, and in the same direction (Fig. 11A). Under the hypothesis of a single optimum, $D_p = 0.0976$ between that optimum and *D. jeffersoni*; under the hypothesis of two optima, $D_p = 0.112$ and 0.102 between *D. jeffersoni* to the callosciurine and sciurine optima, respectively. The shape features that most distinguish the living species from *D. jeffersoni* are their lesser robustness, especially of the angular process and distal diastema, as well as a more subtle but functionally relevant displacement of the masseteric fossa, positioned more anteriorly in the living species (Fig 11B and 11C).

## MODEL MISSPECIFICATION RATES

For each clade, data were simulated under each model (for that clade) and fit to each model for that clade. Each column of Table 9 contains the proportion of cases in which data simulated under one model had the lowest AICc when fit to that (true) model and every other one. Considering all cases, including

**Table 6.** Estimates of the parameters of the selected OU models, $\sigma^2$ and $\alpha$, for size. Given are the values obtained from the data and the mean or median (indicated by *) and confidence intervals obtained from simulations on a tree scaled to unit height.

| | Data | | Simulations | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | $\sigma^2$ | $\alpha$ | Mean $\sigma^2$ | CI | Mean $\alpha$ | CI |
| Callosciurinae | 0.418 | 15.78 | 0.398 | 0.368–0.428 | 16.71 | 15.49–17.93 |
| Sciurini | 0.407 | 10.39 | 0.220 | 0.202–0.247 | 7.45 | 6.75–8.15 |

**Table 7.** Relative support for the alternative models for exceptionally low evolutionary rates: BM1 and OU1 fit to the core callosciurines and sciurines.

| Model | pars | logLik | AICc | $\Delta$AICc |
| --- | --- | --- | --- | --- |
| **Core callosciurines** | | | | |
| BM1 (uncon) | 27 | 549.76 | −1036.30 | 53.58 |
| BM1 | 12 | 543.49 | −1061.24 | 28.64 |
| OU1 | 33 | 585.04 | −1089.88 | 0 |

| Model | pars | logLik | AICc | $\Delta$AICc |
| --- | --- | --- | --- | --- |
| **Core sciurines** | | | | |
| BM1 (uncon) | 27 | 395.34 | −722.92 | 30.72 |
| BM1 | 12 | 381.63 | −736.77 | 16.87 |
| OU1 | 33 | 420.61 | −753.64 | 0 |

**Table 8.** Relative support for hypotheses of one or two primary optima for shape.

| Model | pars | logLik | AICc | $\Delta$AICc |
| --- | --- | --- | --- | --- |
| OU1 | 52 | 1346.21 | −2573.62 | 5.16 |
| OU2 | 60 | 1359.47 | −2578.78 | 0 |

equivocal ones (when $\Delta$AICc $\leq$ 4), model misspecification rates are frequently higher than 5%; in two cases, they are close to 30% (Table 9A). As expected, when the true model is not preferred, it is usually a more complex model that is favored. In the most extreme cases (in Sciurini), the OU3 model is frequently preferred for data simulated under the OU2.Rheithro model; however, for data simulated under the OU3 model, the misspecification rate is also high but the simpler OU2.Mini model is most often preferred over the more complex true model. In the other case of misspecification rates approaching 30%, the OU7 model in Callosciurinae, the simpler OU6 model is most often incorrectly favored. Excluding equivocal cases in which $\Delta$AICc $\leq$ 4.0 (Table 9B), model misspecification rates are lower but the same two models have very high misspecification rates and the same incorrect models are favored in most cases: OU3

over OU2.Rheithro in Sciurini (30%) and OU6 over OU7 in Callosciurinae (27%), consistent with parametric bootstrapping. The competing models are very similar because the optima are very close. The miniature sciurines are not as extreme as miniature callosciurines (or as extreme as the hard-endocarp specialist) so two peaks in the OU3 model are close to each other. Similarly, the only difference between the OU6 and OU7 models for callosciurines is the position of one species, *M. berdmorei*.

## Discussion

Persistence of an optimum over millions of years is challenging to document much less to explain, but our analysis provides compelling evidence for long-term stasis of jaw shape in two lineages of tree squirrels (Callosciurinae and Sciurini). Not only are both lineages static but also their optima have diverged little, if at all, over approximately 30 million years. In both static lineages, approximately 15% of the species, typically dietary specialists, have diverged from the phenotype characteristic of the morphologically specialized dietary generalists. The dietary specialists contribute disproportionately to the disparity of both groups, accounting for nearly half of the disparity of Callosciurinae. In contrast, the flying squirrels (Pteromyini) are not static and no flying squirrel is strikingly divergent. The evidence for stasis of the core lineages, excluding the divergent dietary specialists, lies not only in the lower likelihood of the data under the alternative model of a low Brownian rate ($\Delta$AICc > 100) but also in that model's grossly inaccurate predictions of disparity. Although the predicted disparities under that model are nearly three times the observed values, the discrepancy of the estimates under a single-peak OU model is less than 10% of the observed value. More surprising, the observed value for the joint disparity of the two lineages lies just slightly outside the confidence interval for data simulated under a model of a single optimum. Considering that all models of the evolutionary dynamics of phenotypes predict the accumulation of disparity over time, a model that fails to predict disparity accurately cannot be considered adequate. Judged by that criterion, our results strongly argue for stasis of one or two slightly divergent optima for jaw shape.

**Table 9.** Model misspecification rates. The models under which the data are simulated are in the first column; the models to which those data are fit are in the rows. (A) The proportion of times that the data fit to a given model has a ΔAIC of zero; (B) the proportion of times that the data fit to each model fit to a given model has a ΔAIC ≥ 4.0.

**A**

**Callosciurinae**

| Fit | Simulated | | | | | | |
|---|---|---|---|---|---|---|---|
| | BM1 | OU1 | OU2 | OU3 | Eco | OU6 | OU7 |
| BM1 | 0.84 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| OU1 | 0.00 | 0.94 | 0.01 | 0.05 | 0.00 | 0.00 | 0.00 |
| OU2 | 0.00 | 0.05 | 0.93 | 0.00 | 0.00 | 0.00 | 0.00 |
| OU3 | 0.00 | 0.00 | 0.00 | 0.95 | 0.01 | 0.00 | 0.00 |
| Ecomorph | 0.01 | 0.00 | 0.01 | 0.01 | 0.88 | 0.00 | 0.00 |
| OU6 | 0.00 | 0.00 | 0.00 | 0.00 | 0.10 | 0.99 | 0.28 |
| OU7 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.72 |

**Sciurini**

| Fit | Simulated | | | | |
|---|---|---|---|---|---|
| | BM | OU1 | OU2.Rheithro | OU2.Mini | OU3 |
| BM1 | 0.95 | 0.01 | 0.00 | 0.00 | 0.00 |
| OU1 | 0.04 | 0.95 | 0.00 | 0.01 | 0.00 |
| OU2.Rheithro | 0.00 | 0.02 | 0.69 | NA | 0.05 |
| OU2.Mini | 0.01 | 0.02 | NA | 0.95 | 0.10 |
| OU3 | 0.00 | 0.00 | 0.31 | 0.05 | 0.84 |

**B**

**Callosciurinae**

| Fit | Simulated | | | | | | |
|---|---|---|---|---|---|---|---|
| | BM1 | OU1 | OU2 | OU3 | Eco | OU6 | OU7 |
| BM1 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| OU1 | 0.00 | 1.00 | 0.02 | 0.03 | 0.00 | 0.00 | 0.00 |
| OU2 | 0.00 | 0.00 | 0.98 | 0.00 | 0.00 | 0.00 | 0.00 |
| OU3 | 0.00 | 0.00 | 0.00 | 0.96 | 0.01 | 0.00 | 0.00 |
| eco | 0.00 | 0.00 | 0.00 | 0.00 | 0.91 | 0.00 | 0.00 |
| OU6 | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | 1.00 | 0.27 |
| OU7 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.73 |

**Sciurini**

| Fit (Unequiv) | Simulated | | | | |
|---|---|---|---|---|---|
| | BM | OU1 | OU2.Rheithro | OU2.Mini | OU3 |
| BM1 | 0.96 | 0.01 | 0.00 | 0.00 | 0.00 |
| OU1 | 0.03 | 0.98 | 0.01 | 0.00 | 0.00 |
| OU2.Rheithro | 0.00 | 0.01 | 0.70 | NA | 0.05 |
| OU2.Mini | 0.01 | 0.01 | NA | 0.98 | |
| OU3 | 0.00 | 0.00 | 0.30 | 0.02 | 0.95 |
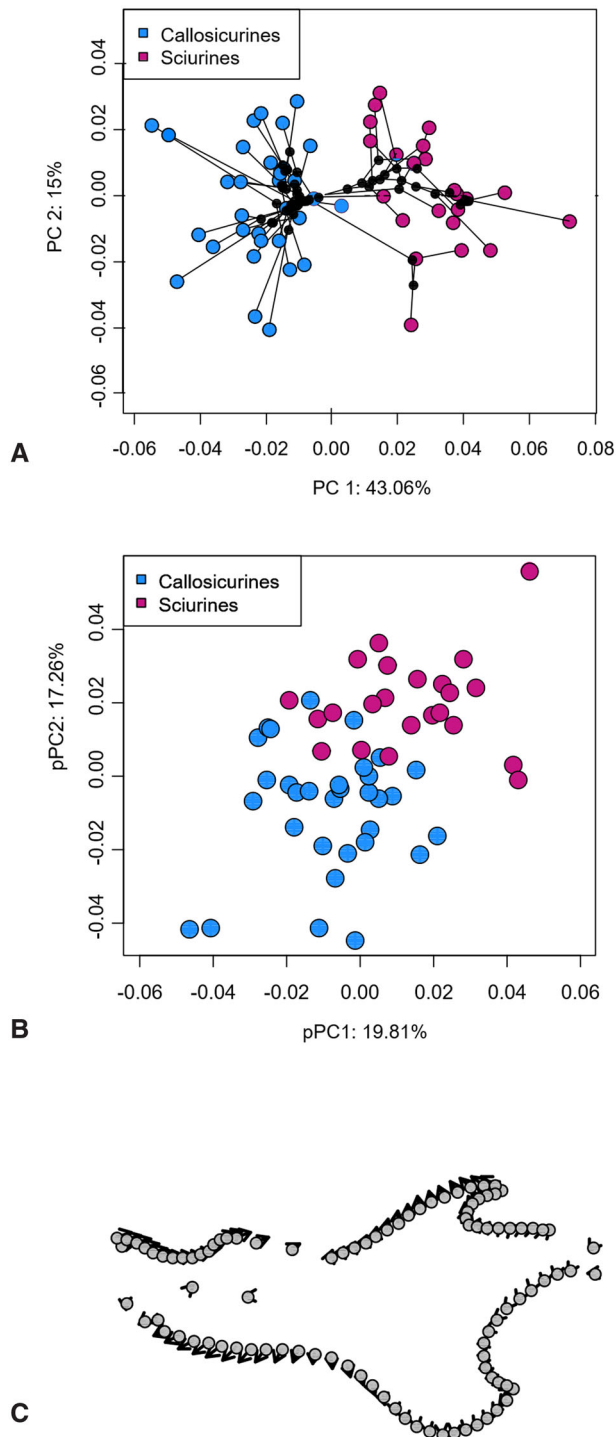
**Figure 10.** The distribution of jaw shapes of the two core lineages within the plane of (A) the first two principal components (PCs), and (B) the first two phylogenetic principal components (pPCs) (extracted given a tree rescaled by the estimated value for $\alpha$ on the one-peak tree) and (C) the difference between the mean shapes of the two lineages.

Our conclusions presuppose that our results are not an artifact of our methods. The most important methodological limitation is that we had to reduce the dimensionality of the data and the number of parameters of the models. We included as many dimensions as possible, given the diversity of each lineage, but that falls short of including all the shape information and relies on a subjective, arguably arbitrary criterion (Monteiro 2013). However, no objective criterion determines how many dimensions are necessary for accurately modeling the evolution of shape. Instead, the most widely used criteria determine how many can be individually interpreted as "meaningful" (e.g., Horn 1965). It may seem more justifiable to reduce the number of model parameters, considering that no information was lost by constraining "traits" to be adaptively independent, but the traits in our analysis are simply axes of the coordinate system. We do find that model misspecification rates are often higher than 5%, up to an alarmingly high 31%, but the models most difficult to differentiate make very similar predictions. Although phylogenetic comparative methods based on Brownian motion display acceptable Type I error and statistical power (Adams and Collyer 2018), their results are nonetheless misleading when the data do not fit that model, predicting disparities up to three times higher than observed and obscuring the processes causing lineages to differ in disparity. Pteromyini differs from Callosciurinae and Sciurini in evolutionary mode not rate, and Callosciurinae differs from Sciurini in both the number of, and distance between, adaptive peaks, which are often unique to a single dietary specialist.

Our results support a variant of the classic hypothesis that generalists are most likely to be static because they remain well-adapted to their environment (Simpson 1944, 1953). What we add to that hypothesis is an explanation for the persistence of the optimum. That explanation is needed because, if ecological generalists are well-adapted to a broad spectrum of resources, their optimum could shift according to the range of resources locally available. According to our hypothesis, what limits those shifts is functional constraints on the morphological specialists whose specializations expand niche breadth. Departures from the optimum that lessen their ability to exploit resources for which they are morphologically specialized are not likely to be tolerated, and they may be even less tolerable for ecologically versatile specialists when those departures reduce niche breadth. Morphological generalists might also occupy steep-sided peaks if their optimum is determined by an invariant balance among trade-offs, but in the absence of functional constraints, they are more likely to drift around their broad adaptive plateau without being repeatedly pulled back to the ancestral optimum. Other lineages of rodents have also been singled out for their limited morphological divergence accompanying speciation, such as *Rattus*, *Peromyscus*, *Microtus*, and Oryzomyinae (Rowe et al. 2011) but analyses of phenotypic evolution of two of them,
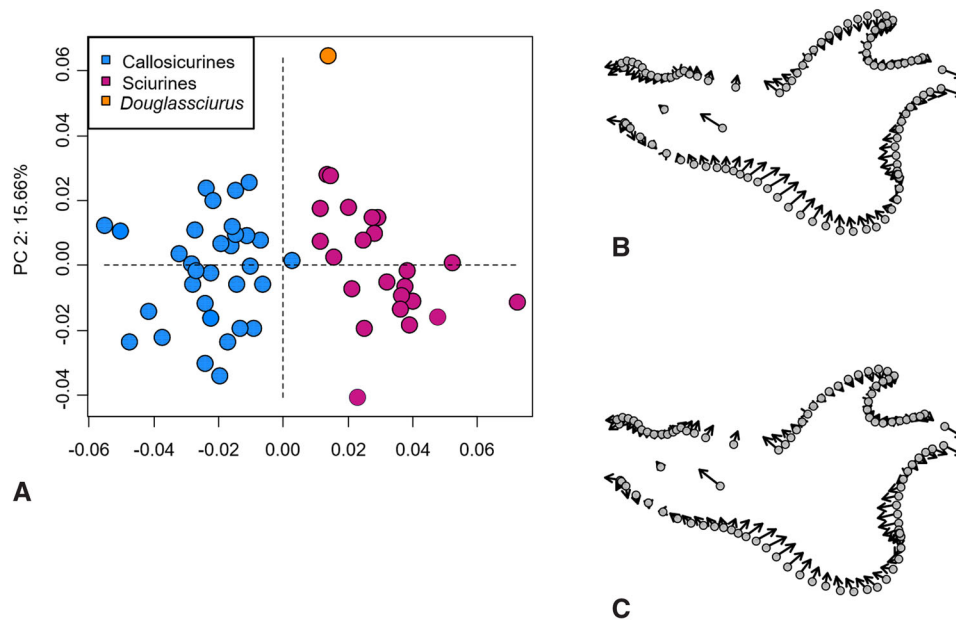
**Figure 11.** The distribution of jaw shapes of the two core lineages plus late Eocene *Douglassciurus jeffersoni* (A); the differences between (B) the optimal shape of the core lineage of Callosciurinae and *D. jeffersoni* and (C) the optimal shape of the core lineage of Sciurini and *D. jeffersoni*.

*Rattus* (Rowe et al. 2011) and Oryzomyalia (Maestri et al. 2017), show that neither is static. For Oryzomyalia, the model that best fit cranial and mandibular shape is a two-rate Brownian model (Maestri et al. 2017), supporting the expectation that, in the absence of functional constraints, morphology may drift around an adaptive plateau. Stochastic peak shifts are a more likely explanation for drifting around an adaptive plateau than is random genetic drift given the tremendous disparity that random genetic drift can produce over macroevolutionary time scales. That tremendous disparity is evident from Lynch's (1990) conclusion that size measurements of a disparate assemblage of North American squirrels, ranging in body size from a small flying squirrel, *Glaucomys volans* (57 g) to the giant *Marmota monax* (2754 g), evolved at just 3% of the minimum neutral expectation.

Of the various hypotheses that ascribe stasis to ecological factors, including environmental stability, biotic interactions, and low speciation rates, none explains why trophic morphology of Callosciurinae and Sciurini is static, unlike that of Pteromyini. Both Callosciurinae and Pteromyini are largely co-distributed in the squirrel-rich Paleotropical forests but Sciurini mainly inhabits squirrel-poor communities of Holarctic-Neotropical forests (Fig. 12). Therefore, explanations that ascribe stasis to the slowly shifting tropical forest belt (Simpson 1944, 1953) or to biotic interactions, either spatial segregation of competitors (Darwin 1859; Lindholm 2014) or persistent competition locking species into their ecological roles (Boucot 1983, 1990; Morris et al. 1995; Kozak et al. 2005), would predict the same evolutionary mode for both Paleotropical lineages and a different one for Sciurini.

Two other explanations are more difficult to rule out: (1) internal stabilizing selection that maintains coherence and functionality of an integrated system (Wagner and Schwenk 2000; Schwenk and Wagner 2001) and (2) intrinsic genetic/developmental constraints. The first is a possibility because the mandible is one component of a functionally integrated trophic apparatus that comprises muscles and teeth as well as bones. The powerful bite of squirrels is partly due to the arrangement of the jaw masseter muscles (sciuromorphy), which is usually singled out as the main component of their specialization (Thorington and Darrow 1996; Druzinsky 2010; Cox et al. 2012). That arrangement arose by displacing the origin of the lateral masseter forward and upward to the zygomatic plate, altering the direction of the anterior lateral masseter and strengthening the forward component of its contraction (Wood 1965). Sciuromorphy, however, is not enough to explain their powerful bites because all squirrels are sciuromorphic but not all have powerful bites; the hard-nut eating tree squirrel, *Sciurus niger*, has an exceptionally powerful bite for its body size but the insectivorous/small seed-eating ground squirrel *Ictidomys tridecemlineatus* does not (Freeman and Lemen 2008). The squirrels that eat the hardest foods typically have a more anterior origin of the anterior deep master, a deeper angular process (increasing mechanical advantage of the superficial masseter), the most robust incisors, and a deep mandibular corpus (Thorington and Darrow 1996). Taken together, morphological specializations for hard-nut feeding implicate adaptations of muscles, bone, and teeth. Because *Douglassciurus jeffersoni* predates the origin of

## Callosciurinae



## Pteromyini

## Sciurini
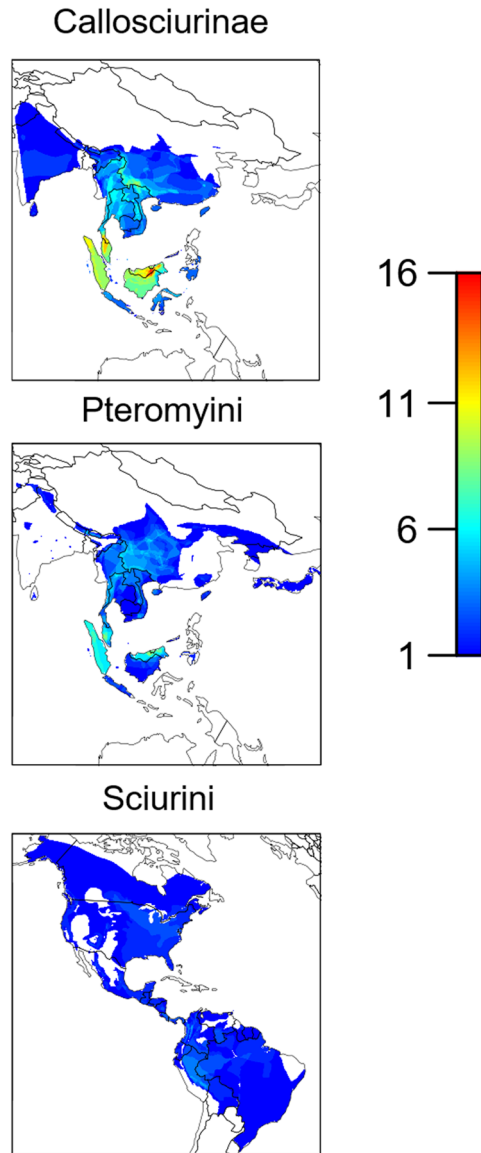
**Figure 12.** Geographic distribution of the diversity of tree squirrels of each lineage based on the International Union for Conservation of Nature (2019) expert range maps at 10-km resolution.

sciuromorphy, it lacks the muscular adaptations that enhance bite-force and that less efficient anatomy might have required larger muscles generating larger forces, hence its more robust jaw (especially the angular process). Once the muscular component of the trophic apparatus evolved, a more slender skeletal structure may have been feasible. Internal stabilizing selection that maintains the functionality of the trophic apparatus may have subsequently contributed to stasis. We cannot rule out the possibility that intrinsic genetic/developmental constraints also play a role in the stasis of tree squirrels; intrinsic constraints can influence macroevolutionary patterns, including the amount, range, and direction of disparity (e.g., Goswami and Polly 2010; Haber 2016; Machado et al. 2018; Rossoni et al. 2019). Even if it

seems unlikely that such constraints can explain both stasis and the diverse directions in which dietary specialists evolved, we cannot rule out that possibility without measuring flexibility and other variational properties.

The specializations of the trophic apparatus, once assembled, enabled tree squirrels to consume the hard nuts present in all environments inhabited by tree squirrels without limiting their ability to consume softer foods. Their specializations thus expand their niche-breadth as well as steepening their adaptive peak. Morphologically specialized, ecological generalists may seem paradoxical, even as paradoxical as stasis. Yet, even though specialized generalists seem paradoxical, their adaptive peaks may be remarkably stable.

### LITERATURE CITED

Adams, D. C. 2014. Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. Syst. Biol. 63:166–167.

Adams, D. C., and M. L. Collyer. 2018. Multivariate phylogenetic comparative methods: evaluations, comparisons, and recommendations. Syst. Biol. 67:14–31.

Adams, D. C., M. L. Collyer, and A. Kaliontzopoulou. 2019. Geomorph: software for geometric morphometric analyses. R package version 3.1.0.

Blows, M. W., and A. A. Hoffmann. 2005. A reassessment of genetic limits to evolutionary change. Ecology 86:1371–1384.

Boettiger, C., G. Coop, and P. Ralph. 2012. Is your phylogeny informative? Measuring the power of comparative methods. Evolution 66:2240–2251.

Bolstad, G. H., T. F. Hansen, C. Pelabon, M. Falahati-Anbaran, R. Perez-Barrales, and W. S. Armbruster. 2014. Genetic constraints predict evolutionary divergence in *Dalechampia* blossoms. Philos. Trans. R. Soc. B 369:20130255.

Bookstein, F. L. 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med. Image Anal. 1:97–118.

Boucot, A. J. 1983. Does evolution take place in an ecological vacuum? II. J. Paleontol. 57:1–30.

———.1990. Evolutionary paleobiology of behaviour and coevolution. Elsevier, Amsterdam, the Netherlands.

Clavel, J., G. Escarguel, and G. Merceron. 2015. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Methods Ecol. Evol. 6:1311–1319.

Collar, D. C., B. C. O'Meara, P. C. Wainwright, and T. J. Near. 2009. Piscivory limits diversification of feeding morphology in centrarchid fishes. Evolution 63:1557–1573.

Cooper, N., G. H. Thomas, C. Venditti, A. Meade, and R. P. Freckleton. 2016. A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. Biol. J. Linn. Soc. 118:64–77.

Cox, P. G., E. J. Rayfield, M. J. Fagan, A. Herrel, T. C. Pataky, and N. Jeffery. 2012. Functional evolution of the feeding system in rodents. PLoS ONE 7:11.

Darwin, C. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, Lond.

de Alencar, L. R. V., M. Martins, G. Burin, and T. B. Quental. 2017. Arboreality constrains morphological evolution but not species diversification in vipers. Proc. R. Soc. B Biol. Sci. 284:20171775.

deVries, M. S., B. C. Stock, J. H. Christy, G. R. Goldsmith, and T. E. Dawson. 2016. Specialized morphology corresponds to a generalist diet: linking form and function in smashing mantis shrimp crustaceans. Oecologia 182:429–442.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7:214.

Druzinsky, R. E. 2010. Functional anatomy of incisal biting in *Aplodontia rufa* and sciuromorph Rodents - Part 2: sciuromorphy is efficacious for production of force at the incisors. Cells Tissues Organs 192:50–63.

Eldredge, N., and J. Cracraft. 1980. Phylogenetic patterns and the evolutionary process: method and theory in comparative biology. Columbia Univ. Press, New York.

Ellerman, J. R. 1949. The families and genera of living rodents. Vol. 3, part 1. Trustees of the British Museum (Natural History). Lond.

Emmons, L., and F. Feer. 1997. Neotropical rainforest mammals, a field guide. Univ. of Chicago Press, Chicago.

Emry, R. J., and R. W. Thorington, Jr. 1982. Descriptive and comparative osteology of the oldest fossil squirrel, *Protosciurus* (Rodentia: Sciuridae). Smithson. Contrib. Paleobiol. 47:1–35.

———. 1984. The tree squirrel *Sciurus* (Sciuridae, Rodentia) as a living fossil. Pp. 23–31 In N. Eldredge and S. M. Stanley, eds. Living fossils. Springer-Verlag, New York.

Emry, R. J., and W. W. Korth. 1996. The Chadronian squirrel ''*Sciurus*'' *jeffersoni* Douglass, 1901: a new generic name, new material, and its bearing on the early evolution of Sciuridae (Rodentia). J. Vertebr. Paleontol. 16:775–780.

Ercoli, M. D., A. Álvarez, and A. M. Candela. 2019. Sciuromorphy outside rodents reveals an ecomorphological convergence between squirrels and extinct South American ungulates. Commun. Biol. 2:202.

Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. Am. Nat. 169:227–244.

Foote, M. 1993. Contributions of individual taxa to overall morphological disparity. Paleobiology 19:403–419.

Freeman, P. W., and C. A. Lemen. 2008. A simple morphological predictor of bite force in rodents. J. Zool. 275:418–422.

Goswami, A., and P. D. Polly. 2010. The Influence of modularity on cranial morphological disparity in Carnivora and Primates (Mammalia). PLoS ONE 5:e9517.

Green, W. D. K. 1996. The thin-plate spline and images with curving features. Pp. 79–87 In K. V. Mardia, C. A. Gill, and I. L. Dryden, eds. Image fusion and shape variability. Univ. of Leeds Press, Leeds, U.K.

Haber, A. 2016. Phenotypic covariation and morphological diversification in the ruminant skull. Am. Nat. 187:576–591.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Hansen, T. F., and D. Houle. 2004. Evolvability, stabilizing selection, and the problem of stasis. Pp. 130–154 In M. Pigliucci and K. Preston, eds. Phenotypic integration: studying the ecology and evolution of complex phenotypes. Oxford Univ. Press, New York.

Hawkins, M. T. R., K. M. Helgen, J. E. Maldonado, L. L. Rockwood, M. T. N. Tsuchiya, and J. A. Leonard. 2016. Phylogeny, biogeography and systematic revision of plain long-nosed squirrels (genus *Dremomys*, *Nannosciurinae*). Mol. Phylogenet. Evol. 94:752–764.

Hine, E., K. McGuigan, and M. W. Blows. 2014. Evolutionary constraints in high-dimensional trait sets. Am. Nat. 184:119–131.

Ho, L. S. T., and C. Ane. 2014. Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. Methods Ecol. Evol. 5:1133–1146.

Horn, J. L. 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30:179–185.

International Union for Conservation of Nature. 2019. IUCN red list of threatened species. available at http://www.iucnredlist.org/ (accessed 15 March, 2018).

Houle, D., G. H. Bolstad, K. van der Linde, and T. F. Hansen. 2017. Mutation predicts 40 million years of fly wing evolution. Nature 548:447–450.

Khabbazian, M., R. Kriebel, K. Rohe, C. Ané, and T. Hansen. 2016. Fast and accurate detection of evolutionary shifts in Ornstein-Uhlenbeck models. Methods Ecol. Evol. 7:811–824.

Kozak, K. H., A. Larson, R. M. Bonett, and L. J. Harmon. 2005. Phylogenetic analysis of ecomorphological divergence, community structure, and diversification rates in dusky salamanders (Plethodontidae: Desmognathus). Evolution 59:2000–2016.

Lanfear, R., B. Calcott, S. Y. Ho, and S. Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29:1695–1701.

Liem, K. F. 1980. Adaptive significance of intraspecific and interspecific differences in the feeding repertoires of cichlid fishes. Am. Zool. 20:295–314.

———.1984. Functional versatility, speciation, and niche overlap: are fishes different? Pp. 269–305 *in* D. G. Meyers and J. R. Strickler, eds. Trophic interactions within aquatic ecosystems. American Association for the Advancement of Science, Washington, D.C.

———.1990. Aquatic versus terrestrial feeding modes: possible impacts on the trophic ecology of vertebrates. Am. Zool. 30:209–221.

Lindholm, M. 2014. Morphologically conservative but physiologically diverse: the mode of stasis in Anostraca (Crustacea: Branchiopoda). Evol. Biol. 41:503–507.

Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. Am. Nat. 136:727–741.

Machado, F. A., T. M. G. Zahn, and G. Marroig. 2018. Evolution of morphological integration in the skull of Carnivora (Mammalia): changes in Canidae lead to increased evolutionary potential of facial traits. Evolution 72:1399–1419.

Maestri, R., L. R. Monteiro, R. Fornel, N. S. Upham, B. D. Patterson, and T. R. de Freitas. 2017. The ecology of a continental evolutionary radiation: is the radiation of sigmodontine rodents adaptive? Evolution 71:610–632.

Mercer, J. M., and V. L. Roth. 2003. The effects of Cenozoic global change on squirrel phylogeny. Science 299:1568–1572.

Meyer, E. R. 1984. Crocodilians as living fossils. Pp. 105–131 *in* N. Eldredge and S. M. Stanley, eds. Living fossils. Springer-Verlag, New York.

Moller, H. 1983. Foods and foraging behavior of red (*Sciurus vulgaris*) and grey (*Sciurus carolinensis*) squirrels. Mammal Rev. 13:81–98.

Monteiro, L. R. 2013. Morphometrics and the comparative method: studying the evolution of biological shape. Hystrix. It. J. Mamm. 24: 25–32.

Morris, P. J., L. C. Ivany, K. M. Schopf, and C. E. Brett. 1995. The challenge of paleoecological stasis: reassessing sources of evolutionary stability. Proc. Natl. Acad. Sci. USA 92:11269–11273.

Musser, G. G., L. A. Durden, M. E. Holden, and J. E. Light. 2010. Systematic review of endemic Sulawesi squirrels (Rodentia, Sciuridae) with descriptions of new species of associated sucking lice (Insecta, Anoplura), and phylogenetic and zoogeographic assessments of sciurid lice. Bull. Am. Mus. Nat. Hist. 339:1–260.

Muul, I., and B. L. Lim. 1978. Comparative morphology, food habits, and ecology of some Malaysian arboreal rodents. Pp. 361–368 In G. G. Montgomery, ed. The ecology of arboreal folivores. Smithsonian Institution Press, Washington, D.C.

Pečnerová, P., and N. Martínková. 2012. Evolutionary history of tree squirrels (Rodentia, Sciurini) based on multilocus phylogeny reconstruction. Zool. Scr. 41:211–219.

Pečnerová, P., J. C. Moravec, and N. Martínková. 2015. A skull might lie: modeling ancestral ranges and diet from genes and shape of tree squirrels. Syst. Biol. 64:1074–1088.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6:7–11.

R Development Core. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabosky, D. L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS ONE 9:e89543.

Rossoni, D. M., B. M. A. Costa, N. P. Giannini, and G. Marroig. 2019. A multiple peak adaptive landscape based on feeding strategies and roosting ecology shaped the evolution of cranial covariance structure and morphological differentiation in phyllostomid bats. Evolution 73:961–981.

Roth, V. L., and J. M. Mercer. 2008. Differing rates of macroevolutionary diversification in arboreal squirrels. Curr. Sci. 95:857–861.

Rowe, K. C., K. P. Aplin, P. R. Baverstock, and C. Moritz. 2011. Recent and rapid speciation with limited morphological disparity in the genus *Rattus*. Syst. Biol. 60:188–203.

Schwenk, K., and G. P. Wagner. 2001. Function and the evolution of phenotypic stability: connecting pattern to process. Am. Zool. 41:552–563.

Sheldon, P. R. 1996. Plus ca change—a model for stasis and evolution in different environments. Palaeogeogr. Palaeoclimatol. Palaeoecol. 127:209–227.

Simpson, G. G. 1944. Tempo and mode of evolution. Columbia Univ. Press, New York.

———.1953. The major features of evolution. Columbia Univ. Press, New York.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Thorington, R., J. L. Koprowski, M. A. Steele, and J. F. Whatton. 2012. Squirrels of the World. Johns Hopkins Univ. Press, Baltimore, MD.

Thorington, R. W., and K. Darrow. 1996. Jaw muscles of old world squirrels. J. Morphol. 230:145–165.

Uyeda, J. C., D. S. Caetano, and M. W. Pennell. 2015. Comparative analysis of principal components can be misleading. Syst. Biol. 64:677–689.

Vrba, E. S. 1987. Ecology in relation to speciation rates: some case histories of Miocene-Recent mammal clades. Evol. Ecol. 1:283–300.

Wagner, G. P., and K. Schwenk. 2000. Evolutionarily stable configurations: functional integration and the evolution of phenotypic stability. Pp. 155–217 *in* M. K. Hecht, R. J. Macintyre, and M. T. Clegg, eds. Evolutionary biology. Springer, Boston, MA.

Wake, D. B., G. Roth, and M. H. Wake. 1983. On the problem of stasis in organismal evolution. J. Theor. Biol. 101:211–224.

Walsh, B., and M. W. Blows. 2009. Abundant genetic variation plus strong selection = Multivariate genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. 40:41–59.

Wood, A. E. 1965. Grades and clades among rodents. Evolution 19:115–130.

Zelditch, M. L., H. D. Sheets, and W. L. Fink. 2003. The ontogenetic dynamics of shape disparity. Paleobiology 29:139–156.

Zelditch, M. L., D. L. Swiderski, and H. D. Sheets. 2012. Geometric morphometrics for biologists: a primer. Academic Press, Lond.

Zelditch, M. L., J. Li, L. A. P. Tran, and D. L. Swiderski. 2015. Relationships of diversity, disparity and their evolutionary rates in squirrels (Sciuridae). Evolution 69:1284–1300.

Zelditch, M. L., J. Ye, J. S. Mitchell, and D. L. Swiderski. 2017. Rare ecomorphological convergence on a complex adaptive landscape: body size and diet mediate evolution of jaw shape in squirrels (Sciuridae). Evolution 71:633–649.

Associate Editor: M. Collyer
Handling Editor: D. W. Hall

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1**.
**Table S2**.
**Table S3**. Relative support for all models fit to mandibular shape.
**Table S4**. Estimates of the four evolutionary rates (s1: s4) for shape of Pteromyini, obtained from the data and simulations.