Iterative adaptive radiations of fossil canids show no evidence for diversity-dependent trait evolution

Graham J. Slater[1]

Department of Paleobiology, National Musuem of Natural History, Smithsonian Institution, Washington, DC 20001

Edited by David Jablonski, The University of Chicago, Chicago, IL, and approved October 20, 2014 (received for review April 19, 2014)

A long-standing hypothesis in adaptive radiation theory is that ecological opportunity constrains rates of phenotypic evolution, generating a burst of morphological disparity early in clade history. Empirical support for the early burst model is rare in comparative data, however. One possible reason for this lack of support is that most phylogenetic tests have focused on extant clades, neglecting information from fossil taxa. Here, I test for the expected signature of adaptive radiation using the outstanding 40-My fossil record of North American canids. Models implying time- and diversity-dependent rates of morphological evolution are strongly rejected for two ecologically important traits, body size and grinding area of the molar teeth. Instead, Ornstein–Uhlenbeck processes implying repeated, and sometimes rapid, attraction to distinct dietary adaptive peaks receive substantial support. Diversity-dependent rates of morphological evolution seem uncommon in clades, such as canids, that exhibit a pattern of replicated adaptive radiation. Instead, these clades might best be thought of as deterministic radiations in constrained Simpsonian subzones of a major adaptive zone. Support for adaptive peak models may be diagnostic of subzonal radiations. It remains to be seen whether early burst or ecological opportunity models can explain broader adaptive radiations, such as the evolution of higher taxa.

macroevolution | adaptive radiation | Canidae | phylogenetics | rates

A central prediction of modern adaptive radiation theory is that rates of diversification and phenotypic evolution are fastest early in clade history and subsequently slow as niches become saturated (1). This prediction is derived in large part from the writings of Simpson (2, 3), who suggested that fast rates of phenotypic evolution are required during the early phases of adaptive radiation to move lineages rapidly through inadaptive phases of the adaptive landscape to a new peak. Although the fossil record provides many examples of rapid early accumulation of morphological disparity (4–10), direct evidence for early rapid rates of phenotypic evolution, a so-called "early burst," has proved rare in phylogenetic comparative data (11). There are several plausible explanations for why early bursts are seldom observed. For example, temporally declining rates are difficult to detect in datasets comprising only extant taxa, and comparative methods are extremely sensitive to noise from convergence or measurement error (11, 12). Incorporation of fossil taxa in phylogenetic tests for early bursts can improve detection of these patterns (13), although isolated analyses of subclades within a larger adaptive radiation may fail to show evidence of declining disparity with time if rates have already significantly decreased, even when fossil species are sampled (14).

An alternative explanation for the lack of early bursts in comparative data is that ecological opportunity, not time, is the key determinant of rates of morphological evolution. If opportunity is the dominant force driving rates of morphological evolution in adaptive radiation, then the early burst model should be a particularly poor fit when clade age is a weak predictor of species richness. Patterns of diversity through time in the fossil record strongly suggest that diversity-dependent speciation and extinction dynamics are common (15–17) and weak or negative relationships between clade age and species richness are increasingly recognized in molecular phylogenies (18–20). To test the role of ecological opportunity in driving the morphological component of adaptive radiation more explicitly, Mahler et al. (21) developed a novel approach to model rates of morphological evolution as function of estimated past diversity at nodes of a time-calibrated molecular phylogeny (a similar method can be found in ref. 22). Using estimated lineage diversity as a proxy for past ecological opportunity in island communities of *Anolis* lizards, they found strong support for diversity dependence of evolutionary rates for body size and limb bone lengths, both of which influence habitat use and performance (21). This result would seem to lend support to a primary role for ecological opportunity in regulating rates of morphological evolution during adaptive radiation. It remains to be tested, however, whether diversity dependence can provide a compelling mechanism to explain patterns of morphological disparity through time in more general contexts, such as on continents, or over longer geological times scales.

In this article, I test for diversity dependence of rates of body size and dental evolution in living and fossil members of the dog family, Canidae. Canids are an attractive system for such a study. Living canids exhibit a range of dietary and predatory behaviors that can be readily diagnosed for fossil species on the basis of craniodental traits (23, 24). The canid fossil record is also well sampled, yielding a diverse radiation of ∼140 species (25–27) that spans the Late Eocene (40 Mya) through present day. Perhaps most importantly, however, for the first 35 My of their evolutionary history, canids were restricted to and diversified exclusively within North America. If ecological opportunity plays a prominent role in regulating rates of morphological diversification over geological time scales, we should expect to find

Significance

Many of the most diverse lineages of animals and plants are hypothesized to have arisen via the process of adaptive radiation. Most modern definitions of adaptive radiation focus on the role of ecological opportunity in regulating rates of morphological diversification. Using the rich fossil record of North American canids (wolves, foxes, and their relatives), I found no relationship between past diversity and rates of body size or tooth shape evolution. Instead, phenotypic evolution is directed toward adaptively optimal morphologies, such as large size or broad molars, without changes in the rate at which diversity accumulates. If these kinds of processes are common, we may need to rethink many of our ideas about the origins of morphological diversity.

Author contributions: G.J.S. designed research, performed research, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The phylogenetic data, morphological measurements, and scripts to replicate analyses have been deposited in the Dryad Digital Repository, datadryad.org (DOI no. 10.5061/dryad.9qd51).

[1]Email: slaterg@si.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1403666111/-/DCSupplemental.

support for a link between diversity and the tempo of ecomorphological diversification in continental radiations with excellent fossil records, such as Canidae.

Results and Discussion

Canids have traditionally been divided into three sequential, subfamilial radiations: Hesperocyoninae (Late Eocene–Middle Miocene), Borophaginae (Early Oligocene–Late Pliocene), and Caninae (Late Oligocene–present). Phylogenetic analysis of a morphological character matrix representing 120 species of living and fossil North American canids using Bayesian inference (28, 29) recovered all three clades, although only Caninae was strongly supported as monophyletic [posterior probability (PP) = 1; *Supporting Information*]. Time calibration of canid phylogeny using a relaxed, uncorrelated morphological clock calibrated by tip last appearance dates and 16 informative node age priors (29) yielded a maximum clade credibility tree that was highly unresolved toward the base of each subfamily (*Supporting Information*). I therefore used 500 randomly sampled trees from the Bayesian PP distribution, as well as the tree with the highest natural log likelihood [ln(likelihood)] (hereafter, the "best" tree; Fig. 1) for subsequent macroevolutionary analyses.

To understand how ecological opportunity may have influenced rates of morphological evolution, I estimated dietary diversity through time, using the method of Mahler et al. (21). Hypercarnivorous canids are those canids for which large-vertebrate prey comprise >70% of their diet, mesocarnivorous canids have diets composed of 50–70% vertebrate material, and hypocarnivorous canids consume <30% vertebrate materials and may heavily use invertebrate or plant material instead. Based on a series of linear discriminant function analyses (LDAs), I was able to classify 91 fossil canids to one of these categories. A symmetrical model, where transitions directly between hypercarnivory and hypocarnivory are predicted to be extremely rare, is the best-supported model of dietary evolution [median small sample-corrected Akaike information criterion weight (AICcW) = 0.69; *Supporting Information*], although model-averaged rates indicate less frequent transitions from hypercarnivory into mesocarnivory than vice versa (*Supporting Information*). Ancestral state estimation under the symmetrical model on the best tree from the Bayesian posterior distribution suggests that the most recent common ancestor of living and fossil canids was most likely a hypocarnivore but that transitions to mesocarnivory and hypercarnivory occurred relatively rapidly and multiple times (Fig. 1).

To determine whether morphological diversification in canids was influenced by ecological opportunity, clade age, or other factors, I fit a series of macroevolutionary models to two ecomorphological traits that have been shown previously to associate with resource use in carnivores. Body mass is positively correlated with maximum prey size in extant canids, and it is widely viewed as an important predictor of ecological niche in vertebrates. For ln(body mass), a model where evolutionary rates decline as a function of increasing diversity within dietary regimes (median $\psi = -0.001$, range = –0.012–0.049) received slightly higher support than models allowing for time- or diet-dependent variation in rates but received lower relative support than a constant-rate Brownian motion (BM) model (Table 1). An Ornstein–Uhlenbeck (OU) model with separate adaptive body mass peaks for the three dietary categories is preferred over all models, including diversity-dependent rates (median AICcW over all models compared = 0.73, Table 1). Based on the need to balance energy expended during foraging with energy obtained from prey, carnivores are predicted to switch to obligate hypercarnivory at sizes of 14.5–21 kg (30). The hypercarnivore optimum is here inferred to be 20.7 kg (median, range = 14.5–35.8 kg), in agreement with these predictions. Smaller optimal masses of 4.7 kg (2.9–9 kg) and 3.2 kg (2.6–4.8 kg) were estimated for mesocarnivorous and hypocarnivorous species, respectively (Table 1).

Further dissection of these results lends support to the view that body size plays an unpredictable role in macroevolutionary pattern (31), and casts doubt on its general utility in studying adaptive radiation. The α-parameter of the OU model describes the strength with which traits are attracted to their associated optima. It is often more intuitive, however, to reparameterize an OU model in terms of the phylogenetic half-life, $t_{1/2} = \ln(2)/\alpha$, which describes the time required for adaptation to a new selective regime to become more influential than constraints implied by an ancestral regime (32). For body mass, the median estimated α-parameter of 0.00074 leads to a $t_{1/2}$ of 936 My. Even α estimated on the best tree, an order of magnitude larger at 0.008, leads to a $t_{1/2}$ of 86 My. Such weak attraction suggests that canid body size evolution might be better explained as a slow sustained trend toward larger body sizes across the three dietary categories (32). Indeed, a trended random walk, albeit with a single trend for all species regardless of diet, provides the second best fit to these data (Table 1).

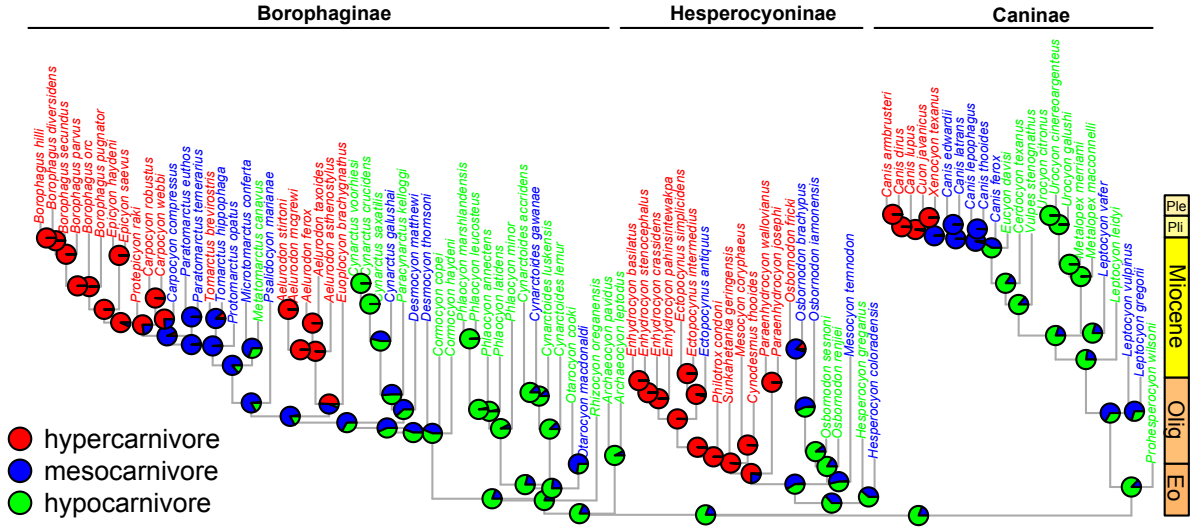

Fig. 1. Reconstruction of ancestral diet under a model of symmetrical transition rates on the time-calibrated tree with the highest log-likelihood from the Bayesian PP distribution. Eo, Eocene; Olig, Oligocene; Ple, Pleistocene; Pli, Pliocene.

Table 1. Median results from macroevolutionary models fitted to 500 trees sampled at random from the Bayesian posterior distribution

Trait	Model	lnLk	AICc	AICcW	σ^2	Scalar	Hypercarnivore	Mesocarnivore	Hypocarnivore
ln(mass)	BM	−43.46	91.06	<0.01	0.018	—	—	—	—
	ACDC	−43.40	93.08	<0.01	0.020	−0.002	—	—	—
	Diversity dependence	−43.35	92.96	<0.01	0.017	0	—	—	—
	Diversity-dependent diet	−42.65	91.56	<0.01	0.024	−0.001	—	—	—
	BMV	−42.72	93.89	<0.01	—	—	0.016	0.023	0.016
	Trend	−34.56	75.39	0.25	0.015	0.041	—	—	—
	OU	−31.33	73.34	0.73	0.014	7.4×10^{-5}	3.03	1.55	1.16
RLGA	BM	57.73	−111.32	<0.01	0.002	—	—	—	—
	ACDC	61.93	−117.59	<0.01	5.7×10^{-4}	0.052	—	—	—
	Diversity dependence	61.33	−116.37	<0.01	0.001	1×10^{-5}	—	—	—
	Diversity-dependent diet	58.69	−111.11	<0.01	0.002	-8×10^{-5}	—	—	—
	BMV	66.27	−124.08	<0.01	—	—	0.001	0.001	0.0035
	Trend	57.73	−109.19	<0.01	0.002	−0.001	—	—	—
	OU	77.01	−143.32	>0.99	0.004	0.182	0.63	0.80	0.97

Relative support is indicated by small sample AICcWs. The evolutionary rate is given by σ^2, whereas scalar refers to model-specific parameters [accelerating/decelerating rate (ACDC) = r; diversity dependent models = ψ, trended random walk (Trend) = μ, OU = α]. For variable rate (BMV) and multipeak OU models, diet-specific values are provided in the final three columns. AICc, corrected Akaike information criterion; lnLk, ln(likelihood).

The true picture of body size evolution is likely much more complex. A plot of canid body masses through time in relation to the putative adaptive optima (Fig. 2*A* and *Supporting Information*) reveals two striking patterns. First, although a large optimal mass of >20 kg is inferred for hypercarnivores, no canid achieves this size until the Middle Miocene, with *Osbornodon fricki* being the only hesperocyonine to do so. Second, there is a general increase in both mean and minimum canid body sizes toward the present that affects the size distribution of species within all three niches (33, 34). This pattern is indicative of a driven (35) or active (36) trend, Cope's rule in its strictest sense (37), and suggests that canid body size niches have themselves been shifting over the past 40 My (2, 3). The differences in maximum size of hypercarnivores before and after the Middle Miocene, combined with an apparent increase in minimum sizes through the Cenozoic (Fig. 2*A*), possibly represent environmentally induced changes in optimal body mass associated with post-Oligocene cooling and the expansion of grasslands, similar to the pattern found for North American horses (38). Regardless of the cause, overlaying of adaptive trends on simultaneously evolving niches complicates simple interpretation of macroevolutionary dynamics for canid body sizes. Many previous studies of adaptive radiation models (21, 39, 40) have examined patterns of body size evolution with the justification that it is a significant predictor of ecological niche (41). Although this claim may have physiological or biomechanical justification in some specific contexts, the incorporation of fossil data here, as in other studies (34, 42, 43), highlights that more complex and temporally variable patterns are common for body size, and serves as a warning against using it to draw general conclusions about the tempo and mode of ecological diversification within a clade, particularly when only extant members are sampled.

The second trait, relative lower grinding area (RLGA), is the combined areas of the talonid basin of the lower first molar (m1) and of the entire lower second molar, relative to body size (represented by m1 length) (23, 24). Hypercarnivorous canids should exhibit lower values of RLGA as the slicing function of the m1 trigonid is emphasized over grinding, whereas hypocarnivorous canids should exhibit large grinding areas for processing tough materials, such as plant matter and arthropod exoskeletons (23, 24). Considering rate models only, diversity-dependent models received low support (median AICcW < 0.01 for diversity and dietary diversity-dependent models) and models allowing for time-dependent increases (median AICcW = 0.02) or diet-specific rates (median AICcW = 0.95) were preferred. Considering the entire suite of models, RLGA is also best explained by an adaptive peak model (median AICcW > 0.99; Table 1). This model predicts a median optimal relative grinding area of 0.63 (range = 0.37–0.7) for hypercarnivorous canids, whereas hypocarnivorous canids are attracted to a larger area of 0.97 (0.90–1.03). Mesocarnivorous species are again intermediate, with optimal grinding areas of 0.8 (0.66–0.87).

In contrast to the results for body mass, the median estimated α-parameter over the posterior sample of trees (α = 0.18) implies strong and rapid attraction of RLGA toward these peaks ($t_{1/2}$ = 3.85 My). Furthermore, these zones remained remarkably stable throughout canid history and seem to have been unaffected by the biotic and abiotic factors that promoted macroevolutionary increases in canid body sizes. Despite the evolution of a few morphologically extreme hypocarnivores in the Middle to Late Miocene (*Cynarctus* species; Fig. 2*B*), most species cluster about their inferred optimal trait values. Presumably, the properties of vertebrate flesh have not changed much over the Cenozoic, and the functional

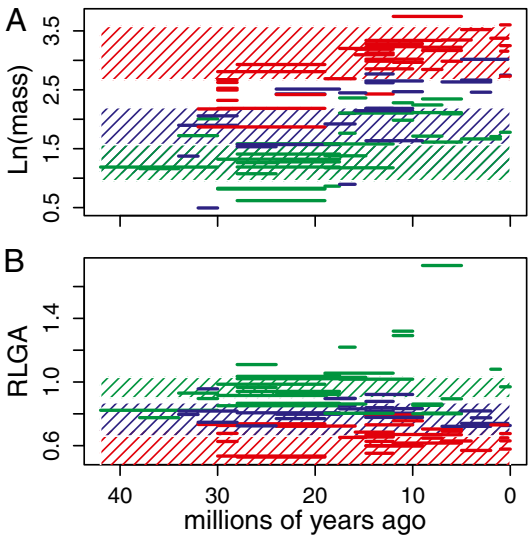

Fig. 2. Trait values through time for ln(body mass) (*A*) and RLGA (*B*). Hatched areas correspond to the range of optimal values estimated for a random sample of 500 trees from the Bayesian posterior distribution.

demands of processing meat are the same now as they were in the Late Eocene.

Selection of a multipeak OU process with a relatively strong attraction parameter for RLGA suggests that the 40-My radiation of North American canids might be best viewed as an iterative set of replicated ecological radiations (1, 23, 44, 45) occurring in and among at least three constant and bounded dietary zones. The concept of replicate adaptive radiation is generally viewed as being at odds with classic predictions of temporally declining or diversity-dependent rates of morphological evolution (46, 47). Their expected outcomes, at least in terms of the rate at which morphospace is filled, are remarkably consistent, however. Simulations show that early peak disparity is consistent with expectations of both early rapid evolutionary rates and evolution occurring at a constant rate but within a bounded morphospace (14). Under this second scenario, a limited range of morphological outcomes, due to developmental, ecological, or functional constraints (45), cause diversifying clades to saturate morphospace rapidly, after which disparity ceases to increase, regardless of subsequent increases in taxonomic diversity (6, 48, 49). Temporally declining rates and time-homogeneous rates in a bounded space cannot be distinguished based on disparity patterns alone (14), but if our attempts to interpret patterns of disparity through time focus on rate variation as the explanatory factor, as we have tended to do (1, 50, 51), then variation in evolutionary rates will seem all too prevalent (49, 52).

Simpson suggested the concept of subzones to differentiate fine-scale ecological diversification within adaptive zones from that among major ["totally distinct, essentially discontinuous" (ref. 2, p. 189)] adaptive zones. To the extent that subzones represent a highly restricted and narrowly circumscribed set of niches relative to the broader adaptive zone of the parent clade, then iterative or replicated patterns of deterministic ecomorphological evolution are perhaps to be expected. The fossil record of North American canids, along with many classic examples of adaptive radiation from the biological literature, such as anole ecomorphs, fits well within this definition, and support for multipeak OU processes may be diagnostic of subzonal diversification. If this hypothesis is correct, then much of what we have learned from empirical studies about the macroevolutionary dynamics of quantitative traits during adaptive radiation has come from a focus on deterministic subzonal radiations that may or may not scale to higher level processes. If early burst- or diversity-dependent–like patterns are to be found in comparative data, they may be restricted to higher level clades (2, 3, 14, 53–55). A renewed focus on understanding patterns of morphological diversification in the fossil record using phylogenetic methods provides our best hope of resolving this question.

It is possible that preference for a multipeak OU model is an artifact of correlations between RLGA and some of the variables used to classify fossil canids to dietary groups (*Supporting Information*). High convergence rates are expected for univariate traits evolving under BM-like processes (56, 57), and a worst-case scenario in which species are simply divided into regimes based on trait values simulated under BM results in false-positive rates of 100% in favor of a multipeak OU model (*Supporting Information*). Furthermore, predictive tests suggest that fitted model parameter estimates do not allow BM evolution to be ruled out for RLGA (*Supporting Information*). I suspect such an artifact is not driving the results obtained here. A large body of ecomorphological work indicates that RLGA convergently responds to macroevolutionary changes in diet among carnivores (58), providing a crucial phenotype environment correlation upon which an adaptive hypothesis can be based (45). Furthermore, the asymmetry of model-averaged rates between hypercarnivory and mesocarnivory (*Supporting Information*) would not be expected under a Brownian model. Nonetheless, this situation presents a particular challenge for modeling adaptive evolution in the fossil record, where ecological habits must often be inferred directly from traits of macroevolutionary interest or from alternative traits that show substantial correlations with them.

Even if the preferred multipeak OU models are discounted, time-homogeneous or diet-specific rate scenarios receive greater weight than traditional adaptive radiation models for body mass and RLGA. My results therefore add to a growing body of work that provides little support for the generality of time- or diversity-dependent controls on quantitative trait variation (11, 47, 59). There are reasonable explanations for why diversity-dependent regulation of evolutionary rates might go undetected in North American canids. An implicit assumption in this study is that ecological opportunity is determined only by the number of coexisting canid species. The fossil record of North American carnivorans is exceptionally diverse, and species from several clades, including Amphicyonidae, Ursidae, Felidae, and Procyonidae, almost certainly competed with fossil canids for dietary resources (23). Consideration of rates of ecomorphological evolution in a broader phylogenetic context could generate alternative interpretations to those made here. Similarly, the methods used in this study assume a relatively stable environment where the effect of variation in ecological opportunity on evolutionary rate is uniform across dietary groups and through time. This assumption seems questionable, given the trends observed for body size, and more refined models that allow for variation in these parameters would be useful for future tests.

More broadly, ecological opportunity could exert controls on the generation of morphological disparity in subzonal adaptive radiations without acting directly on trait evolutionary rates by regulating transitions between niches. Although I did not directly test for a signal of diversity-dependent transition rates, there is some evidence to suggest that such a process may have been operating. Semilog plots of reconstructed dietary diversity through time show that for most of the Cenozoic, dietary diversity in the North American canid guild remained relatively constant (Fig. 3), despite substantial turnover at the subfamilial and tribal levels (23) (Fig. 1). The most striking example of

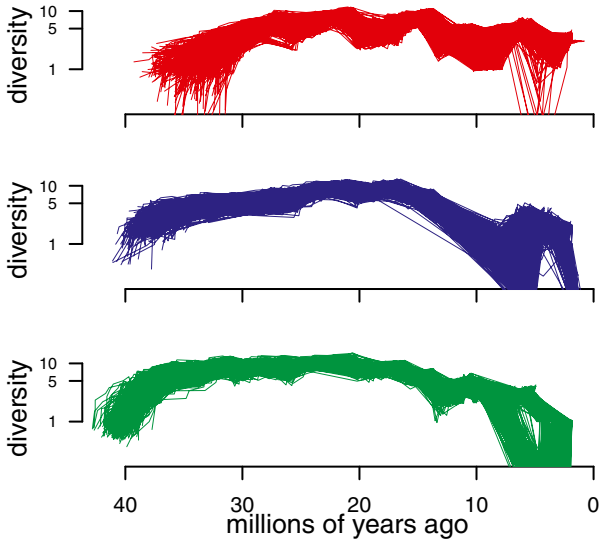

Fig. 3. Plots of diversity through time for hypercarnivores (*Top*), mesocarnivores (*Middle*), and hypocarnivores (*Bottom*), estimated using the method of Mahler et al. (21). Diversity is computed from marginal ancestral state estimates, and is therefore more conservative than are raw species counts with respect to uncertainty in dietary history. Trajectories are computed for 500 trees from the Bayesian posterior distribution to accommodate the effects of phylogenetic uncertainty. Note the logarithmic scale on the ordinate, rendering semilog plots.

replicate adaptive radiation, Caribbean anoles, shows exactly this signal, in fact, with strong evidence for diversity-dependent diversification rates on different islands (60) but support for multipeak OU models over diversity-dependent rates or early burst models for ecomorphological traits (47). Tests for diversity-dependent controls on niche occupancy, rather than underlying ecomorphological traits, would be a fruitful avenue for further tests of adaptive radiation in the fossil record.

Materials and Methods

Phylogenetic Inference. I assembled a matrix of 123 morphological characters coded for 121 canid species plus one outgroup (*SI Text*). The matrix was largely based on characters described elsewhere (25–27, 61). Although there is some overlap in characters among these sources, character states and codings often differed. I redefined character states, where possible, after merging similar characters and recoded taxa to fit with the new scheme. Some remaining gaps in the matrix were filled in based on primary examination of specimens or by reference to descriptions in the literature. For the final matrix, 30.9% of cells were coded as "?".

I used fossil tip dating (29) with an independent gamma rate (IGR) uncorrelated relaxed clock model, as implemented in MrBayes version 3.2 (28), to produce a distribution of time-calibrated canid phylogenies for comparative analyses. The stratigraphic range of each species was recorded from work by Wang and coworkers (25–27), and the last appearance date was assigned as a point constraint on occurrence time. I placed a relatively broad exponential prior ($\lambda = 127$) on the variance of the gamma distribution for the IGR model, and a log-normal prior on the clock rate (mean = −6, SD = 0.1). Tip dating neglects information from species' stratigraphic ranges, and preliminary analyses revealed that not accounting for first appearance dates resulted in unreasonably old divergence times (e.g., a Cretaceous origin for Canidae). I used a topology-only analysis before time calibration to identify unambiguous (PP = 1) nodes that could be constrained. The oldest first appearance date for members of each of the 15 clades, plus the root node, was used to define the offset for an exponential age prior on that node. The mean of each exponential distribution was chosen to generate a reasonable age range, typically with a mean of ∼2 My older and a 95% upper bound within 5 My of the first appearance date. I imposed slightly narrower priors on younger divergences. For both analyses, I used the default Markov model for morphological data, with coding set to variable and gamma-distributed rates. Two runs with four chains (three heated and one cold) were run for 1 million (topology only) and 5 million (time calibration) generations, sampling from the chain every 100 generations. After ensuring convergence using MrBayes diagnostics and trace plots of posterior sample values, I discarded the first 25% of samples. For the topology search, node support was assessed from a maximum clade credibility tree produced from the retained sample.

Morphological Data and Analyses. I collected morphological measurements from fossil and living canid specimens housed in the collections of the American Museum of Natural History and National Museum of Natural History. Measurements were taken using Mitutoyo digital calipers to 0.01-mm precision. I estimated body masses for fossil canids from lower m1 lengths, using Van Valkenburgh's regression equation (62). Mass estimates were ln-transformed before computing species means. Sample sizes ranged from 1 to 96 individuals per species for m1 length (median = 4) and 1–77 individuals per species for RLGA (median = 4).

I classified fossil canids to one of three dietary categories (hypercarnivore, mesocarnivore, or hypocarnivore) using LDA implemented in the Modern Applied Statistics with S (MASS) (63) package for R (64). The training set consisted of five well-represented and informative craniodental variables (relative blade length, blade size relative to jaw length, relative size of the lower second molar, mechanical advantage of the temporalis muscle, and robustness of the lower fourth premolar) measured for 21 extant canid species, nine procyonids (representing all genera except *Nasuella*), and *Ailurus fulgens*. Noncanid species were included because fossil canids exhibit more extreme adaptions to hypocarnivory than is realized in the extant radiation (25). This suite of craniodental traits permitted classification of fossil canid species without direct reference to the traits analyzed for evolutionary pattern (i.e., body size, RLGA), although correlations among all traits should be borne in mind when interpreting results (*SI Text*). Due to the fragmentary nature of much fossil material, I performed a series of discriminant analyses, iteratively reducing the number of variables to maximize species coverage. Two taxa, *Enhydrocyon stenocephalus* and *Urocyon galushi*, were classified using alternative variables. The classifications for these taxa were consistent with expectations based on congeners.

I estimated transition rates between dietary regimes under a Markov model using the fitDiscrete() function in geiger version 2.0 (65). Five hundred trees were sampled at random, without replacement, from the Bayesian posterior distribution, and models were fitted under equal ($k = 1$), symmetrical ($k = 3$), and all different ($k = 6$) rates models. Model fit was assessed by computing small sample AICcWs, which give a measure of relative support for each model, given the candidate pool. I also computed model-averaged rates (*SI Text*). To visualize the evolution of dietary strategy, I computed marginal ancestral state probabilities (66) on the best tree using the phytools library (67).

To understand the dynamics of body mass and RLGA evolution, I fit models of continuous trait evolution to the random sample of 500 trees. Constant-rate BM, temporally accelerating/decelerating rate, and trended random walk models were fitted using the fitContinuous() function in geiger. I fit diversity-dependent rate models using the fitDiversityModel() function in the phytools library (67). The ψ-parameter of the diversity-dependent models determines the magnitude and direction of change in rate per unit change in diversity, and is expected to be negative for a classic diversity-dependent process (21). I fit two permutations of the diversity-dependent rate model. In the first, I assumed a single regime, corresponding to a general lineage diversity dependence. For the second, lineage diversity was estimated based on dietary regimes, thereby allowing rates to vary as a function of dietary diversity through time. I fit multipeak OU and diet-specific Brownian rate (BMV) models using the OUwie library (68). Although it is possible to fit multipeak OU models that also allow α and σ^2 to vary across selective regimes, simulations show that parameter estimates are unreliable unless the number of terminals in the tree is exceptionally large (68). Because I was specifically interested in estimating the parameters of the OU model, I therefore restricted my analyses to models constraining α and σ^2 to be constant over states. I based model selection on computation of AICcW.

ACKNOWLEDGMENTS. I thank D. Jablonski and N. Shubin for inviting me to contribute this article, and D. Bohaska, J. Galkin, D. Lunde, and N. Pyenson for access to specimens. I am especially grateful to G. Hunt for much useful discussion of ideas, and to G. Hunt, J. Finarelli, and an anonymous reviewer for helpful comments and criticisms on earlier versions of the manuscript. This work was supported by a Peter Buck Post-Doctoral Fellowship at the National Museum of Natural History (Smithsonian Institution).

1. Schluter D (2000) *The Ecology of Adaptive Radiation* (Oxford Univ Press, Oxford).
2. Simpson GG (1944) *Tempo and Mode of Evolution* (Columbia Univ Press, New York).
3. Simpson GG (1953) *Major Features of Evolution* (Columbia Univ Press, New York).
4. Westoll TS (1949) On the evolution of the Dipnoi. *Genetics, Paleontology, and Evolution*, eds Jepsen GL, Simpson GG, Mayr E (Princeton Univ Press, Princeton), pp 121–184.
5. Gould SJ (1989) *Wonderful Life* (W. W. Norton, New York).
6. Foote M (1994) Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. *Paleobiology* 20(3):320–344.
7. Wagner PJ (1997) Patterns of morphological diversification among the Rostroconchia. *Paleobiology* 23(1):115–150.
8. Ruta M, Wagner PJ, Coates MI (2006) Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proc Biol Sci* 273(1598):2107–2111.
9. Erwin DH (2007) Disparity: Morphological pattern and developmental context. *Paleontology* 50(1):57–73.
10. Hughes M, Gerber S, Wills MA (2013) Clades reach highest morphological disparity early in their evolution. *Proc Natl Acad Sci USA* 110(34):13875–13879.
11. Harmon LJ, et al. (2010) Early bursts of body size and shape evolution are rare in comparative data. *Evolution* 64(8):2385–2396.
12. Slater GJ, Pennell MW (2014) Robust regression and posterior predictive simulation increase power to detect early bursts of trait evolution. *Syst Biol* 63(3):293–308.
13. Slater GJ, Harmon LJ, Alfaro ME (2012) Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66(12):3931–3944.
14. Foote M (1996) Models of morphological diversification. *Evolutionary Paleobiology*, eds Jablonski DJ, Erwin DE, Lipps JH (Univ of Chicago Press, Chicago), pp 62–88.
15. Raup DM, Gould SJ, Schopf TJM, Simberloff DS (1973) Stochastic models of phylogeny and the evolution of diversity. *J Geol* 81(5):525–542.
16. Alroy J (1988) Equilibrial diversity dynamics in North American mammals. *Biodiversity Dynamics: Turnover of Populations, Taxa, and Communities*, eds McKinney ML, Drake JA (Columbia Univ Press, New York), pp 233–287.
17. Foote M (2000) Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. *Paleobiology* 26(4):578–605.
18. Rabosky DL (2009) Ecological limits and diversification rate: Alternative paradigms to explain the variation in species richness among clades and regions. *Ecol Lett* 12(8):735–743.
19. Rabosky DL (2009) Ecological limits on clade diversification in higher taxa. *Am Nat* 173(5):662–674.
20. Rabosky DL, Slater GJ, Alfaro ME (2012) Clade age and species richness are decoupled across the eukaryotic tree of life. *PLoS Biol* 10(8):e1001381.

21. Mahler DL, Revell LJ, Glor RE, Losos JB (2010) Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* 64(9):2731–2745.
22. Weir JT, Mursleen S (2013) Diversity-dependent cladogenesis and trait evolution in the adaptive radiation of the auks (Aves: Alcidae). *Evolution* 67(2):403–416.
23. Van Valkenburgh B (1991) Iterative evolution of hypercarnivory in canids. *Paleobiology* 17(4):340–362.
24. Van Valkenburgh B, Koepfli K-P (1993) Cranial and dental adaptations to predation in canids. *Symposium of the Zoological Society of London* 65:15–37.
25. Wang X, Tedford RH, Taylor BE (1999) Phylogenetic systematics of the Borophaginae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* 243:1–391.
26. Wang X (1994) Phylogenetic systematics of the Hesperocyoninae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* 221:1–207.
27. Tedford RH, Wang X, Taylor BE (2009) Phylogenetic systematics of the North American fossil Caninae (Carnivora: Canidae). *Bulletin of the American Museum of Natural History* 325:1–218.
28. Ronquist F, et al. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61(3):539–542.
29. Ronquist F, et al. (2012) A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Syst Biol* 61(6):973–999.
30. Carbone C, Teacher A, Rowcliffe JM (2007) The costs of carnivory. *PLoS Biol* 5(2):e22.
31. Jablonksi D (1996) Body size and macroevolution. *Evolutionary Paleobiology*, eds Jablonski DJ, Erwin DE, Lipps JH (Univ of Chicago Press, Chicago), pp 256–289.
32. Hansen TF (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51(5):1341–1351.
33. Van Valkenburgh B, Wang X, Damuth J (2004) Cope's rule, hypercarnivory, and extinction in North American canids. *Science* 306(5693):101–104.
34. Finarelli JA (2007) Mechanisms behind active trends in body size evolution of the Canidae (Carnivora: Mammalia). *Am Nat* 170(6):876–885.
35. McShea DW (1994) Mechanisms of large-scale evolutionary trends. *Evolution* 48(6):1747–1763.
36. Wagner PJ (1996) Contrasting the underlying patterns of active trends in morphologic evolution. *Evolution* 50(3):990–1007.
37. Stanley SM (1973) An explanation for Cope's rule. *Evolution* 27(1):1–26.
38. MacFadden BJ (1986) Fossil horses from "Eohippus" (*Hyracotherium*) to *Equus*: Scaling, Cope's law, and the evolution of body size. *Paleobiology* 12(4):355–369.
39. Cooper N, Purvis A (2010) Body size evolution in mammals: Complexity in tempo and mode. *Am Nat* 175(6):727–738.
40. Venditti C, Meade A, Pagel M (2011) Multiple routes to mammalian diversity. *Nature* 479(7373):393–396.
41. Calder WA (1984) *Size, Function, and Life History* (Harvard Univ Press, Cambridge, MA).
42. Hunt G (2007) The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proc Natl Acad Sci USA* 104(47): 18404–18408.
43. Finarelli JA, Goswami A (2013) Potential pitfalls of reconstructing deep time evolutionary history with only extant data, a case study using the canidae (mammalia, carnivora). *Evolution* 67(12):3678–3685.
44. Losos JB (2010) Adaptive radiation, ecological opportunity, and evolutionary determinism. *Am Nat* 175(6):623–639.
45. Losos JB (2011) Convergence, adaptation, and constraint. *Evolution* 65(7):1827–1840.
46. Frédérich B, Sorenson L, Santini F, Slater GJ, Alfaro ME (2013) Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). *Am Nat* 181(1):94–113.
47. Mahler DL, Ingram T, Revell LJ, Losos JB (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341(6143):292–295.
48. Foote M (1995) Morphological diversification of Paleozoic crinoids. *Paleobiology* 20(3):424–444.
49. Hunt G (2012) Measuring rates of phenotypic evolution and the inseparability of tempo and mode. *Paleobiology* 38(3):351–373.
50. Harmon LJ, Schulte JA, 2nd, Larson A, Losos JB (2003) Tempo and mode of evolutionary radiation in iguanian lizards. *Science* 301(5635):961–964.
51. Freckleton RP, Harvey PH (2006) Detecting non-Brownian trait evolution in adaptive radiations. *PLoS Biol* 4(11):e373.
52. Slater GJ (2013) Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods Ecol Evol* 4(8): 734–744.
53. Osborn HF (1902) The law of adaptive radiation. *Am Nat* 36:353–363.
54. Valentine JW (2004) *On the Origin of Phyla* (Univ of Chicago Press, Chicago).
55. Humphreys AM, Barraclough TG (2014) The evolutionary reality of higher taxa in mammals. *Proc Biol Sci* 281(1783):20132750.
56. Polly PD (2004) On the simulation of the evolution of morphological shape: Multivariate shape under selection and drift. *Palaeontol Electronica* 7:7A.
57. Stayton CT (2008) Is convergence surprising? An examination of the frequency of convergence in simulated datasets. *J Theor Biol* 252(1):1–14.
58. Van Valkenburgh B (2007) Deja vu: The evolution of feeding morphologies in the Carnivora. *Integr Comp Biol* 47(1):147–163.
59. Ingram T, Harmon LJ, Shurin JB (2012) When should we expect early bursts of trait evolution in comparative data? Predictions from an evolutionary food web model. *J Evol Biol* 25(9):1902–1910.
60. Rabosky DL, Glor RE (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc Natl Acad Sci USA* 107(51):22178–22183.
61. Tedford RH, Taylor BE, Wang X (1995) Phylogeny of the Caninae (Carnivora, Canidae): The living taxa. *Am Mus Novit* 3146:1–37.
62. Van Valkenburgh B (1990) Skeletal and dental predictors of body mass in carnivores. *Body Size in Mammalian Paleobiology: Estimation and Biological Implications*, eds Damuth J, MacFadden BJ (Univ of Cambridge Press, Cambridge, UK), pp 181–205.
63. Venables WN, Ripley BD (2002) *Modern Applied Statistics with S* (Springer, New York), 4th Ed.
64. R Core Team (2013) *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna). Available at www.R-project.org/.
65. 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.
66. Yang Z (2006) *Computational Molecular Evolution* (Oxford Univ Press, Oxford).
67. Revell LJ (2012) phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3(2):217–223.
68. Beaulieu JM, Jhwueng DC, Boettiger C, O'Meara BC (2012) Modeling stabilizing selection: Expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* 66(8):2369–2383.

Supporting Information

Slater 10.1073/pnas.1403666111

SI Text

Phylogenetic Analysis

Morphological characters and their associated states are provided at the end of this section. Figs. S1 and S2 show the maximum clade credibility trees derived from a topology-only analysis and a simultaneous analysis of topology and divergence times, respectively.

Estimating Transition Rates Among Dietary Regimes

Fitting Markov models to the dietary regime data over 500 trees sampled from the Bayesian PP distribution revealed that a symmetrical model was most preferred. Table S1 shows the median transition rate estimates for each model, along with their ln(likelihoods), Akaike information criterion (AIC) scores, and AICcWs. The symmetrical model implies that forward and backward transition rates are identical between a given pair of states but differ among pairs of states. Although this model received the most support (median AICcW = 0.69), an all rates different model, where rates were free to vary for all transitions, received a moderate amount of relative support (median AICcW = 0.3).

To account better for uncertainty in model choice and parameter estimation, I computed model-averaged transition rates (Q) from parameter estimates for the 500 sampled trees. A model-averaged rate matrix is given in Table S2 and illustrated graphically in Fig. S5. The main effect of averaging over all three models is to decrease the rate of transitions from hypercarnivory into mesocarnivory ($Q_{12} = 0.031$) relative to the reverse ($Q_{21} = 0.053$). This result is intuitive, given that increasing adaptation to hypercarnivory has been viewed as a macroevolutionary ratchet that prevents reversion to more generalized diets and increases extinction risk (1).

Simulation Tests

Fitting OU models to comparative data requires that each terminal taxon in the tree, as well as each internal edge, is assigned to a particular macroevolutionary regime [2–4, but see ref. 5]. Extant species are typically assigned to a particular regime based on direct observation. For example, species can be assigned to dietary regimes for a subsequent analysis of dental evolution based on feeding observations taken from wild populations. For extinct taxa, ecology cannot be directly observed and must instead be inferred from fossilized remains. Circularity can arise, however, when the trait(s) used to classify species are the same as, or are correlated with, the trait(s) of macroevolutionary interest. This circularity need not be the case if an independent regime is imposed, for example, evolution before and after some temporally constrained event (2, 6, 7).

Arguably, we are often more interested in understanding whether trait evolution within a given clade is guided by selection toward distinct but temporally coincident adaptive peaks than by temporally distinct peaks. In such cases, an a priori classification method, such as discriminant analysis, must be used to classify species for which ecology is not known and cannot be observed via indirect means, such as fossilized stomach contents. In the best-case scenario, we might have a set of traits, for example, craniodental variables, that can be used to classify species to an evolutionary regime and a second set of traits, for example, postcranial metrics, that we wish to model according to the regimes previously inferred. This scenario is a best-case scenario because the classification variables are, by most standards, independent of the variables of macroevolutionary interest. Correlations may exist, but limbs are developmentally distinct from skulls and teeth, and so circularity is not implied. In the worst-case scenario, only one trait is available to classify species, and it happens to be the same one that we are interested in modeling. In this case, a circularity is implied; by classifying species based on the trait of interest, we would appear to bias ourselves toward finding support for evolutionary models that allow for distinct trait means according to ecology.

My discriminant analyses were based on set of traits that did not include the traits of interest, although they each exhibit some degree of correlation with one of the traits (RLGA) among the extant species sampled (Fig. S6 and Table S3). An ordination of the first two discriminant functions shows that both are required to discriminate between canids with different diets, particularly between mesocarnivores and hypocarnivores but also, to some extent, between hypercarnivores and mesocarnivores (Fig. S7). Only the first discriminant function correlates with RLGA, however, mostly due to the large negative loading of the relative length of the m1 blade on this axis, and species with different diets overlap in RLGA values along both axes (Fig. S8).

I examined the effect of a worst-case scenario, as described above, on model selection via simulation. Using the time-calibrated canid tree with the highest log-likelihood (the "best" tree; Fig. 1), I simulated 500 sets of trait data under a time-homogeneous, single-rate BM model. I used the maximum likelihood estimate (MLE) of σ^2 in all simulations, with no optimal trait values or attraction parameters. For each replicate, I then assigned species to one of three evolutionary regimes based on their simulated trait values by dividing the distribution of simulated trait values at the 0.33 and 0.66 quantiles. I then compared the fit of a time-homogeneous, single-rate BM model with the fit of a multipeak OU model with regimes determined according to species' trait values, as implemented in OUwie (4). I assessed model fit using small AICcWs.

The OU model was strongly preferred over the BM model for all simulated datasets (mean AICcW = 0.99, range = 0.99–1; Fig. S9). A false-positive rate of 100% is clearly a nonoptimal outcome and implies that support for OU models where terminals have been classified based on the trait being analyzed should be viewed with skepticism.

Further dissection of the results revealed a significant positive relationship between the estimated number of regime shifts in the simulated data and the MLEs of the α-parameter for the fitted OU model (Fig. S10). This result suggests a predictive test, where BM could be viewed as an insufficient model to explain the accumulation of morphological variation within and among dietary regimes if the MLE of α from the empirical dataset is greater than the MLE of α predicted by simulation, given an inferred number of transitions. For RLGA on the best canid tree (Fig. 1), the estimated number of transitions between dietary regimes is 24, a relatively but not significantly ($P = 0.16$; Fig. S11) low number compared with the BM simulations. The MLE of α is 0.3352. This model implies very rapid rates of adaptation in RLGA ($t_{1/2} = 2.1$ My). However, the predicted value of α for 24 dietary regime transitions under BM is 0.3346 (95% prediction interval = 0.25–0.41; Fig. S10). If we had classified canids to dietary regimes on the basis of RLGA values alone, there would clearly be a basis for rejecting, or at least being skeptical of, the best-fitting OU model.

Morphological Characters

Abbreviations: C, canine; I, incisors; M, molar; P, premolar. Uppercase refers to upper dentition, and lowercase refers to lower dentition.

1. Basal cusps on I1-2: absent (0); present (1) (8)
2. Size of I3: unenlarged relative to I1-2 (0); enlarged relative to I1-2 (1); greatly enlarged, possibly caniniform (2) (8–10)
3. Incisor row: curved, parabolic (0); straight line (1) (9)
4. I3 lateral cusps: I3 without lateral cusps (0); I3 with one lateral cusp (1); I3 with two lateral cusps (2); I3 with three lateral cusps (3) (9)
5. I1-3 medial cusps: present (0); absent on I3 only (1); weak or absent on I1-2 (2) (10)
6. c1 lateral groove: absent (0); present (1) (9)
7. Recurved c1: absent (0); present (1) (9)
8. Canine shape: long, slender (0); short, robust (1) (10)
9. Width of premolars: premolars not widened (0); premolars widened (1); premolars very wide and robust (2), premolars narrowed and slender (3) (8–10)
10. Premolar diastemata: closed premolar row (0); premolars separated by diastemata (1) (9, 10)
11. Anterior cingular premolar cusps: present on p2–4 (0); weak to absent on p2–4 (1); present only on p4 (2) (10)
12. Premolar crown height: premolars normal crown height (0); anterior premolars low crowned (1); high crowned (2) (9, 10)
13. Premolar length: short (0); elongate (1) (10)
14. P1: present (0); absent (1) (8)
15. p1: present (0); absent (1) (8, 9)
16. p3 posterior accessory cuspids: absent (0); present, moderately developed (1); present, enlarged (2) (8–10)
17. Height of principal cusp of p3 vs. p2, p4: forms ascending series or at same height (0); lies below p2 and p4 (1) (10)
18. Position of p3 crown base vs. crown base of p4: approximately same level as p4 when ramus is viewed laterally (0); crown base of p3 lies mostly below that of p4 (1) (10)
19. Second posterior accessory cusp of p3: absent (0); present (1) (9)
20. Relative size of p4: not greatly enlarged relative to p3 (0); greatly enlarged relative to p3 (1); further enlarged and reclined toward m1 (2) (8, 9)
21. p4 posterior accessory cusp position: located along midline of tooth (0); shifted laterally (1) (9)
22. p4 posterior accessory cusp size: moderate (0); enlarged (1); absent/lost (2) (9)
23. p4 second posterior accessory cusp: absent (0); present, lies between first posterior accessory cusp and cingulum (1); undifferentiated from posterior cingulum (2) (10)
24. p4 height relative to m1 paraconid: equals or exceeds m1 paraconid height (0); lower than m1 paranoid (1) (8, 9)
25. P4 protocone position: extends beyond anterior edge of paracone (0); medial to paracone (1) (8, 10)
26. P4 protocone size: unenlarged (0); enlarged (1); reduced (2); markedly reduced to small bulge with small root (3) (8–10)
27. P4 protocone and parastyle connection: protocone not connected to parastyle by a ridge (0); protocone connected to parastyle by a ridge (1) (9)
28. P4 parastyle: no P4 parastyle on anterior cingulum (0); parastyle originating from anterior cingulum separate from anterior ridge of paracone (1); strong ridge on anterior face of paracone (2); distinct parastyle as delineated by a notch on anterior ridge of paracone (3); parastyle prominently enlarged (4) (9)
29. P4 lingual cingulum or hypocone: internal cingulum weak or absent (0); cingulum thickened (1); cingulum raised to become hypocone (2) (9)
30. P4 shape: broad with strong anterior cingulum (0); narrow with weak anterior cingulum (1) (10)
31. M1 parastyle: large and salient (0); united with well-developed preparacrista (1); subdued, but remains united with preparacrista (2); preparacrista directed more anteriorly, lingual to parastyle (3) (8–10)
32. M1 paracone height: low, subequal to metacone (0); high and larger than metacone (1) (8–10)
33. M1 lingual cingulum: surrounds protocone anteriorly (0); posteriorly positioned and not surrounding protocone (1); anteriorly thickened (2) (8–10)
34. M1 labial cingulum at metacone: present at metacone (0); absent lateral to metacone (1) (9)
35. M1 labial cingulum at paracone: present at paracone (0); absent or subdued lateral to paracone (1) (9, 10)
36. M1 shape: anteroposteriorly short but transversely wide (0); subquadrate (1); longitudinally elongated (2) (9, 10)
37. M1 metaconule: weak or absent (0); present (1); large (2); metaconule split into two cusps (3) (9, 10)
38. M1 paraconule: weak or absent (0); distinct (1); enlarged (2) (9)
39. M1 posterior border shape: gently concave (0); sharply concave due to posterior extension of lingual cingulum (1) (9)
40. M1 hypocone: absent, lingual cingulum undifferentiated (0); hypocone present as swelling of posterior lingual cingulum (1); conical M1 hypocone present (2); conical M1 hypocone surrounded by cingulum (3) (9, 10)
41. M2 metacone: present, unreduced relative to paracone (0); reduced relative to paracone (1); extremely reduced or absent (2) (8, 9)
42. M2 hypocone: conical M2 hypocone absent (0); conical M2 hypocone present (1); conical M2 hypocone enlarged and posteriorly expanded (2) (9)
43. M2 metaconule and internal cingulum: metaconule and internal cingulum not connected (0); metaconule and internal cingulum connected by ridge (1); metaconule weak or absent (2) (9, 10)
44. M2 posterior cingulum: absent or weakly developed (0); present and well developed (1) (9)
45. M2 postprotocrista: present (0); incomplete or absent (1) (10)
46. M3: present (0); absent (1) (8)
47. m1 metaconid: present (0); unreduced, subequal to protoconid (1); present, reduced, almost vestigial (2); absent (3) (8, 10)
48. m1 protostylid: absent, weak ridge only (0); present, small (1); present, large, isolated from protoconid (2) (9, 10)
49. m1 anterior edge of paraconid: nearly linear and vertical (0); inclined posteriorly, may be curved (1) (10)
50. m1 trigonid elongation: m1 trigonid short (0); m1 trigonid elongated and open (1); m1 trigonid further shortened (2) (9)
51. m1 hypoconid form: hypoconid of m1 ridge-like (0); hypoconid of m1 conical (1) (8)
52. m1 entoconid form: poorly differentiated low crest on lingual border of talonid (0); discrete conical cusp (1) (9)
53. m1 talonid transverse cristid connecting hypoconid to entoconid: absent (0); present (1) (9)
54. m1 entoconulid: continuous with metaconid anteriorly or with small entoconulid between them (0); m1 entoconid deeply notched anteriorly, resulting in lingually open talonid (1); elevated, enlarged entoconulid anterior to entoconid (2) (9)
55. m1 talonid width: subequal with trigonid (0); wide relative to trigonoid (1); narrow relative to trigonid (2) (9)
56. m1 hypoconulid shelf: absent (0); present (1) (10)
57. m2 metaconid height: subequal in height with paraconid (0); shorter than paraconid (1); absent (2); much higher than protoconid (3) (8–10)

58. m1–2 selenodont: not selenodont (0); selenodont by development of crescentic labial talonid cusps (1) (9)
59. m2 size: unreduced/enlarged: smaller than m1 (0); enlarged relative to m1 (1); reduced in size relative to m1 (2) (8, 9)
60. m2 protostylid: protostylid absent (0); m2 protostylid present, small (1); m2 protostylid present, large (2) (9, 10)
61. m2 paraconid: present (0); weak or absent (1) (10)
62. m2 talonid length: talonid <90% trigonid (0); talonid >90% trigonid (1) (10)
63. m2 anterolabial cingulum: weak (0); well-developed, often reaching labial side of protocone (1) (10)
64. m3 trigonid cusps: two (0); single, centrally placed cusp (1); m3 absent (2) (8, 10)
65. m3 posterior shelf-like cingulum: absent (0); present (1) (10)
66. Rostrum length: mesocephalic (0); slightly dolichocephalic (1); extremely dolichocephalic, premolars widely spaced (2); relatively brachycephalic (3); extremely brachycephalic, premolars crowded (4) (8, 9)
67. Anterior palatine foramina length: short, posterior border lies at or anterior to posterior end of canine alveolus (0); long, posterior border lies posterior to canine alveolus (1) (10)
68. Palatine length: extends posterior to, or just anterior to, end of tooth row (0); extends beyond end of tooth row (1) (11)
69. Anterior process of frontal: smoothly curved frontal process (0); frontal process laterally pinched with sharp corner at its base (1) (8)
70. Nasal process of frontal: long (0); short (1) (10)
71. Premaxillary contact with frontal: premaxilla does not meet frontal, nasal, and maxilla in contact (0); premaxilla forms short contact with frontal, no nasal-maxilla contact (1); premaxilla forms broad contact with frontal (2) (9)
72. Nasal length: long, usually extending posteriorly beyond most posterior position of frontomaxillary suture (0); short, does not extend beyond frontomaxillary suture (1) (10)
73. Infraorbital foramen shape: rounded or oval shaped (0); compressed into a vertical slit (1) (8)
74. Elongation of orbital portion of skull: not elongated in middle section of skull (0); skull elongated, postorbital constriction situated well posterior of postorbital process (1) (8)
75. Foramen ovale and alisphenoid canal: separate (0); in common pit (1) (10)
76. Alisphenoid canal: present (0); absent (1) (12)
77. Frontal sinus: no frontal sinus, presence of a depression on dorsal surface of postorbital process (0); small frontal sinus that does not invade postorbital process or extend beyond postorbital constriction, depression may be retained (1); sinus invades postorbital process and may extend posteriorly to frontoparietal suture (2); sinus extends posteriorly beyond frontoparietal suture (3); sinus penetrates far back over top of entire braincase (4) (8–10)
78. Dorsal inflation of frontal sinus: flat forehead without dorsal inflation of sinus (0); small dome in forehead due to slight dorsal inflation (1); prominently domed forehead (2) (9)
79. Maxillojugal suture shape: obtuse (0); acute (1) (10)
80. Masseteric scar on zygomatic arch: masseteric scar wide and deep, particularly anteriorly, occupies >1/2 of lateral surface of zygomatic arch (0); masseteric scar narrow, restricted mostly to ventral face of arch and <1/3 depth of arch in lateral view (1) (9)
81. Orbital margin of zygomatic arch: laterally flared and everted (0); not laterally flared or everted (1) (9, 10)
82. Widest point of zygomatic arch: zygomatic arch gently curved laterally in dorsal view, widest point close to middle (0); widest point posteriorly shifted close to glenoid fossa, creating a more angled than arched appearance in dorsal view (1) (9)
83. Lateral expansion of zygomatic arches: unexpanded (0); laterally expanded (1) (8)
84. Zygomatic arch shape: nearly flat or moderately arched in lateral view (0); strongly dorsoventrally arched (1) (10)
85. Form of temporal crests: single sagittal crest formed by merging of temporal crests behind postorbital process (0); double crested, often lyrate, but with little or no reinforcement of crests (1); strong, widely separated and parallel temporal crests enclosing a longitudinal valley (2) (9)
86. Height of sagittal crest: sagittal crest low (0); sagittal crest high (1); sagittal crest very high (2) (9)
87. Sagittal crest profile: dorsally arched or straight (0); concave (1) (9)
88. Sagittal crest location: confined to parietal (0); extends on to frontal (1) (10)
89. Lambdoidal crest constriction: lambdoidal crest not constricted (0); nuchal portion of crest constricted, forming a rectangular plate (1) (9)
90. Postparietal foramen: present (0); absent (1) (10)
91. Dorsal exposure of cerebellum: significant exposure dorsoposteriorly between cerebrum and lambdoidal crest (0); completely overlapped by cerebrum, not exposed dorsoposteriorly (1) (10)
92. Supraoccipital shield form: rectangular or fan-shaped, inion not overhanging condyles (0); triangular in shape, inion often pointed and overhanging condyles (1) (9, 10)
93. Insertion for rectus capitis dorsalis muscle above occipital condyles: no fossae present, smooth surface (0); rounded fossae present (1) (8)
94. Suprameatal fossa: absent (0); presence of a small suprameatal fossa (1); suprameatal fossa present and enlarged (2) (8, 9)
95. Entotympanic bulla: unossified or absent (0); ossified and present (1) (8)
96. Bulla size: uninflated for canid type (0); bulla inflated (1); bulla hypertrophied (2); bulla shortened (3) (8, 9)
97. Internal carotid artery: intrabullar (transpromontorial) (0); extrabullar, between entotympanic and petrosal, and located dorsal to basioccipital-entotympanic suture (1); extrabullar, embedded within entotympanic, and located ventral to basioccipital-entotympanic suture (2) (8)
98. Promontorium shape: globular and isolated (0); medially and anteriorly expanded, in contact with surrounding bone (1) (8)
99. Ectotympanic ring: ectotympanic forms half-ring, dorsal roof of external auditory meatus formed by squamosal (0); ectotympanic ring complete, forms entire external auditory opening (1) (9)
100. Meatal tube: absence of tubular meatus (0); presence of a short tube (1); presence of an elongate tube (2) (9)
101. Direction of paraoccipital processes: posteriorly oriented and not fused with bulla (0); ventrally directed, not widely fused with bulla (1); ventrally directed, fused with bulla along entire length (2) (8–10)
102. Size of paraoccipital process: short and slender (0); elongate and robust (1); further hypertrophied and keeled posteriorly (2) (8–10)
103. Lateral expansion of paraoccipital process: no lateral expansion (0); laterally extended (1); further extension into a wide blade (2) (8, 10)
104. Posterior expansion of paraoccipital: no posterior extension (0); posteriorly extended (1); further extended, leading to development of a longitudinal plate (2) (8, 11)
105. Mastoid process form and size: small, crest-like (0); large, knob-like, inflated (1); prominently inflated beyond squamosal shelf (2); ventrally expanded (3); mastoid process very reduced, further receded under squamosal shelf (4) (8–10)

106. Lateral ridge of mastoid: absent (0); present (1) (9)
107. Postglenoid foramen and ectotympanic: postglenoid foramen enclosed posteriorly by ectotympanic (0); postglenoid foramen not enclosed posteriorly (1); ectotympanic extensively fused with postglenoid process (2) (8)
108. Optic foramen and anterior lacerate foramen: separate (0); in common pit (1) (10)
109. Medial pocket of angular process: absence of a pocket (0); presence of a deep medial pocket on angular process formed by lateral and dorsal extension of internal ridge (1) (8)
110. Angular process shape: attenuated, often with dorsal hook (0); shortened, blunt, deep (1) (8–10)
111. Relative size of fossa for superior ramus of medial pterygoideus muscle: small (0); fossa for superior ramus much larger than inferior (1) (10)
112. Subangular lobe of mandible: absent (0); present, rounded, smooth (1); present, sharply distinguished (2); present, dorsoventrally deep, angular (3) (9, 10)
113. Masseteric fossa ventral margin excavation: shallow (0); deeply excavated into ventral rim (1) (8)
114. Masseteric fossa ventral margin position: close to lower border of horizontal ramus (0); elevated, creating space between lower border of ramus and masseteric fossa (1) (9)
115. Masseteric fossa anterior margin: anterior margin of fossa not excavated (0); deeply excavated anteriorly (1) (9)
116. Shape of horizontal ramus: deep and strong (0); shallow and slender (1) (9, 10)
117. Symphyseal flange: ramus without flange (0); ramus with a symphyseal flange (1) (8)
118. Position of condyle: low (0); elevated above tooth row (1) (8)
119. Entepicondylar foramen of humerus: present (0); absent (1) (10)
120. Supratrochlear foramen of humerus: absent (0); present (1)
121. Metatarsal I: present with phalanges (0); reduced to rudiment, lacking phalanges (1) (10)
122. Radial-tibial ratio: <80% (0); 80–90% (1); >90% (2) (10)
123. Baculum size and shape: long, curved (0); short, straight (1) (8)

1. Van Valkenburgh B (1991) Iterative evolution of hypercarnivory in canids. *Paleobiology* 17(4):340–362.
2. Hansen TF (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51(5):1341–1351.
3. Butler MA, King AA (2004) Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *Am Nat* 164(6):683–695.
4. Beaulieu JM, Jhwueng DC, Boettiger C, O'Meara BC (2012) Modeling stabilizing selection: Expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* 66(8):2369–2383.
5. Mahler DL, Ingram T, Revell LJ, Losos JB (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341(6143): 292–295.
6. Slater GJ (2013) Phylogenetic evidence for a shift in the mode of mammalian body size evolution at the Cretaceous-Palaeogene boundary. *Methods Ecol Evol* 4(8):734–744.
7. Benson RBJ, Frigot RA, Goswami A, Andres B, Butler RJ (2014) Competition and constraint drove Cope's rule in the evolution of giant flying reptiles. *Nat Commun* 5:3567.
8. Wang X (1994) Phylogenetic systematics of the Hesperocyoninae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* 221:1–207.
9. Wang X, Tedford RH, Taylor BE (1999) Phylogenetic systematics of the Borophaginae (Carnivora, Canidae). *Bulletin of the American Museum of Natural History* 243:1–391.
10. Tedford RH, Wang X, Taylor BE (2009) Phylogenetic systematics of the North American fossil Caninae (Carnivora: Canidae). *Bulletin of the American Museum of Natural History* 325:1–218.
11. Tedford RH, Wang X, Taylor BE (1995) Phylogeny of the Caninae (Carnivora, Canidae): The Living Taxa. *Am Mus Novit* 3146:1–37.
12. Wesley–Hunt GD, Flynn JJ (2005) Phylogeny of the Carnivora: Basal relationships among the carnivoramorphans, and assessment of the position of 'Miacoidea' relative to Carnivora. *J Syst Palaeontology* 3(1):1–28.

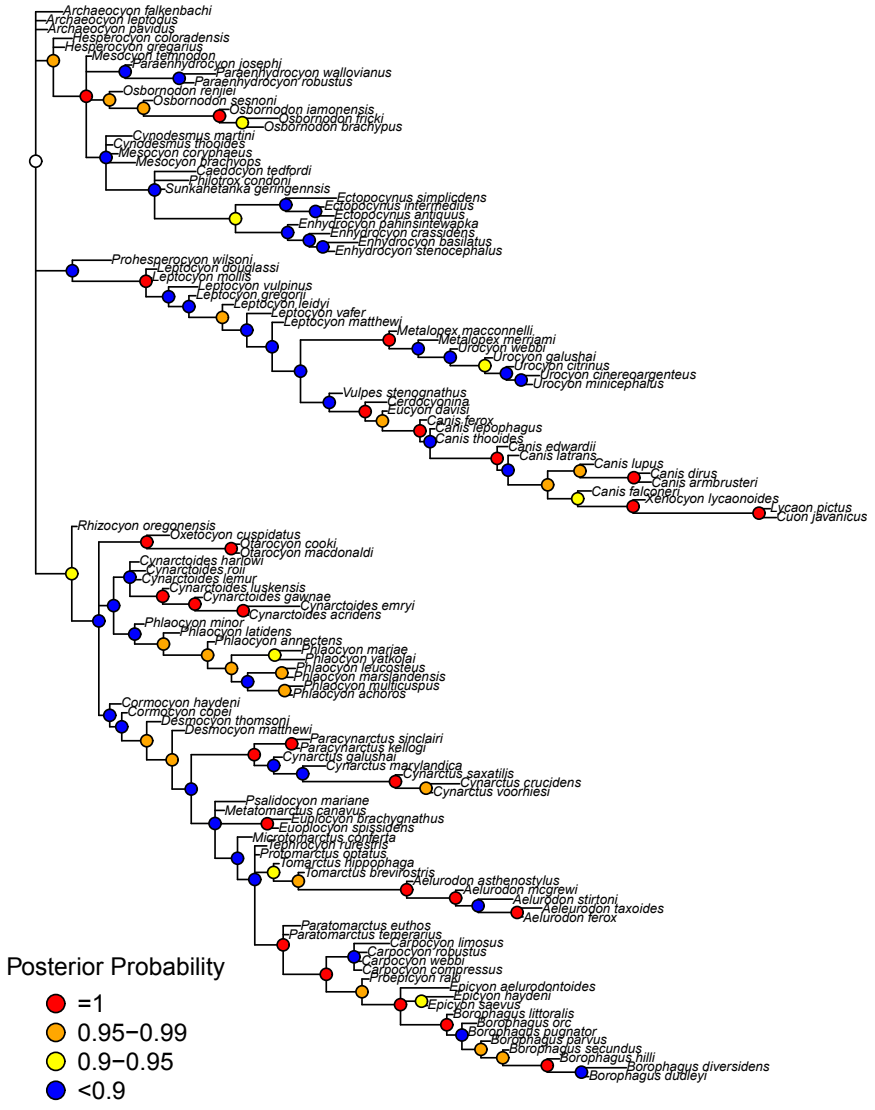

Fig. S1. Maximum clade credibility phylogram from a Bayesian analysis of 124 characters coded for 121 living and fossil canids plus one outgroup (removed).

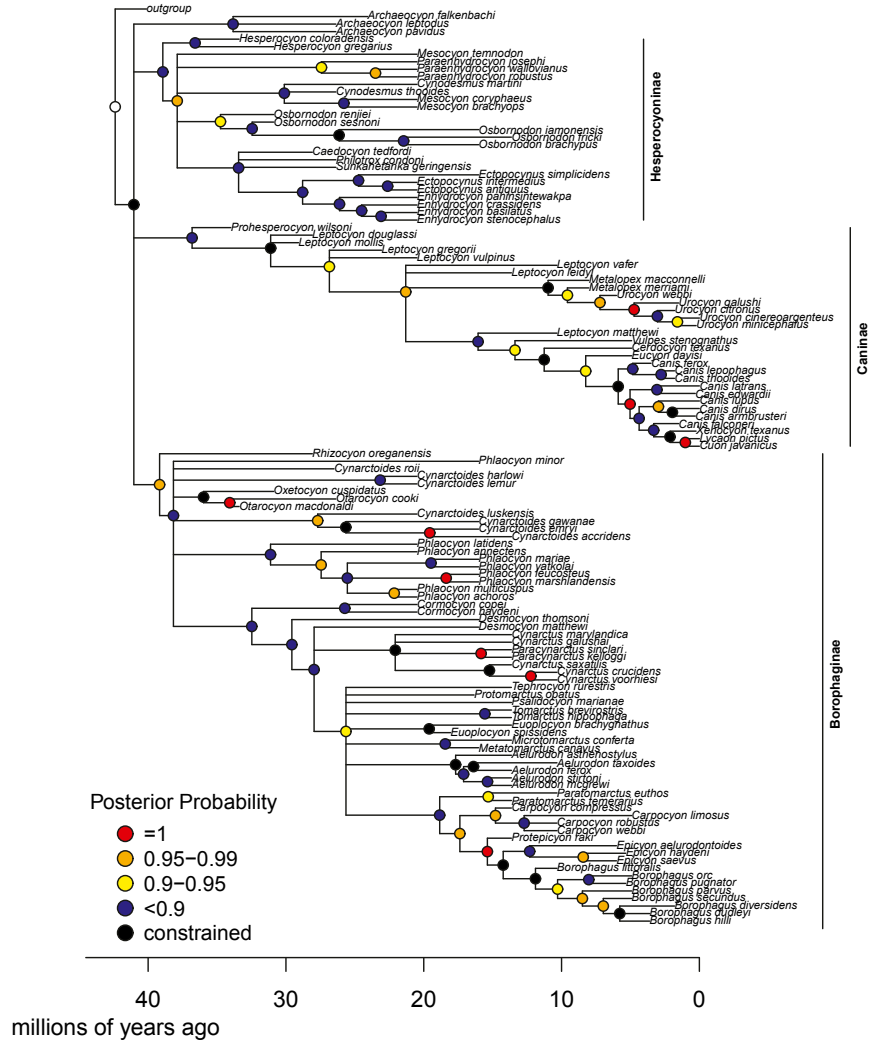

Fig. S2. Maximum clade credibility chronogram from a Bayesian tip dating analysis of 124 characters coded for 121 living and fossil canids plus one outgroup. Fossil species are plotted at their last appearance date. Divergence times are estimated under a relaxed, uncorrelated morphological clock, and therefore may occur after the first appearance date(s) in the fossil record of one or more daughter taxa. Constrained nodes are those nodes that were unambiguously supported in the topology analysis (PP = 1; Fig. S1), and where the oldest first appearance date of a descendant taxon was used to derive an informative node age prior.

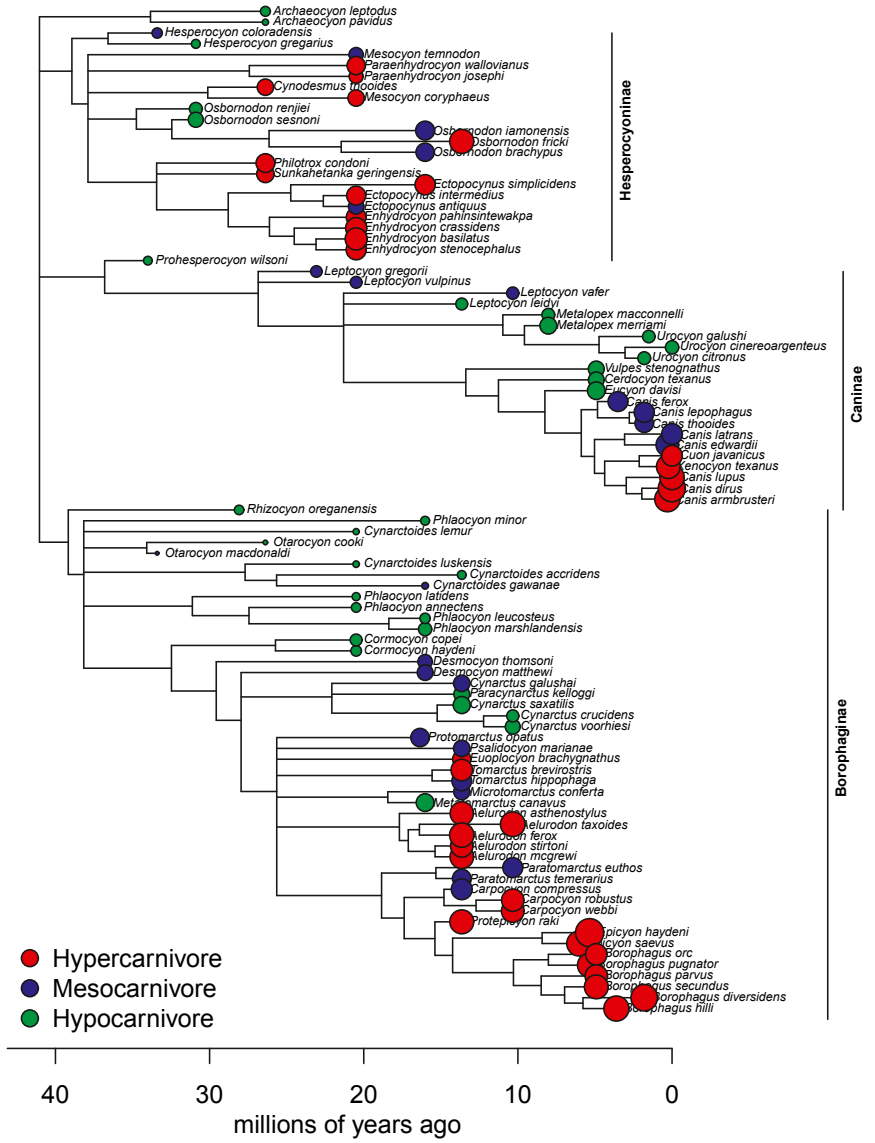

Fig. S3. Maximum clade credibility chronogram showing distribution of body size variation over canid subfamilies and through time. Circles are colored according to diet, with size scaled according to ln(mass).

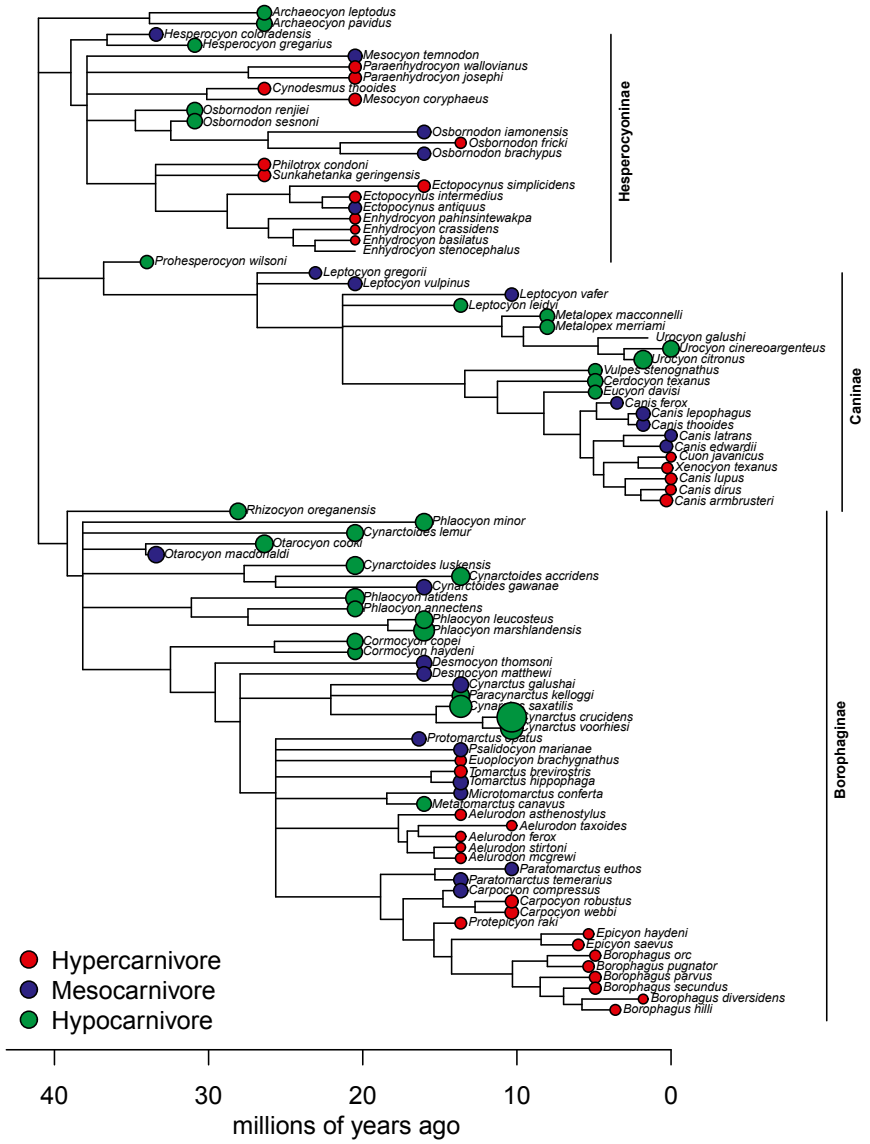

Fig. S4. Maximum clade credibility chronogram showing distribution of RLGA variation over canid subfamilies and through time. Circles are colored according to diet, with size scaled according to RLGA.

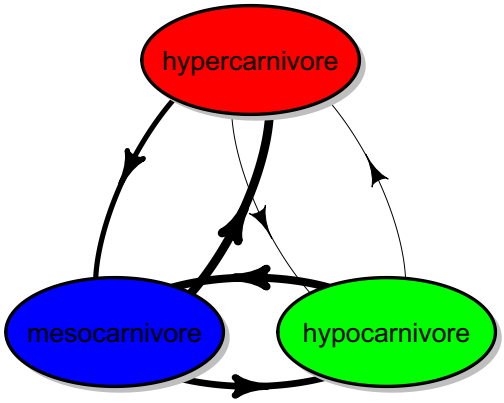

Fig. S5. Illustration of mean model-averaged rates between dietary states derived from 500 model fits. Although a symmetrical model is most strongly supported, model averaging of rates shows that transitions from hypercarnivore to mesocarnivore occur less frequently than transitions from mesocarnivore to hypercarnivore.

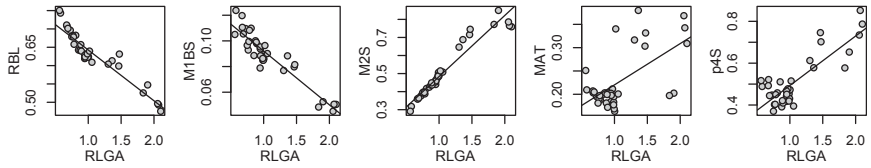

Fig. S6. Regression of the variables used to classify fossil canids to dietary groups on RLGA reveals significant correlations for all five traits (Table S3). MAT, moment arm of temporalis; M1BS, lower first molar blade size (relative to jaw length); M2S, relative size of upper second molar; p4S, shape of lower first premolar; RBL, relative length of the m1 blade.

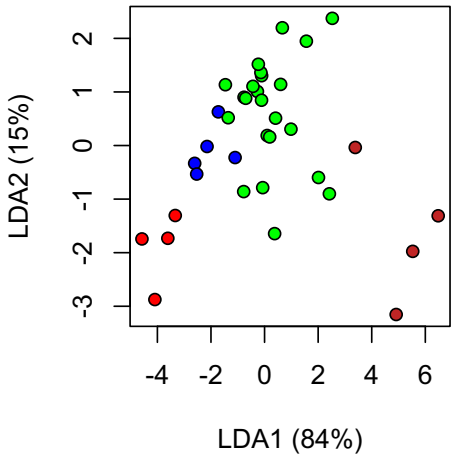

Fig. S7. Ordination of extant canids and procyonids on the first two discriminant functions. Colors correspond to dietary regimes: red, hypercarnivore; blue, mesocarnivore; green, hypocarnivore; brown, herbivore. Only procyonids and the red panda *Ailurus* are herbivorous. Note the substantial overlap of mesocarnivores and hypocarnivores on DF1 and of all groups on DF2. DF, discriminant function.

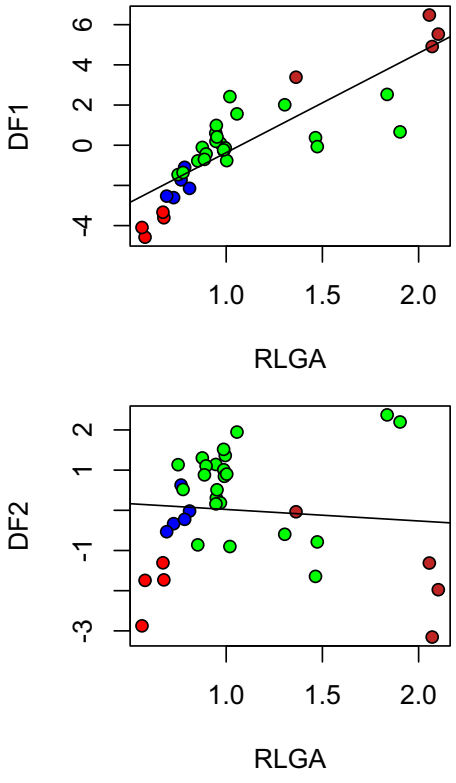

Fig. S8. Scatter plots showing the relationship between RLGA and scores on the first two DFs. The regression of DF1 on RLGA is significant ($R^2 = 0.72, P < 0.001$), but the regression of DF2 is not ($R^2 = -0.02, P = 0.6$).

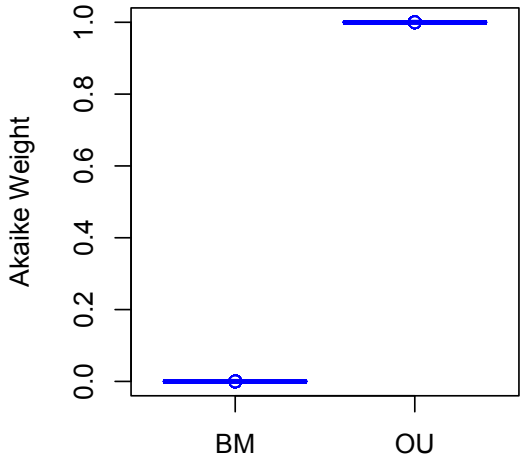

Fig. S9. AICcWs for BM and OU models fitted to 500 datasets simulated under BM and divided into three regimes based on trait values. The false-positive rate is 100%.

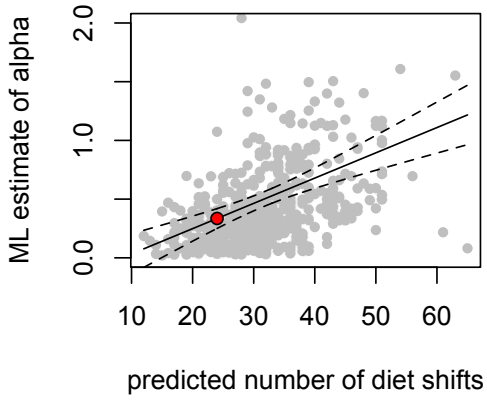

Fig. S10. There is a significant positive relationship ($R^2 = 0.06, P < 0.001$) between the predicted number of regime shifts in BM-simulated data and the MLE of the α-parameter of the OU model. However, the MLE of α from the empirical fossil canid dataset (red circle) falls within the 95% prediction envelope of this model, meaning that we cannot rule out BM in favor of OU in a worst-case scenario. The linear model is $\alpha = -0.18 + 0.215$.

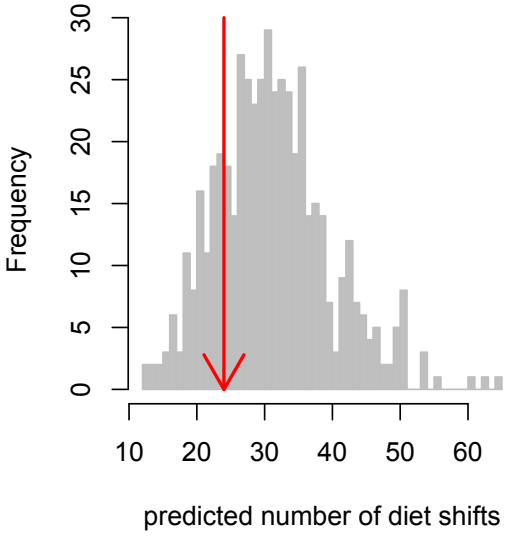

Fig. S11. Frequency distribution for the number of regime shifts (inferred via ancestral state estimation) implied in data simulated under BM and divided according to trait values. The number of regime shifts inferred for North American canids is shown by the red arrow; this number is not significantly different from the null distribution.

Table S1. Median transition rates and relative model support for discrete models of dietary evolution

Model	Q_{12}	Q_{13}	Q_{21}	Q_{23}	Q_{31}	Q_{32}	LnLk	AICc	AICcW
ER	0.0247	0.0247	0.0247	0.0247	0.0247	0.0247	−78.98	159.96	<0.01
SYMM	0.0412	0.0000	0.0412	0.0478	0.0000	0.0478	−72.07	150.14	0.69
ARD	0.0139	0.0000	0.0717	0.0292	0.0000	0.0466	−69.38	150.77	0.30

ARD, all rates different; ER, equal rates; SYM, symmetric rates.

Table S2. Model-averaged transition rates between dietary states

State	1	2	3
1	—	0.03	0.0
2	0.05	—	0.04
3	0.00	0.05	—

1, hypercarnivory state; 2, mesocarnivory state; 3, hypocarnivory state.

Table S3. Variable loadings on the three linear discriminant (LD) functions used to classify fossil canids to dietary regimes, and the strength and significance of their correlations with RLGA

Trait	LD1	LD2	LD3	r^2	*P* value
RBL	−45.49	1.35	−31.12	0.88	<0.001
M1BS	1.48	1.70	79.09	0.86	<0.001
M2S	−3.32	12.75	−4.74	0.93	<0.001
MAT	24.71	−8.26	13.08	0.36	<0.001
p4S	−9.92	−15.26	−6.59	0.69	<0.001

MAT, moment arm of temporalis; M1BS, lower first molar blade size (relative to jaw length); M2S, relative size of upper second molar; p4S, shape of lower first premolar; RBL, relative length of the m1 blade.