Supporting Information

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 (Q21 = 0.031) relative to the reverse (Q12 = 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 $\sigma^2$ 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 $\text{AICcW} = 0.99$, 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 $\alpha$-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 $\alpha$ from the empirical dataset is greater than the MLE of $\alpha$ 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 ($P = 0.16; \text{Fig. S11}$) low number compared with the BM simulations. The MLE of $\alpha$ is 0.3352. This model implies very rapid rates of adaptation in RLGA ($t_{1/2} = 2.1$ My). However, the predicted value of $\alpha$ 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. Curved 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 diastema: closed premolar row (0); premolars separated by diastema (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 inclined to trigonid (2) (9, 10)
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 shape: broad with strong anterior cingulum (0); narrow with weak anterior cingulum (1) (10)
25. 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)
26. M1 paracone height: low, subequal to metacone (0); high and larger than metacone (1) (8–10)
27. M1 lingual cingulum: surrounds protocone anteriorly (0); posteriorly positioned and not surrounding protocone (1); anteriorly thickened (2) (8–10)
28. M1 labial cingulum at metacone: present at metacone (0); absent lateral to metacone (1) (9)
29. M1 lingual cingulum at paracone: present at paracone (0); absent or subdued lateral to paracone (1) (9, 10)
30. M1 shape: anteroposteriorly short but transversely wide (0); subquadrangular (1); longitudinally elongated (2) (9, 10)
31. M1 metacone: weak or absent (0); present (1); large (2); metaconule split into two cusps (3) (9, 10)
32. M1 paracrista: weak or absent (0); distinct (1); enlarged (2) (9)
33. M1 posterior border shape: gently concave (0); sharply concave due to posterior extension of lingual cingulum (1) (9)
34. 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)
35. M2 metacone: present, unreduced relative to paracone (0); reduced relative to paracone (1); extremely reduced or absent (2) (8, 9)
36. M2 hypocone: conical M2 hypocone absent (0); conical M2 hypocone present (1); conical M2 hypocone enlarged and posteriorly expanded (2) (9)
37. 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)
38. M2 posterior cingulum: absent or weakly developed (0); present and well developed (1) (9)
39. M2 postprotocrista: present (0); incomplete or absent (1) (10)
40. M3: present (0); absent (1) (8, 9)
41. M1 metaconid: present (0); unreduced, subequal to protocristid (1); present, reduced, almost vestigial (2); absent (3) (8, 10)
42. M1 protostylid: absent, weak ridge only (0); present, small (1); present, large, isolated from protoconid (2) (9, 10)
43. M1 anterior edge of paraconid: nearly linear and vertical (0); inclined posteriorly, may be curved (1) (10)
44. M1 trigonid elongation: m1 trigonid short (0); m1 trigonid elongated and open (1); m1 trigonid further shortened (2) (9)
45. M1 hypoconid form: hypoconid of m1 ridge-like (0); hypoconid of m1 conical (1) (8)
46. M1 entoconid position: poorly differentiated low crest (0); discrete conical cusp (1) (9)
47. M1 paraconid: absent (0); present, reduced, almost vestigial (1); present, large, isolated from protoconid (2) (9, 10)
48. M1 lingual border of talonid: continuous with metaconid anteriorly or with small entoconulid between them (0); entoconulid deeply notched anteriorly, resulting in lingually open talonid (1); elevated, enlarged entoconulid anterior to entoconid (2) (9)
49. M1 molar length: subequal to trigonid (0); wide relative to trigonid (1); narrow relative to trigonid (2) (9)
50. M1 entoconulid shelf: absent (0); present (1) (10)
51. M2 metaconid height: subequal in height with paraconid (0); shorter than paraconid (1); absent (2); much higher than paraconid (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 protoconid: absent (0); m2 protoconid present, small (1); m2 protoconid large, double (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 palateine 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, posterior orbital constriction situated well posterior of postorbital process (1) (8)
75. Infratemporal fossa 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); prominent domed forehead (2) (9)
79. Maxillolugal 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 dorsally posteriorly between cerebrum and lambdoidal crest (0); completely overlapped by cerebrum, not exposed dorsally posteriorly (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. Ento tympanic 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-ento tympanic suture (1); extrabullar, embedded within entotympanic, and located ventral to basioccipital-ento tympanic suture (2) (8)
98. Promontorium shape: globular and isolated (0); medially and anteriorly expanded, in contact with surrounding bone (1) (8)
99. Ecotympanic ring: ecotympanic forms half-ring, dorsal roof of external auditory meatus formed by squamosal (0); ecotympanic 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 paroccipital 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 paroccipital process: short and slender (0); elongate and robust (1); further hypertrophied and keeled posteriorly (2) (8–10)
103. Lateral expansion of paroccipital process: no lateral expansion (0); laterally extended (1); further extension into a wide blade (2) (8, 10)
104. Posterior expansion of paroccipital: 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, crescent-like (0); large, knob-like, inflated (1); prominently inflated beyond squamosal shelf (2); ventrally expanded (3); mastoid process very reduced, further reeded 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)

---

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).
Hesperocyon wilsoni
Hesperocyon gregarius
Hesperocyon coloradensis
Mesocyon temnodon
Mesocyon coryphaeus
Cynodesmus thooides
Sunkahetanka geringensis
Philotrox condoni
Enhydrocyon pahinsintewakpa
Enhydrocyon crassidens
Enhydrocyon stenocephalus
Enhydrocyon basilatus
Osbornodon renjiei
Osbornodon sesnoni
Osbornodon iamonensis
Osbornodon brachypus
Osbornodon fricki
Paraenhydrocyon josephi
Paraenhydrocyon wallovianus
Ectopocynus antiquus
Ectopocynus intermedius
Ectopocynus simplicidens
Archaeocyon pavidus
Archaeocyon leptodus
Otarocyon macdonaldi
Otarocyon cooki
Rhizocyon oreganensis
Cynarctoides lemur
Cynarctoides luskensis
Cynarctoides gawanae
Cynarctoides accridens
Phlaocyon minor
Phlaocyon latidens
Phlaocyon annectens
Phlaocyon marshlandensis
Phlaocyon leucosteus
Cormocyon haydeni
Cormocyon copei
Desmocyon thomsoni
Desmocyon matthewi
Paracynarctus kelloggi
Cynarctus galushai
Cynarctus saxatilis
Cynarctus voorhiesi
Cynarctus crucidens
Metatomarctus canavus
Euoplocyon brachygnathus
Psalidocyon marianae
Microtomarctus conferta
Protomarctus opatus
Tomarctus hippophaga
Tomarctus brevirostris
Aelurodon asthenostylus
Aelurodon mcgrewi
Aelurodon stirtoni
Aelurodon ferox
Aelurodon taxoides
Paratomarctus temerarius
Paratomarctus euthos
Carpocyon compressus
Carpocyon webbi
Carpocyon robustus
Protepicyon raki
Epicyon saevus
Epicyon haydeni
Borophagus pugnator
Borophagus orc
Borophagus secundus
Borophagus hilli
Borophagus diversidens
Leptocyon vulpinus
Leptocyon gregorii
Leptocyon leidyi
Leptocyon vafer
Vulpes stenognathus
Metalopex macconnelli
Metalopex merriami
Urocyon galushi
Urocyon citronus
Urocyon cinereoargenteus
Cerdocyon texanus
Eucyon davisi
Canis ferox
Canis thooides
Canis lepophagus
Canis edwardii
Canis latrans
Canis armbrusteri
Canis dirus
Canis lupus
Xenocyon texanus
Cuon javanicus
40 30 20 10 0
millions of years ago

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 $\alpha$-parameter of the OU model. However, the MLE of $\alpha$ 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

<table>
<thead>
<tr>
<th>Model</th>
<th>Q_{12}</th>
<th>Q_{13}</th>
<th>Q_{21}</th>
<th>Q_{23}</th>
<th>Q_{31}</th>
<th>Q_{32}</th>
<th>LnLk</th>
<th>AICc</th>
<th>AIcW</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>0.0247</td>
<td>0.0247</td>
<td>0.0247</td>
<td>0.0247</td>
<td>0.0247</td>
<td>0.0247</td>
<td>−78.98</td>
<td>159.96</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>SYMM</td>
<td>0.0412</td>
<td>0.0000</td>
<td>0.0412</td>
<td>0.0478</td>
<td>0.0000</td>
<td>0.0478</td>
<td>−72.07</td>
<td>150.14</td>
<td>0.69</td>
</tr>
<tr>
<td>ARD</td>
<td>0.0139</td>
<td>0.0000</td>
<td>0.0717</td>
<td>0.0292</td>
<td>0.0000</td>
<td>0.0466</td>
<td>−69.38</td>
<td>150.77</td>
<td>0.30</td>
</tr>
</tbody>
</table>

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

Table S2. Model-averaged transition rates between dietary states

<table>
<thead>
<tr>
<th>State</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>—</td>
<td>0.03</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>—</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.05</td>
<td>—</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Trait</th>
<th>LD1</th>
<th>LD2</th>
<th>LD3</th>
<th>r^2</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBL</td>
<td>−45.49</td>
<td>1.35</td>
<td>−31.12</td>
<td>0.88</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>M1BS</td>
<td>1.48</td>
<td>1.70</td>
<td>79.09</td>
<td>0.86</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>M2S</td>
<td>−3.32</td>
<td>12.75</td>
<td>−4.74</td>
<td>0.93</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MAT</td>
<td>24.71</td>
<td>−8.26</td>
<td>13.08</td>
<td>0.36</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>p4S</td>
<td>−9.92</td>
<td>−15.26</td>
<td>−6.59</td>
<td>0.69</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

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.