

Pulsed evolution shaped modern vertebrate body sizes

Michael J. Landis ^{*}, Joshua G. Schraiber ^{† ‡}

^{*}Department of Ecology and Evolutionary Biology, Yale University, [†]Department of Biology, Temple University, and [‡]Institute for Genomics and Evolutionary Medicine, Temple University

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The relative importance of different modes of evolution in shaping phenotypic diversity remains a hotly debated question. Fossil data suggest that stasis may be a common mode of evolution, while modern data suggest some lineages experience very fast rates of evolution. One way to reconcile these observations is to imagine that evolution proceeds in pulses, rather than in increments, on geological time scales. To test this hypothesis, we developed a novel maximum likelihood framework for fitting Lévy processes to comparative morphological data. This class of stochastic processes includes both an incremental and a pulsed component. We found that a plurality of modern vertebrate clades examined are best fit by pulsed processes over models of incremental change, stationarity, and adaptive radiation. When we compare our results to theoretical expectations of the rate and speed of regime shifts for models that detail fitness landscape dynamics, we find that our quantitative results are broadly compatible with both microevolutionary models and with observations from the fossil record.

Macroevolution, adaptive shift, phylogenetic comparative method, Lévy process

Abbreviations:

A key debate in evolutionary biology centers around the seeming contradictions regarding the tempo and mode of evolution as seen in fossil data compared to ecological data. Fossil data often support models of stasis, in which little evolutionary change is seen within lineages over long timescales [1, 2], while ecological data show that rapid bursts of evolution are not only possible, but potentially common [3, 4]. At face value, these observations seem to contradict one another, an observation known as the “paradox of stasis” [5]. These observations are often reconciled through a descriptive model of pulsed evolution, entailing stasis interrupted by pulses of rapid change, as famously articulated by Simpson [6].

On macroevolutionary timescales, pulses of rapid change are expected to look roughly instantaneous. Only recently have statistical methods grown sophisticated enough to model pulsed evolution as a stochastic process, with advances showing that punctuation is detectable in some fossil time series [7] and between pairs of living and extinct taxa [8].

While these studies establish the existence of pulsed evolution, it is still unknown whether the evolutionary mode is common or rare. How many clades in the Tree of Life were shaped by abrupt pulses of rapid evolution? If these evolutionary pulses are common, then that should inform our expectations about how traits evolved for clades that left no fossils, and the potential for vulnerable species to adapt rapidly to climate change [9]. To this end, phylogenetic models—models of trait evolution that account for the shared ancestry of species—have played a vital role in measuring the relative support of competing Simpsonian modes of evolution.

A pioneering meta-analysis [10] fitted a collection of phylogenetic models to 49 animal clades, finding preference for modes of incremental, but not explosive, evolutionary change. Their work predated the advent of phylogenetic models of pulsed evolution [11, 12], so its frequency could not have been measured. Moreover, there remains a concern that models of

incremental and pulsed change leave similar patterns of trait variation in neontological data [13, 14]. While recent methodological developments show that there is power in comparative data to identify pulsed evolution [15, 11, 16], little is known about either the statistical properties of these methods, nor is much known about the prevalence of pulsed change throughout some of Earth’s most intensely studied clades.

Here, we examine evidence for pulsed evolution across vertebrate taxa using a method for fitting Lévy processes to comparative data. These processes can capture both incremental and pulsed modes of evolution in a single, simple framework. We apply this method to analyze 66 vertebrate clades containing 8,323 extant species for evidence of pulsed evolution by comparing the statistical fit of several varieties of Lévy jump processes (modeling different types of pulsed evolution) to three models that emphasize alternative macroevolutionary dynamics. Under these models, the adaptive optimum of a lineage may wander incrementally and freely (Brownian motion), it may change incrementally but remain stationary (Ornstein-Uhlenbeck), or it may change most rapidly following the initial diversification of a clade while decelerating over time, e.g. during an adaptive radiation (Early Burst). Beyond simple model comparison, we show that the parameter estimates corresponding to the microevolutionary and macroevolutionary mechanisms of the model have biologically meaningful interpretations [2], illuminating previously hidden features of Simpson’s adaptive grid.

Significance

The diversity of forms found among animals on Earth is striking. Despite decades of study, it has been difficult to reconcile the patterns of diversity seen between closely related species with those observed when studying single species on ecological timescales. We propose a set of models, called Lévy processes, to attempt to reconcile rapid evolution between species with the relatively stable distributions of phenotypes seen within species. These models, which have been successfully used to model stock market data, allow for long periods of stasis followed by bursts of rapid change. We find that many vertebrate groups are well-fitted by Lévy models when compared to models for which traits evolve towards a stationary optimum, or evolve in an incremental and wandering manner.

Reserved for Publication Footnotes

Results

Maximum likelihood method has power to distinguish pulsed evolution from comparative data. We developed a maximum likelihood method for fitting Lévy processes to phylogenetic comparative data using restricted maximum likelihood estimation (REML), by analyzing the phylogenetically independent contrasts [17] (*Materials and Methods*). The Lévy processes we apply in this work consist of two components: a Brownian motion and a pure jump process. The Brownian motion is characterized by a rate parameter, σ^2 , and the pure jump process is characterized by a Lévy measure, $\nu(\cdot)$, where $\nu(dx)dt$ can be thought of as the probability of a jump with a size in the interval $(x, x + dx)$ occurring in the short time dt . Note

that this model does not couple pulses of evolution to cladogenesis, as in the classical theory of punctuated equilibrium [18]. Instead, pulses may occur at any time, sometimes known as “punctuated anagenesis” [19].

Both Lévy processes with jumps and pure Brownian motion accumulate variance proportionally to time (*SI Appendix*), leading to speculation that it is impossible to distinguish between pulsed and certain incremental models from comparative data [13, 14]. For simulations with moderately sized clades (>100 taxa), we had sufficient power to differentiate pulsed evolution from other Simpsonian modes of evolution. This is due to the impact of rare, large jumps resulting in a heavy-tailed distribution of trait change (*SI Appendix*). Moreover, we saw low false positive rates for identifying pulsed evolution, even in the presence of phylogenetic error (4% for clades with ~ 100 taxa, 7% for clades with ~ 300 taxa; *SI Appendix*).

Extant vertebrate body sizes evolved by rapid bursts. We assembled comparative datasets from time-scaled tree estimates and body size measurements for 66 clades across five major vertebrate groups (Table 1). We computed the sample size-corrected Akaike Information Criteria weights (wAICc) [20, 21] for each dataset from a panel of models. We used a Brownian motion (BM) to model incremental phenotypic change, in which a lineage’s phenotype follows a wandering optimum. We used the Ornstein-Uhlenbeck (OU) model to incremental evolution around a single stationary optimum and the Early Burst (EB) model to capture adaptive radiation, in which the tempo of evolution slows over time. We also compared two different types of jump processes: the compound Poisson with normally distributed jumps (JN or Jump Normal) and the Normal Inverse Gaussian (NIG). The JN process waits for an exponentially distributed amount of time before jumping to a new value; we use this to represent stasis followed by large-scale shifts between adaptive zones. On the other hand, the NIG process is constantly jumping, with larger jumps requiring longer waiting times; this model captures the dynamics of constant phenotypic change, where the majority of change occurs within the width of an adaptive zone, but rare shifts between adaptive zones are still possible. Finally, we examined the combination of the Brownian and jump processes, BM+JN and BM+NIG. We further modeled intraspecific variation and sampling noise using a normal distribution. Examples of how each different Lévy process models trait evolution can be found in the *SI Appendix*.

Figure 1 shows that that 32% of clades were best fit by models of pulsed evolution. Brownian motion, Early Burst, and Ornstein-Uhlenbeck were each selected substantially less often (18%, 14%, and 2%, respectively). About 34% of clades lacked decisive support for any single mode of trait evolution. We did not find any evidence that a particular evolutionary mode is enriched in any of the five vertebrate groups (χ^2 test, $p \approx 0.10$). Lévy jump models and Brownian models fitted to any given clade yield nearly identical estimates of process variance (*SI Appendix*), indicating that the tempo and mode of evolution leave distinct patterns in comparative data.

We speculated that clades would favor models that produced similar patterns in comparative data. To characterize this, we applied a principal components analysis to the vectors of wAICc scores of Figure 1, then clustered wAICc profiles using k-means ($k=4$). The first three principal components explain 85.6% of the variance. Figure 2 shows that four clusters form around clades that select models of incremental evolution (Brownian motion and Ornstein-Uhlenbeck), explosive evolution (Early Burst), and the two models representing pulsed evolution (Jump Normal and Normal Inverse Gaussian). Interestingly, the third component separates clades that favored



Fig. 1. Model selection profiles for 66 vertebrate clades. Clade colors indicate their order: black, fish; purple, amphibians; green, reptiles; blue, birds; red, mammals. Each clade was fit to seven models, classified into four groups: incremental change (BM), incremental stationarity (OU), explosive change (EB), and pulsed change (JN, NIG, BM+JN, BM+NIG). AICc weights were computed using only the best fitting model within each class. A model class is selected only if its AICc weight is twice as large than that of any other model class (circles indicate selection counts: 12 incremental change, 1 incremental stationarity, 9 explosive change, 21 pulsed change, 23 ambiguous). Alternative model classifications are provided in *SI Appendix*.

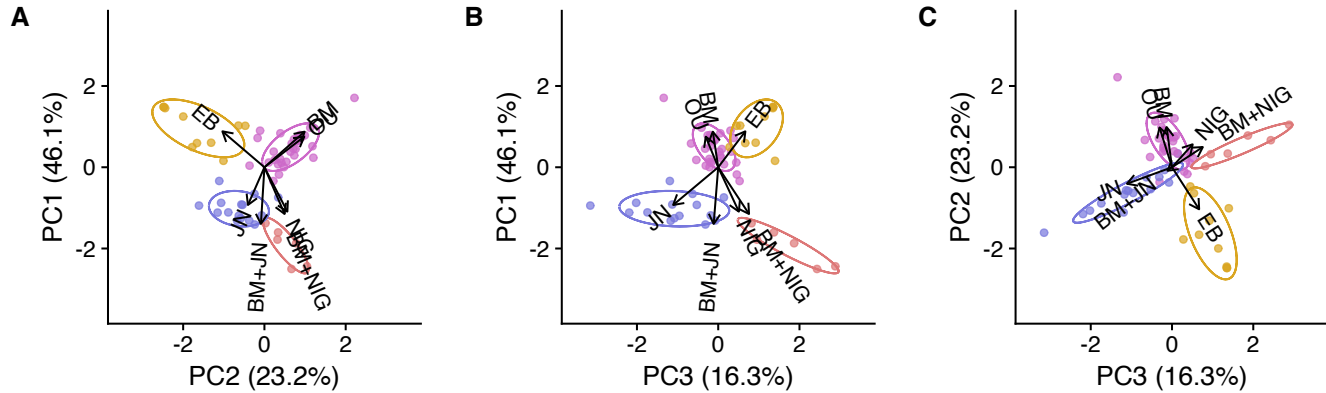


Fig. 2. Empirical axes of trait evolution. The principal components and k-means clusters ($k = 4$) for wAICc scores reported in Figure 1. Panels A, B, and C plot the first three principal component axes as pairs, explaining 85.6% of the model-fit variance. Both PCA and the k-means clusters identify three major axes of trait evolution: incremental change and incremental stationarity (purple), explosive (gold), and pulsed models, with the pulsed axis being divisible into subtypes of jump processes with finite (blue) and infinite (red) activity.

the two different kinds of jump models we explored, with the most pulsed process (Jump Normal) running opposite to the infinitely active processes (Normal Inverse Gaussian). This suggests that with more power it will be possible to choose among different models of pulsed evolution with greater confidence.

Waiting times to transition between adaptive zones. Under a microevolutionary model of stabilizing selection, intraspecific variation is distributed normally around the adaptive optimum [22]. This intraspecific variation in turn approximates the width of a given adaptive zone [6, 23]. When fitting models of trait evolution, we inferred a parameter that corresponds to the standard deviation of intraspecific variation. Because we have only one sample per species, this is a noisy measure of intraspecific variation that also includes measurement error. We therefore regressed our estimates of the intraspecific variation using bird data with known phenotypic standard deviations, and used the regressed estimates as a measure of intraspecific variation, σ_{intra} (*Materials and Methods, SI Appendix*).

We sought to estimate the rate at which lineages escaped their current adaptive zones via shifts. To do so, we computed the expected waiting time, $\bar{t}(z\sigma_{\text{intra}})$, until the lineage jumps at least z intraspecific standard deviations away from the current phenotype (*Materials and Methods*). Figure 3A shows these waiting times for the pure Jump Normal process. The value at $z = 0$ shows the waiting time for a jump of any size to occur, while we take $z \geq 2$ to be a jump outside of the current adaptive zone, which we define as the interval containing roughly 95% of intraspecific variation per lineage. We found that lineages waited between 1 and 100 million years before shifting to new adaptive zones ($z \geq 2$), with a median waiting time of 2.5×10^7 years.

The waiting time between adaptive shifts depends both on the magnitude of the shift and on the underlying jump model (Figure 3B). Models with jumps differ in two ways: 1) the relative frequencies of jumps within and between adaptive zones (e.g. $z < 2$ vs $z \geq 2$); and 2) whether the fitted model is a pure jump process or a jump-diffusion process. JN models experience long periods of stasis, and jumps tend to be large. Thus, small changes within a lineage’s adaptive zone occur on approximately the same time scale as jumps between adaptive zones. In contrast, infinitely active processes, as represented by NIG, jump continuously. Under these models,

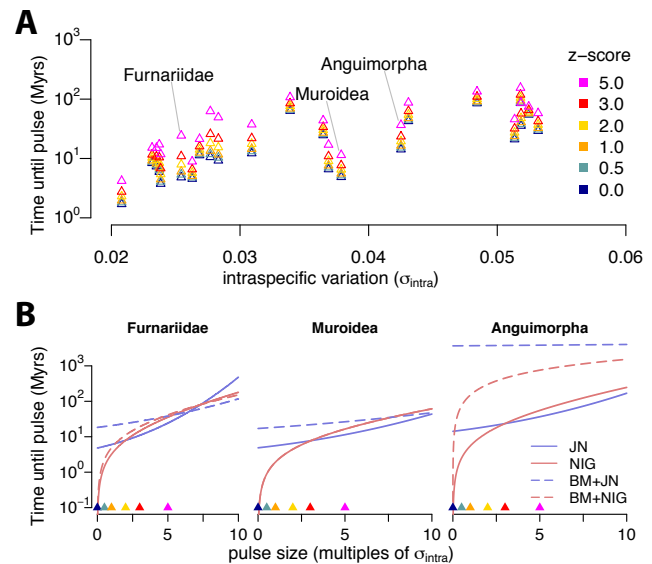


Fig. 3. Expected waiting time to shift between adaptive zones. A) Plot shows the expected waiting of time needed to produce an evolutionary pulse that is at least z times that of the standard deviation of the intraspecific variance (σ_{intra}). Times shown for the pure jump-normal process for those clades that select jump processes and $\sigma_{\text{intra}} > 0$ by any amount. B) Waiting times until pulses at least as large as given size differs between the four fitted jump models for three clades shown in (A). Jump-normal (blue) has finite activity, and cannot easily produce small amounts of change quickly, unlike the normal-inverse Gaussian (red), which is infinitely active. AIC selects pure jump models (solid lines) for *Furnariidae* (left) and *Muroidea* (center), and a jump-diffusion model (dotted lines) for *Anguimorpha* (right).

shifts within an adaptive zone occur frequently, which could reflect rapid adaptation to small-scale environmental perturbations, while larger shifts between adaptive zones occur on much longer time scales. Interestingly, we find that both finite and infinite activity pure jump models predict similar rates of shifting between adaptive zones, as indicated by the fact that the solid curves are near each other for $z \geq 2$. When looking at a clade that is best fit by a jump-diffusion model, such as *Anguimorpha* (Figure 3B, right), we find that they explain most of the morphological disparity among taxa via the Brownian motion component of the model, and we see that jumps occur extremely rarely compared to pure jump models.

Jump-diffusion models with small diffusion components produce jump size-time curves that are essentially identical to the jump size-time curves for the corresponding pure jump model (compare NIG and BM+NIG for *Muroidea*).

Discussion

The pulses of rapid evolution proposed by Simpson [6] are evident in ecological [3, 4] and paleontological [24] observations. Because of their rapidity relative to geological timescales, such phenotypic “jumps” are difficult to observe directly [25, 26], but their lasting impressions are detectable in both simulated and empirical comparative phenotypic datasets (Figure 1). Among Simpsonian modes of phenotypic macroevolution—including unconstrained incremental evolution, evolutionary stationarity, explosive evolution, and pulsed evolution—we found that pulsed evolution is not only detectable, but is a preferred explanation for how body sizes evolved among diverse vertebrate groups. Phenotypic variation accrues in time at a rate that is nearly identical when assuming models of non-stationary incremental evolution versus pulsed evolution; that is, they differ in evolutionary mode rather than evolutionary tempo.

To test for the signatures of different modes of evolution, we assembled neontological comparative datasets for 66 vertebrate clades, composed of time-calibrated phylogenies and body size measurements (Table 1). Importantly, we undertook the call for “broad comprehensive sampling” that is necessary to characterize tempo and mode of evolution without bias [27]. We developed a method to fit a model of phylogenetic evolution of continuous traits for processes that model macroevolutionary jumps by exploiting the mathematical properties of Lévy processes. Applying our method to the vertebrate dataset, we found that body size evolution was best explained by jump models in 32% of vertebrate clades when fitting incremental, stationary, and explosive models of evolution along with four flavors of jump processes. Taking a broader definition of support for models that concentrate, rather than evenly distribute, evolutionary change, pulsed and Early Burst models are selected for 45% of all clades, and for 70% among those clades with strong model preference of any kind. This suggests future work to explore the interaction of Early Burst and pulsed models, in which the rate and/or magnitude of evolutionary pulses decays following the initial radiation of a clade.

Simulated datasets with phylogenetic error do not result in a systematic bias towards preferring Lévy processes over competing models (Figure S2). We found that, if anything, our data filtering and model selection procedures tip the balance to incorrectly favor non-pulsed over pulsed models of evolution. The false positive rate seen in simulations are four to seven times lower than the rate of clades that are best fit by Lévy processes in the empirical data, suggesting a very low false discovery rate. These findings contradicted concerns that incremental and pulsed models would be difficult to distinguish, because Brownian motion models and jump processes result in the same tempo of evolution (as measured by the variance that the processes accumulate per unit time). Nonetheless, rare, large jumps result in heavy-tailed distributions of trait change along branches (as quantified by the excess kurtosis of the trait change distribution), which can be distinguished from incremental evolution occurring at a similar tempo. In the empirical dataset, we found no support that certain model classes were overrepresented within a particular vertebrate group (Table S2). One possibility is that the five vertebrate groups ex-

plore phenotypic space using the same palette of evolutionary modes in roughly equal proportion.

We did not replicate findings of weak support for Early Burst [28]. Several studies have reported that macroevolutionary models of continuous trait evolution may produce misleading results when intraspecific variation is ignored [29, 30, 31, 32]. Thus, we suspected that Early Burst models that ignore intraspecific variation cannot simultaneously explain explosive variation between deeply diverged species in the clade alongside variation concentrated near the tips of the tree. When we forced the model to ignore intraspecific variation, we found that Ornstein-Uhlenbeck to be better supported than Early Burst or Brownian motion (Table S3). This suggests that many findings of support for models of evolutionary stationarity over models of adaptive radiations may be due to model misspecification (either by ignoring intraspecific variation or assuming a fixed value), rather than evidence for adaptive landscapes with single stationary peaks. We therefore suggest that modeling intraspecific variation is an essential component of analyses of comparative data, although more work is needed to understand the complicated interplay when modeling both intraspecific and evolutionary variation simultaneously [32, 33].

One possible microevolutionary explanation for pulses of rapid evolution may relate to Wright’s shifting balance theory, in which small populations stochastically shift between adaptive zones [34]. We estimate that the median waiting time to jump to new adaptive zones is on the order of 10^7 years across clades, consistent with predictions for the time required to jump between adaptive peaks that are separated by fitness valleys via genetic drift in small populations [25]. Under this model, the primary impediment to peak shifts is escaping the current adaptive zone; once escaped, transitions to new adaptive optima occur rapidly. With biologically reasonable parameters, shifts between adaptive zones are expected to occur every $\sim 10^6 - 10^7$ generations, but only take $\sim 10^2 - 10^3$ generations to complete once initiated. Alternatively, this waiting time may represent the timescale of changes of the adaptive landscape itself. Under this model, macroevolutionary jumps are precipitated by changing biotic or abiotic conditions (such as climate change). Quantitative genetic theory shows that adaptation to new, nearby optima can occur extremely rapidly, again on the order of $\sim 10^2 - 10^3$ generations [26].

Our results may also be compatible the classic theory of punctuated equilibrium, in which pulses of evolution are coincident with cladogenesis [18]. In this scenario, apparent “punctuated anagenesis” in comparative data is due to speciation events corresponding to now-extinct lineages. While recent work showed that jointly modeling lineage dynamics and trait evolution has some power to detect classical punctuated equilibrium from comparative data [35], little is known about the power to distinguish strict punctuated equilibrium from more generic pulsed evolutionary models. If “hidden speciations” occur at an approximately constant rate across a phylogeny, it is likely that it would be difficult to disentangle these two models from neontological data alone. Instead, inhomogeneous speciation dynamics, such as logistic diversification [36], may provide a sufficiently strong signal to disentangle anagenetic and cladogenetic models of pulsed evolution.

We found that different kinds of jump processes, representing different modes of rapid evolution (constant rapid adaptation vs. long periods of stasis broken up by jumps between adaptive zones), leave faint, but unique, signatures in phylogenetic data. By integrating these models into fossil sequences, we suspect that further fine-scale details of macroevolution can be elucidated. Moreover, in quantitative finance, where the

Table 1. Vertebrate datasets. Data were collected for $N = 66$ clades for 8323 species records in total (mean 126 species per clade). The “Phylogeny” column indicates how phylogenies within each vertebrate group were generally acquired: by pruning from a large tree or by collecting independently estimated trees from the literature.

	Clade sizes			Phylogeny	Body size
	# clades	mean	(min, max)		
Fish	12	105	(37, 251)	Collected [28, 47, 48, 49, 50, 51, 52, 53, 54, 55]	Length from [56]
Amphibians	4	176	(103, 231)	Pruned from [57, 41]	Length from [58]
Reptiles	24	127	(53, 220)	Pruned from [59, 60]	Length from [61]
Birds	17	111	(56, 260)	Pruned from [43]	Mass from [62]
Mammals	9	158	(69, 229)	Collected [63, 64, 65, 66, 67, 68, 69]	Mass from [70]
Total	66	126	(37, 260)	—	—

“fossil record” of stock prices through time is perfectly kept, fine-scale dynamics of jump processes can be inferred [37, 38], suggesting that such power exists for suitably densely sampled fossil sequences. Our approach, which uses only modern data but integrates it into a phylogenetic framework, represents an important step towards a fully integrative analysis of macroevolutionary processes.

Materials and Methods

Likelihoods using characteristic functions. Because most Lévy processes are only known by their characteristic function, we developed a REML algorithm that operates on characteristic functions. We computed the likelihood of the independent contrasts by proceeding from the tips to the root of the phylogeny. To do so, we recursively update the estimate of the trait at internal nodes as a linear combination of the trait value at the two daughter nodes. We also include an additional term to model the uncertainty in trait values at the tips of the tree due to both intraspecific variation and measurement error. Details of the algorithm can be found in the *SI Appendix*.

Optimizing model fit using REML. To fit the model to data, we used the R programming language, and relied on several R packages [39, 40, 41]. Our package is available at <https://github.com/Schraiber/pulsR>. Optimal parameter values are obtained by maximizing the likelihood using the Nelder-Mead simplex method [42]. To ensure that we obtained true maximum likelihood estimates, we performed multiple independent parameter searches (10 for empirical data, 4 for simulated data). In addition, we validated that all maximum likelihood estimates agreed with model hierarchies, e.g. the maximum likelihood for Ornstein-Uhlenbeck must be greater than or equal to that for Brownian motion.

Simulation study for model selection power and sensitivity. We assessed the power and false positive rate of our method by manipulating tree size and phylogenetic error in a controlled simulation setting. Each dataset was simulated by sampling one tree from an empirical posterior distribution of trees [43] then simulating trait data under each of the seven candidate models (BM, OU, EB, JN, NIG, BM+JN, BM+NIG). Simulation parameters are given in Table S1. We then computed AIC weights for each of the seven datasets across the same seven candidate models, once under the originally sampled tree (the “true” tree), then again under the maximum clade credibility tree that summarizes the posterior (the “consensus” tree). Repeating this procedure 100 times, we counted how often each model type is selected in the presence and absence of tree error. This experiment was conducted for two clades with sizes representative of our empirical datasets (Procellariidae with 106 taxa; Tityridae, Tyrannidae, and allies with 325 taxa). In total, we simulated 1,400 datasets and performed 39,200 maximum likelihood fittings.

Vertebrate dataset. We compiled numerous phylogenetic and trait measurement resources to measure body size evolution (Table 1). The species relationships within each clade were represented using fixed phylogenies with branch lengths measured in millions of years. Body size data were represented by body length for fish, amphibians, and reptiles; and, mass being a function of volume, the cube root of body mass for mammals and birds. Trait measurements were then log transformed, under the assumption that trait evolution operates on a multiplicative scale [44]. Handling the data in this manner permits the comparison of evolutionary parameters between clades because our body size and tree data share a common scale. Broadly, we applied the same data assembly strategy to all clades within each of the five vertebrate groups. Clades were generally either extracted from large, taxon-rich phylogenies (amphibians, reptiles, birds), or pooled across independently estimated phylogenies (fish, mammals). Availability and quality of trait data varied between vertebrate group, each requiring special handling. **Fish:** Standard, fork, and total lengths were extracted from Fishbase

using RFishbase [45]. For each fish clade, we used an allometric regression to convert all taxa to the most frequently observed length type [46]. **Amphibians:** Mean female snout-vent lengths were used for the three Anuran clades (Ranoidea, Hyloidea I, and Hyloidea II). Caudata traits were drawn from [41]. **Reptiles:** Total lengths were used for all snake (Serpentes) clades. Maximum straight-line carapace lengths were used for turtles (Chelonia). Snout-vent lengths were used for all remaining reptile clades. **Mammals:** Adult unsexed body mass measurements were used. **Birds:** Adult sex-averaged mean body mass measurements were used. **Data treatment:** We corrected for two sources of measurement error that would potentially mislead model selection tests to favor Lévy processes. First, rounding error in the trait data caused some clades to contain contrasts that have the exact value of zero. Zero-valued contrasts artificially favor the JN process, which contains a singularity at 0. Second, grossly inaccurate measurements and/or phylogenetic error will drive some portion of contrasts to appear excessively large, fattening the tail density of the trait distribution. To mitigate such errors, we first pruned away any subclades with zero-valued contrasts for all clades. Then, assuming a Brownian motion model, we pruned away the subclade with the most extreme valued contrast for 63 of 66 clades. The three clades exempted from the second filtering step had either a large, but expected, contrast (Anguimorpha, involving the contrast containing *Varanus komodoensis*) or had the largest contrast fall near the base of the tree (Procellariidae and Thamnophtilidae). The filtering procedures resulted in a 4.7% loss of data (8,323 of 8,729 taxa remained). *SI Appendix* contains further details regarding how the input data were processed.

Empirical model selection. We computed sample size corrected AIC weights (wAICc) for each clade across four model classes: BM, OU, EB, and Lévy processes (Figure 1). The Lévy process with the highest AICc score was chosen to represent all Lévy processes within the class. Model selection required the model class to be at least twice as probable as any competing model class. With Brownian motion being a special case of the remaining six models, this threshold lies on the cusp of the maximum relative probability that it might receive under wAICc. Numerous alternative analyses under various model classifications and assumptions are located in *SI Appendix*.

Expected waiting time between adaptive shifts. For Lévy processes with a jump component, we are interested in the waiting time until a jump outside of the current adaptive zone occurs. Under the symmetric jump models we consider here, the waiting time for a jump larger than x is

$$\bar{t}(x) = \frac{1}{2 \int_x^\infty \nu(ds)}.$$

We characterized the width of adaptive zones by the amount of intraspecific variation in a clade. To provide an accurate estimate of intraspecific variation for all clades, we regressed our observed values of σ_{tip} against direct measurements of intraspecific standard deviation in birds [62] ($n = 16$, $p < 0.005$, $R^2 = 0.53$, slope = 0.200, intercept = 0.059). Further details for this analysis and results using the uncorrected σ_{tip} estimates are given in *SI Appendix*.

ACKNOWLEDGMENTS. We would like to thank Damien Wilburn, Tracy Heath, and Ignacio Quintero for helpful comments on an early version of this manuscript. Jonathan Chang, Peter Cowman, and Nathan Upham provided invaluable advice on constructing the empirical datasets. Discussions with Joe Felsenstein helped direct the early course of this project. The comments of two anonymous reviewers greatly improved the clarity and readability of this manuscript. JGS was supported in the early stages of this work by NSF Postdoctoral Fellowship (DBI-1402120) and subsequently by startup funds from Temple University. MJL was supported initially by the Donnelley Fellowship through the Yale Institute of Biospheric Studies and later through the NSF Postdoctoral Fellowship (DBI-1612153). Analyses were performed on the Yale High Performance Computing clusters.

1. Hunt G (2007) The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proceedings of the National Academy of Sciences* 104(47):18404–18408.
2. Estes S, Arnold SJ (2007) Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. *The American Naturalist* 169(2):227–244.
3. Thompson JN (1998) Rapid evolution as an ecological process. *Trends in Ecology & Evolution* 13(8):329–332.
4. Carroll SP, Hendry AP, Reznick DN, Fox CW (2007) Evolution on ecological time-scales. *Functional Ecology* 21(3):387–393.
5. Hansen TF, Houle D (2004) Evolvability, stabilizing selection, and the problem of stasis in Phenotypic Integration: Studying the Ecology and Evolution of Complex Phenotypes, eds. Pigliucci M, Preston K. (Oxford University Press, Oxford), pp. 130–150.
6. Simpson GG (1944) *Tempo and Mode in Evolution*. (Columbia University Press, New York).
7. Hunt G, Hopkins MJ, Lidgard S (2015) Simple versus complex models of trait evolution and stasis as a response to environmental change. *Proceedings of the National Academy of Sciences* 112(16):4885–4890.
8. Uyeda JC, Hansen TF, Arnold SJ, Pienaar J (2011) The million-year wait for macroevolutionary bursts. *Proceedings of the National Academy of Sciences* 108(38):15908–15913.
9. Quintero I, Wiens JJ (2013) Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecology Letters* 16(8):1095–1103.
10. Harmon LJ et al. (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64(8):2385–2396.
11. Landis MJ, Schraiber JG, Liang M (2013) Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. *Systematic Biology* 62(2):193–204.
12. Duchon P et al. (2017) Inference of evolutionary jumps in large phylogenies using Lévy processes. *Systematic Biology* (Advance Access).
13. Hansen TF, Martins EP (1996) Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* pp. 1404–1417.
14. Clavel J, Morlon H (2017) Accelerated body size evolution during cold climatic periods in the Cenozoic. *Proceedings of the National Academy of Sciences* 114(16):4183–4188.
15. Bokma F (2002) Detection of punctuated equilibrium from molecular phylogenies. *Journal of Evolutionary Biology* 15(6):1048–1056.
16. Duchon P et al. (2017) Inference of evolutionary jumps in large phylogenies using Lévy processes. *Systematic Biology* p. syx028.
17. Felsenstein J (2004) *Inferring phylogenies*. (Sinauer Associates Sunderland).
18. Gould SJ, Eldredge N (1977) Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology* 3(2):115–151.
19. MacLeod N (1991) Punctuated anagenesis and the importance of stratigraphy to paleobiology. *Paleobiology* 17(2):167–188.
20. Akaike H (1981) Likelihood of a model and information criteria. *Journal of Econometrics* 16(1):3–14.
21. Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. *Biometrika* 76(2):297–307.
22. Lande R (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* pp. 314–334.
23. Van Valen L (1973) A new evolutionary law. *Evolutionary Theory* 1:1–30.
24. Hunt G, Bell MA, Travis MP (2008) Evolution toward a new adaptive optimum: phenotypic evolution in a fossil stickleback lineage. *Evolution* 62(3):700–710.
25. Lande R (1985) Expected time for random genetic drift of a population between stable phenotypic states. *Proceedings of the National Academy of Sciences* 82(22):7641–7645.
26. Kirkpatrick M (1982) Quantum evolution and punctuated equilibria in continuous genetic characters. *The American Naturalist* 119(6):833–848.
27. Stanley SM (1979) *Macroevolution: Pattern and Process*. (WH Freeman San Francisco).
28. Harmon LJ et al. (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64(8):2385–2396.
29. Cooper N, Thomas GH, Venditti C, Meade A, Freckleton RP (2015) A cautionary note on the use of Ornstein-Uhlenbeck models in macroevolutionary studies. *Biological Journal of the Linnean Society*.
30. Pennell MW, FitzJohn RG, Cornwell WK, Harmon LJ (2015) Model adequacy and the macroevolution of angiosperm functional traits. *The American Naturalist* 186(2):E33–E50.
31. Silvestro D, Kostikova A, Litsios G, Pearman PB, Salamin N (2015) Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution* 6(3):340–346.
32. Kostikova A, Silvestro D, Pearman PB, Salamin N (2016) Bridging inter- and intraspecific trait evolution with a hierarchical Bayesian approach. *Systematic Biology* 65(3):417–431.
33. Rohlf RV, Harrigan P, Nielsen R (2013) Modeling gene expression evolution with an extended Ornstein-Uhlenbeck process accounting for within-species variation. *Molecular biology and evolution* 31(1):201–211.
34. Wright S (1932) *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*. (na) Vol. 1.
35. Bokma F (2008) Detection of "punctuated equilibrium" by bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. *Evolution* 62(11):2718–2726.
36. Rabosky DL (2009) Ecological limits on clade diversification in higher taxa. *The American Naturalist* 173(5):662–674.
37. Barndorff-Nielsen OE, Mikosch T, Resnick SI (2012) *Lévy processes: theory and applications*. (Springer Science & Business Media).
38. Belomestny D (2010) Spectral estimation of the fractional order of a Lévy process. *The Annals of Statistics* pp. 317–351.
39. Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290.
40. Bihorel S (2015) The neldermead Package-version 1.0-10. CRAN 1:1.
41. Pennell MW et al. (2014) geiger v2. 0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* 30(15):2216–2218.
42. Nelder JA, Mead R (1965) A simplex method for function minimization. *The Computer Journal* 7(4):308–313.
43. Jetz W, Thomas G, Joy J, Hartmann K, Mooers A (2012) The global diversity of birds in space and time. *Nature* 491(7424):444–448.
44. Gingerich PD (2000) Arithmetic or geometric normality of biological variation: an empirical test of theory. *Journal of Theoretical Biology* 204(2):201–221.
45. Boettiger C, Lang D, Wainwright P (2012) rfishbase: exploring, manipulating and visualizing FishBase data from R. *Journal of Fish Biology* 81(6):2030–2039.
46. Gayrusuz Ö et al. (2006) Conversions of total, fork, and standard length measurements based on 42 marine and freshwater fish species (from Turkish waters). *Turkish Journal of Fisheries and Aquatic Sciences* 6(2):79–84.
47. Cowman P, Bellwood D (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *Journal of Evolutionary Biology* 24(12):2543–2562.
48. Santini F, Sorenson L, Alfaro ME (2013) A new phylogeny of tetraodontiform fishes (Tetraodontiformes, Acanthomorpha) based on 22 loci. *Molecular Phylogenetics and Evolution* 69(1):177–187.
49. Santini F, Sorenson L, Alfaro ME (2013) A new multi-locus timescale reveals the evolutionary basis of diversity patterns in triggerfishes and filefishes (balistidae, monacanthidae; tetraodontiformes). *Molecular Phylogenetics and Evolution* 69(1):165–176.
50. Sorenson L, Santini F, Carnevale G, Alfaro ME (2013) A multi-locus timetree of surgeonfishes (Acanthuridae, Percomorpha), with revised family taxonomy. *Molecular Phylogenetics and Evolution* 68(1):150–160.
51. Bloom DD, Lovejoy NR (2014) The evolutionary origins of diadromy inferred from a time-calibrated phylogeny for Clupeiformes (herring and allies). *Proc. R. Soc. B* 281(1778):20132081.
52. Roxo FF et al. (2014) Molecular phylogeny and biogeographic history of the armored Neotropical catfish subfamilies Hypoptopomatinae, Neoplecostominae and Otophyriniae (Siluriformes: Loricariidae). *PLoS One* 9(8):e105564.
53. Campanella D et al. (2015) Multi-locus fossil-calibrated phylogeny of Atheriniformes (Teleostei, Ovalentaria). *Molecular Phylogenetics and Evolution* 86:8–23.
54. Dornburg A, Moore J, Beaulieu JM, Eytan RI, Near TJ (2015) The impact of shifts in marine biodiversity hotspots on patterns of range evolution: Evidence from the Holocentridae (squirrelfishes and soldierfishes). *Evolution* 69(1):146–161.
55. Martin SD, Bonett RM (2015) Biogeography and divergent patterns of body size disparification in North American minnows. *Molecular Phylogenetics and Evolution* 93:17–28.
56. Froese R, Pauly D (2000) *FishBase 2000: Concepts Designs and Data Sources*. (World-Fish) Vol. 1594.
57. Pyron RA, Wiens JJ (2011) A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. *Molecular Phylogenetics and Evolution* 61(2):543–583.
58. De Lisle SP, Rowe L (2013) Correlated evolution of allometry and sexual dimorphism across higher taxa. *The American Naturalist* 182(5):630–639.
59. Wright AM, Lyons KM, Brandley MC, Hillis DM (2015) Which came first: the lizard or the egg? Robustness in phylogenetic reconstruction of ancestral states. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 324(6):504–516.
60. Jaffe AL, Slater GJ, Alfaro ME (2011) The evolution of island gigantism and body size variation in tortoises and turtles. *Biology Letters* p. rsl20101084.
61. Feldman A, Sabath N, Pyron RA, Mayrose I, Meiri S (2016) Body sizes and diversification rates of lizards, snakes, amphisbaenians and the tuatara. *Global Ecology and Biogeography* 25(2):187–197.
62. Dunning Jr JB (2007) *CRC Handbook of Avian Body Masses*, Second Edition. (CRC Press).
63. Vos RA, Mooers AO (2004) A new dated supertree of the Primates. (PhD thesis).
64. Slater GJ, Price SA, Santini F, Alfaro ME (2010) Diversity versus disparity and the radiation of modern cetaceans. *Proceedings of the Royal Society of London B: Biological Sciences* 277(1697):3097–3104.
65. Nyakatura K, Bininda-Emonds OR (2012) Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC Biology* 10(1):12.
66. Bibi F (2013) A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla, Ruminantia) and the importance of the fossil record to systematics. *BMC Evolutionary Biology* 13(1):166.
67. Mitchell KJ et al. (2014) Molecular phylogeny, biogeography, and habitat preference evolution of marsupials. *Molecular Biology and Evolution* 31(9):2322–2330.
68. Schenk JJ, Rowe KC, Steppan SJ (2013) Ecological opportunity and incumbency in the diversification of repeated continental colonizations by muroid rodents. *Systematic Biology* 62(6):837–864.
69. Shi JJ, Rabosky DL (2015) Speciation dynamics during the global radiation of extant bats. *Evolution* 69(6):1528–1545.

70. Jones KE et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90(9):2648–2648.