

Longer in the tooth, shorter in the record? The evolutionary correlates of hypsodonty in Neogene ruminants

P. Raia^{1,*}, F. Carotenuto¹, J. T. Eronen² and M. Fortelius^{2,3}

¹*Dipartimento di Scienze della Terra, Università degli studi Federico II, Largo San Marcellino 10, 80138 Napoli, Italy*

²*Department of Geosciences and Geography, University of Helsinki, PO Box 64, 00014 Helsinki, Finland*

³*Institute of Biotechnology, University of Helsinki, PO Box 56, 00014 Helsinki, Finland*

The acquisition of hypsodont molars is often regarded as a key innovation in the history of ruminant ungulates. Hypsodont ruminants diversified rapidly during the later Neogene, circa 15–2 Myr ago, and came to dominate the ruminant fossil record in terms of species diversity. Here we show that hypsodont clades had higher speciation and diversification rates than other clades. Hypsodont species had, on average, shorter stratigraphic durations, smaller range size and lower occupancy than non-hypsodont species. Within hypsodont clades, some species were very common and acquired large geographical ranges, whereas others were quite rare and geographically limited. We argue that hypsodont clades diversified in an adaptive radiation-like fashion, with species often splitting cladogenetically while still in the expansive phase of their occupancy history.

Keywords: key innovation; hypsodonty; ruminants; Neogene Old-World database; species duration

1. INTRODUCTION

Key innovations are often granted a pivotal role in evolution [1–3]. They are thought to spark bursts of diversification and enable entry into new adaptive zones and extension of geographical range [4–8]. Hypsodonty, or increased tooth crown height, has been recognized as such an innovation in the evolution of the mammalian dentition, enabling especially herbivores to process foods that are abrasive or otherwise conducive to increased dental wear [9,10]. Hypsodont ruminants rapidly diversified during the later Neogene, circa 15–2 Myr ago, in a context of climatic cooling, increasing aridity and expansion of open habitats, including grasslands [11–15]. Although geologically younger than other ruminants, hypsodont species make up the majority of the ruminant fauna today [13,15].

The diversification rate (the rate of increase in species number through time) depends on the difference between speciation and extinction rates. A given phenotypic trait could affect diversification through its effect on either factor. As a key innovation, hypsodonty might have promoted speciation via adaptive radiation. Alternatively, it might have lowered extinction rate if hypsodonts were better ecological competitors. Investigating the ecological circumstances under which hypsodont lineages thrived would then indicate the mechanism for their current dominance. For instance, hypsodont species are expected to be larger sized than non-hypsodont taxa. This is because the resource that hypsodonty makes available to herbivores is a physiologically demanding one. The

symbiotic digestion of ‘fibrous’ foods that consist mainly of cellulose and other polysaccharides requires a slowed-down passage of the ingesta. Thus, grazing ruminants are more intake-limited than browsing ones, and compensate by increasing their body size in order to lower mass-specific energy requirements [16–18]. The relationship between hypsodonty and population size is more ambiguous. If hypsodonty did confer any significant competitive advantage over species with low-crowned molars, hypsodont species are expected to have had larger geographical range size and/or locality occupancy [4,11]. This would also agree with the finding that hypsodont species were shown to have larger niche breadths [19], meaning that they may subsist on graze, browse or any combination of both, whereas low-crowned species are restricted to browse. Alternatively, rapid diversification in hypsodont lineages might have promoted secondary dietary specialization, leading to smaller ranges [9,20].

Our aim here is to test whether higher speciation rate, lower extinction rate, or both were the cause of the rapid diversification of hypsodont lineages. For this, we built a 244 species-level tree of ruminants, including 213 extinct and 31 extant Eurasian species. We first estimated speciation and extinction rates in both hypsodont and non-hypsodont species, and their course through time. Then, we measured the evolutionary and ecological correlates of hypsodonty. We used material from the NOW database (see [21] and the electronic supplementary material) and applied both linear regression models and phylogenetic generalized least-squares (PGLS) regressions of hypsodonty classes (hypsodont, mesodont or brachyodont) on species duration in the fossil record (in millions of years, Myr). We also retrieved from the NOW database information on species body size. Based on the occurrence data, we computed

* Author for correspondence (pasquale.raia@libero.it).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2011.0273> or via <http://rspb.royalsocietypublishing.org>.

occupancy [11,22–25] and geographical range sizes for each species. For PGLS, we applied both the Brownian motion model of evolution, and Pagel's λ transform of the original branch lengths [26,27]. Pagel's λ provides the best fit of the Brownian motion to the data by means of a maximum likelihood approach [28]. For geographical range size, we computed two different estimates. One is the minimum convex polygon (MCP) including all the fossil sites where a species occurs. The other (k local convex hull (k -LOCOH)) is based on the density distribution of the fossil sites [29,30], and is restricted to common species only, as it is not advisable to apply this method to taxa with less than 10 locality occurrences. In the present context, 'common' thus refers to taxa with 10 or more locality occurrences. Stratigraphic ranges were computed as the difference between the oldest and the youngest occurrence of the species in the fossil record, after localities with uncertain age were excluded by means of spectral ordering [31]. Species commonness (locality occupancy) was estimated from the fossil record as the relative proportion of sites where a species is present divided by the number of sites included in the species' stratigraphic range.

Our results shed light on the nature and processes whereby key innovations give rise to evolutionary success.

2. MATERIAL AND METHODS

(a) Rates of speciation and extinction

The extended discussion of our ruminant phylogeny, including a presentation of the tree itself in Newick format is available in the electronic supplementary material.

Different methods have been developed to compute character-associated diversification rates [7,32–34]. Among them, the binary-state speciation and extinction (BiSSE) method is the only one that computes extinction rate, and does not assume that ancestral character states reconstruction is independent from the character influence on speciation and extinction themselves [34].

BiSSE computes the probability of a phylogenetic tree and of the observed distribution of a binary character state (here hypsodont or non-hypsodont molars) among the tree tips, given a model of character evolution, speciation and extinction. BiSSE uses six parameters, speciation rate for states 0 (λ_0) and 1 (λ_1), the corresponding extinction rates μ_0 and μ_1 and the transition rates from one character state to the other, q_{01} and q_{10} . Here state 0 corresponds to the non-hypsodont and 1 corresponds to the hypsodont conditions, respectively. The maximum-likelihood estimates of the parameters are computed. It is possible to fix parameters to be equal to each other by setting constraints (e.g. by setting $\lambda_0 = \lambda_1$) so that alternative models of evolution can be tested and then compared with each other by means of likelihood and Akaike information criterion (AIC) values assessment. When compared with other methods, BiSSE correctly computes transition rates from one character state to the other, thereby avoiding biased estimation of speciation and extinction rates based on incorrect transition frequency estimation [35].

We tested two alternative models. First, we fitted a 'free' model where all the six parameters are estimated. Then, we fitted a second model where $\lambda_0 = \lambda_1$ and $\mu_0 = \mu_1$. Since the diversification rate is $\lambda - \mu$, our second model effectively assumes that diversification rates are character independent. In its initial formulation, the BiSSE method [34] assumed

the phylogeny to be dichotomous and complete (i.e. all extant species of a group and their character states are included in the phylogeny). A later development extended the method to incomplete and unresolved phylogenies [36]. This implementation was applied to our data by using the package *diversitree* in R. We first divided the fossil record into 10, 2 Myr long intervals. We then extracted the species present in each interval and then pruned the phylogeny to an interval phylogeny. The interval phylogenies were used to compute the six parameters in the free model, and the four parameters pertaining to the constrained model. The two were then compared with each other by means of ANOVA, for each interval. Our procedure computes speciation, extinction and transition rates per interval and per character state, so that the course of these parameters through time could be assessed.

(b) Species data

We performed our tests on Neogene Old-World ruminants of Eurasia. We used the NOW database (<http://www.helsinki.fi/science/nov/>) [21] to retrieve information on species occurrence, hypsodonty and body size. Based on these, we computed occupancy (= relative commonness in the fossil record [22]), geographical range size and duration for each species.

We tested the relationships between hypsodonty and each of the variables both as they are, and in a phylogenetically explicit context by using PGLS regression. Body-size data in the NOW database is based on literature and on measured specimen-specific estimates based on equations in Damuth & McFadden [37]. We used the data as it was recorded in the database.

Species duration (= age span) was calculated as the time interval intervening between a species first and last appearances in the fossil record. Fossil localities with unknown or imprecise dating may severely affect the estimation of species duration. To get rid of this problem, we first performed the spectral ordering of the NOW localities considered in this study, including both non-ruminant and ruminant species. Spectral ordering is a mathematical ordination (in time) of fossil localities species lists computed according to their faunal similarity [31]. By means of spectral ordering, localities were given a score, which we regressed against the known age of that subset of localities having an absolute dating. We then removed those localities whose NOW age estimate falls outside the 95% confidence interval of the calculated age (see [38] for the full explanation of this procedure).

To estimate species occupancy, we first divided our longer than 18 Myr long fossil record into 19 time bins each of 1 Myr long. Then, for each species, we calculated the total number of occurrences, and divided it for the total number of localities present in the time bins where the species occurs. This procedure effectively removes sampling inequality [39,40], which is present in our data (the total number of localities per time bin varies between 5 and 221; the total number of species occurrences varies between 12 and 874).

Hypsodonty is usually calculated as an index of crown height [41] standardized to some measure of tooth's breadth at the base of the crown. Although meaningful, the hypsodonty index values are not available for a majority of species. In order to maximize the sample size, we used an ordinal index of hypsodonty, classifying brachydont species as 1, mesodont species as 2 and hypsodont species as 3. We used the classification recorded in the NOW database.

The criteria for assigning species to these classes are ultimately up to the taxonomic coordinators of the NOW advisory board, but the rule of thumb is based on the ratio of height to length of the second molar (upper or lower, either). Brachydont teeth have a ratio of less than 0.8, mesodont teeth a ratio of 0.8–1.2 and hypsodont teeth a ratio greater than 1.2. A potential problem with this classification is that it is not an absolute scale of reference values, since the degree of hypsodonty is not directly comparable between clades. As such, it is exposed to potential misclassifications and is subjective to some extent, especially so for species with no living relatives. Nonetheless, ruminant molars are structurally very conservative, hence the scope for serious misclassification is probably very limited.

We calculated species range sizes by using two different procedures: the MCP and the recently developed fixed *k*-LOCOH method [29]. The former method calculates the smallest convex polygon enclosing all the fossil localities where a species is present. We included all of the species with a number of occurrences greater than 2. The MCP of each species was computed by using ESRI ARCGIS. MCP may provide an underestimation of actual range sizes because of the incompleteness of the fossil record. On the other hand, some overestimation may be hidden in the data if (i) the species range moved over time because of dispersal, or (ii) the MCP includes non-inhabited areas.

The impact of these potential biases to our dataset is uncertain. Therefore, we used *k*-LOCOH [29] to obtain a different estimate of the extent of the territory occupied by a given species. *k*-LOCOH uses the density distribution of the localities where a species occurs. In order to compute *k*-LOCOH, different local convex hulls are drawn, each one associated with a single locality and its $k - 1$ nearest neighbours, where k is an appropriate number of neighbouring localities chosen by the ‘minimum spurious hole covering’ method, as suggested in Getz *et al.* [30]. Then, all the hulls are collated from the smallest to the largest and new polygons (the isopleths) are built by the union of the computed local hulls until encompassing an *a priori* established percentage of the total number of occurrences. The isopleth including 95 per cent of the localities where a species occurs is considered here to represent the actual range size of the species.

k-LOCOH was computed via the R package *Adehabitat* [42] and the related script ‘NNCH’ [30]. *k*-LOCOH gives the benefit of letting geographically outlying localities out of a species’ range computation, thereby diminishing the risk of including areas occupied by range size movement or even by misclassification errors. A potential drawback of the method is that area estimates cannot be produced for rare species, given its computation requires a larger minimum number of localities (here $n_{\min} = 10$) than the MCP. For both methods, non-terrestrial stretches of area were subtracted by the computed polygons.

We collected a total of 244 data points for hypsodonty, species duration and occupancy, 187 data points for MCP geographical range size, 121 data points for *k*-LOCOH geographical range size and 127 data points for body mass (electronic supplementary material, appendix S1).

(c) PGLS regressions

Species phenotypic traits are not independent data points in comparative studies. This happens to be the case because of the effect of shared ancestry on trait diversification [43,44].

A phylogenetically informed comparative analysis uses some measure of *a priori*-defined phenotypic trait covariation based on a phylogenetic tree to account for phylogenetic effects [45]. In PGLS, the residual error in the dependent variable is assumed to be the variance–covariance matrix drawn from the species tree [45]. Diagonal entries in this matrix are the distance in branch lengths (here in Myr) from each species to the tree root. Off-diagonal entries are the sum of branch lengths from the tree root to the most common ancestor for each pair of species. By assuming the most common evolutionary model (Brownian motion), the phenotypic similarity between each pair of species is expected to be proportional to their off-diagonal value (= their covariance). In PGLS regression, the variance–covariance matrix is used to adjust the regression parameters in order to maximize the fit to the Brownian motion model via generalized least squares. Since our independent variable is ordinal, dummy variables were used to fit both linear and PGLS regressions. This was done by using the packages *nlme* and *ape* in R.

We first regressed hypsodonty versus (i) body mass, (ii) geographical range size (both estimates), (iii) occupancy, and (iv) species duration by applying linear models. Then, we replicated regressions in a phylogenetic context via PGLS, by assuming the Brownian motion model of evolution, which assumes a trait to evolve according to a constant-variance random walk. Trait evolution can proceed differently from the Brownian motion. Thus, we repeated each PGLS regression by using a transformed variance–covariance matrix by using Pagel’s λ [27]. Pagel’s λ multiplies the off-diagonal elements of the variance–covariance matrix to a value λ in order to provide the best fit of the Brownian motion to the data by means of a maximum-likelihood approach [28]. When λ is 0, no phylogenetic signal is present in the data. Conversely, $\lambda = 1$ implies Brownian motion. The upper limit of λ is defined by the tree height, because off-diagonal elements of the variance–covariance matrix cannot be larger than the variance [28].

A key innovation is inherently phylogenetic in nature, because the advantage it confers is highly correlated with phylogeny as far as the key trait is itself heritable. In such a situation, PGLS is conservative because it regresses out the phylogenetic component of the regression, which is the relationship between the key trait itself and the advantage it confers. With this caveat in mind, we also calculated the phylogenetic signal for each of the five variables considered by using the *K* statistic [46]. *K* measures the tendency for related species to resemble each other more than expected by chance given the species included in the tree under the Brownian motion model of evolution.

3. RESULTS AND DISCUSSION

As expected from the known rise of hypsodont forms to dominance [11], the diversification rate in hypsodont taxa became higher than in non-hypsodont taxa during the later Neogene, the rate in hypsodont ones surpassing the non-hypsodont rate by about 10–12 Myr (table 1). The takeover was mainly dependent on the much higher speciation rate in hypsodont forms, whereas extinction rates are almost comparable between the two character states throughout our greater than 18 Myr long study period (figure 1). The two BiSSE models we tested perform equally well for the five older time bins, whereas

Table 1. Origination, extinction and diversification rate comparisons of hypsodont versus non-hypsodont taxa per time bin. (λ_0 , speciation rate for state '0' (=non-hypsodont); λ_1 , speciation rate for state '1' (=hypsodont); μ_0 , extinction rate for state '0' (=non-hypsodont); μ_1 , extinction rate for state '1' (=hypsodont); $\ln\text{Lik.free}$, likelihood for the free model (parameters are fitted independently from each other); $\ln\text{Lik.one.rate}$, likelihood for the constrained model (diversification rate is held the same for both character states); AIC.free, Akaike information criterion for the free model (parameters are fitted independently from each other); AIC.one.rate, information criterion for the constrained model (diversification rate is held the same for both character states); p , probability that the two models differ; R_0 , diversification rate for state '0' (=non-hypsodont); R_1 , diversification rate for state '1' (=hypsodont).

interval (Myr ago)	λ_0	λ_1	μ_0	μ_1	$\ln\text{Lik.free}$	$\ln\text{Lik.one.rate}$	AIC.free	AIC.one.rate	p	R_0	R_1
0–2	0.047	0.122	3.94E–07	3.27E–07	–273.551	–282.234	559.102	572.469	1.69E–04	0.047	0.122
2–4	0.049	0.129	7.11E–08	8.23E–08	–160.468	–164.368	332.937	336.736	0.020	0.049	0.129
4–6	0.054	0.177	1.72E–05	9.63E–02	–133.094	–137.161	278.189	282.323	0.017	0.054	0.081
6–8	0.003	0.157	1.13E–06	4.51E–06	–248.736	–263.839	509.473	535.678	2.76E–07	0.003	0.157
8–10	0	0.140	3.58E–06	2.60E–08	–284.718	–298.983	581.435	605.966	6.38E–07	–2.83E–06	0.140
10–12	0.051	0.082	2.34E–07	2.20E–02	–474.653	–475.581	961.306	959.163	0.395	0.051	0.060
12–14	0.090	0	3.22E–07	1.50E–06	–220.435	–218.756	452.870	445.513	0.187	0.090	8.65E–07
14–16	0.114	0.028	7.12E–07	1.22E–07	–69.282	–69.882	150.565	147.765	0.549	0.114	0.028
16–18	0.164	1.20E–07	0.112	0.017	–45.693	–46.051	103.386	100.103	0.699	0.053	–0.017

the model with free parameters performs much better for the five younger bins, from 10 Ma onwards (table 1). According to the free parameters model, speciation rate associated with hypsodonty is much higher than with non-hypsodonty (i.e. $\lambda_1 > \lambda_0$) during the last 10 Myr. No corresponding difference in extinction rates can be seen, however (table 1). These results indicate that hypsodonty conferred significantly higher speciation rate, but not higher survival to high-crowned ruminant species.

The relationship between hypsodonty and species duration in the fossil record is negative and significant ($p = 0.049$; table 2). This still applies when phylogenetic effects are taken into account ($p_{\text{Brownian}} = 0.263$, $p_{\text{Pagel}} = 0.049$ (notice that the Pagel transform works better basing on its much lower AIC score; table 3)). Considering extinct species only, duration is significantly different between hypsodonty classes (ANOVA, $F_2 = 7.486$, $p = 0.001$). Hypsodont species have significantly shorter durations than non-hypsodont (browsers plus grazers) ones (Tukey HSD: hypsodont species duration = 1.05 Myr, grouped comparison $p = 0.336$). The negative relationship between hypsodonty and species duration is consistent with the 'key innovation' hypothesis, in so far as a key innovation is thought to promote diversification via increased speciation rate [5–7].

If hypsodonty conferred any ecological advantage, hypsodont species might be expected to have had larger range size and higher occupancy than low-crowned species. Instead, we found evidence that the relationships between hypsodonty and both of these variables are in fact negative. When MCP range size estimates are used, linear regression models indicate that hypsodonty is marginally associated with range size ($p = 0.078$). The slope of the regression is negative (-0.196). The same negative relationship applies with PGLS regression ($p_{\text{Brownian}} = 0.003$, $p_{\text{Pagel}} = 0.007$; table 3). Similarly, occupancy is negatively associated with hypsodonty class (slope = -0.005 , $p = 0.010$; $p_{\text{Brownian}} = 0.480$, $p_{\text{Pagel}} = 0.049$; again Pagel's transform provides a lower AIC score). This would suggest that hypsodont taxa belong to clades whose species have relatively small geographical ranges (hence higher habitat specialization) [20]. This interpretation agrees with Simpson's original concept of hypsodonty as a key innovation driving adaptive evolution and specialization [9].

However, when k-LOCOH range size estimates are used, the sign of the relationship between range size and hypsodonty class is reversed. This applies both to raw data (slope = 0.225 ; $p = 0.072$), and under PGLS regression ($p_{\text{Brownian}} < 0.001$, $p_{\text{Pagel}} = 0.018$; table 3). This is not an artefact of how the ranges are calculated, because the correlation between the two range size estimates is strong and positive ($r_{119} = 0.880$, $p < 0.001$). Furthermore, when we repeated the PGLS regression on MCP data cut to the same species for which we had k-LOCOH range size estimates, the relationship between MCP and hypsodonty becomes significant and positive both under Brownian motion (slope = 0.256 , $p < 0.001$, AIC = 354.09) and by using Pagel's λ transform (slope = 0.177 , $p < 0.044$, AIC = 286.47). Instead, this applies because k-LOCOH ranges cannot be computed for rare species, whose range size is in fact negatively correlated to hypsodonty (figure 1). Indeed, average occupancy of the taxa with k-LOCOH range size

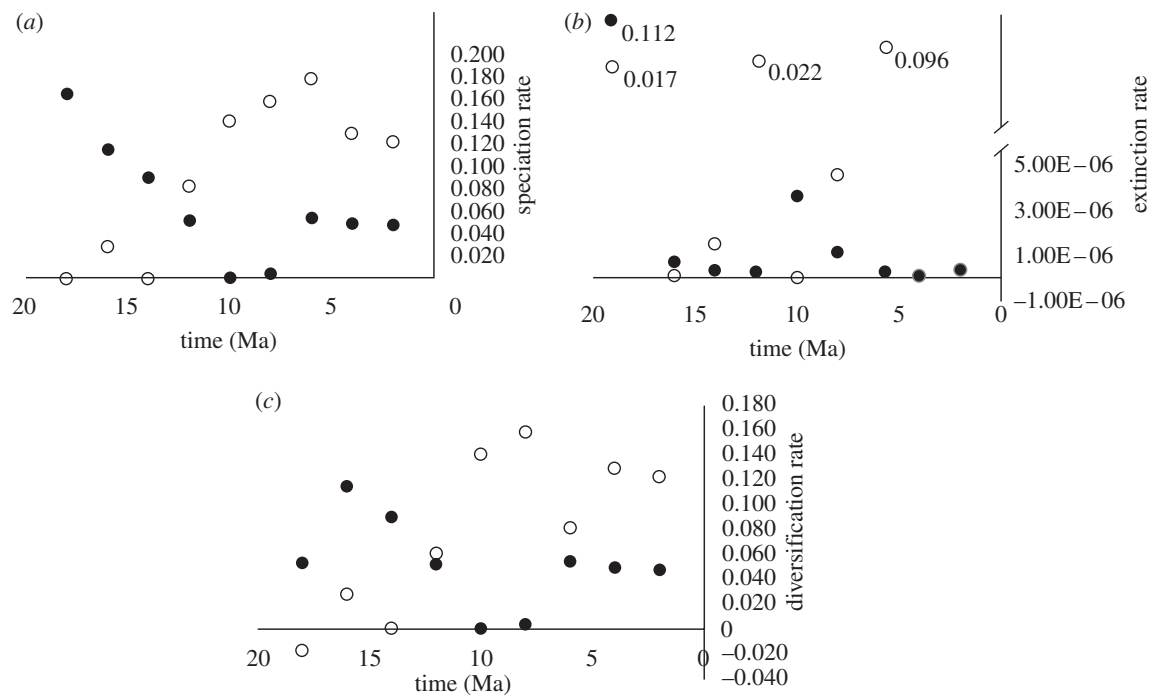


Figure 1. Differences in (a) speciation rate (filled circles, λ_0 ; open circles, λ_1), (b) extinction rate (filled circles, μ_0 ; open circles, μ_1) and (c) diversification rate (filled circles, R_0 ; open circles, R_1), through time, in hypsodont versus non-hypsodont species.

Table 2. Linear model regression statistics between hypsodontology and ecological variables. (Data for body size, range size and occupancy were log-transformed to reassure normality.)

	body size	range size (MCP)	range size (k-LOCOH)	duration	occupancy
slope	0.610	-0.196	0.225	-0.158	-0.005
r^2	0.118	0.011	0.019	0.014	0.010
p	<0.001	0.078	0.072	0.047	0.064
d.f.	108	185	119	211	242

Table 3. PGLS regression statistics between hypsodontology and ecological variables. (Data for body size, range size and occupancy were log-transformed to reassure normality. Both the Brownian motion model and λ transform were fitted to the data to perform PGLS. For each relationship and model, we report Akaike information criterion (AIC).)

	slope	AIC	p	d.f.
Brownian motion				
body size	0.616	289.553	<0.001	108
range size (MCP)	-0.497	679.931	0.003	185
range size (k-LOCOH)	0.318	390.069	<0.001	119
duration	-0.170	811.773	0.263	211
commonness	-0.002	-935.787	0.480	242
Pagel's λ transform				
body size	0.643	287.569	<0.001	108
range size (MCP)	-0.199	613.381	0.007	185
range size (k-LOCOH)	0.259	340.972	0.018	119
duration	-0.171	579.649	0.049	211
commonness	-0.005	-959.230	0.049	242

estimates (0.0396) is about five times higher than in species without k-LOCOH estimates (0.0076). This points to a substantial dichotomy between rare and common species in their range size/hypsodontology relationship. In keeping with this notion, Jernvall & Fortelius [11] found that the most common taxa were almost entirely responsible for the increase in hypsodontology during the spread of open habitats some 11–5 Myr ago.

A possible caveat could be that if ranges shift over species existence, longer lived species may have spuriously larger ranges. Since hypsodont species are shorter lived, they may have spuriously shorter ranges. For this caveat to apply, then, range size and species duration should be correlated. Yet, we found this is not the case. For MCP range size estimates, the correlation is not significant ($t_{185} = -1.069$, $p = 0.286$, $r = -0.078$). With k-LOCOH estimates, the correlation is significant, but negative ($t_{119} = -2.54$, $p = 0.012$, $r = -0.226$). Thus, long-living species do not have significantly larger geographical ranges.

The regression of hypsodontology on body size gives a significant and positive slope ($p = 0.001$). Both the sign and significance of this relationship are confirmed under

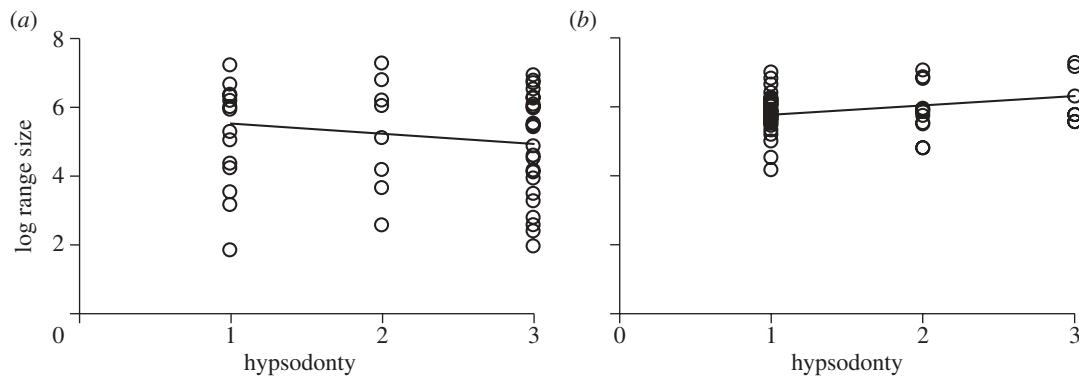


Figure 2. The relationship between hypsodonty class and range size drawn for the species having the (a) 50 least and (b) highest occupancy values.

Table 4. Phylogenetic signal (K statistic). (The p -value refers to the null hypothesis $K = 0$.)

	K	p	n
hypsodonty	0.648	0.001	244
body size	1.492	0.001	127
range size (MCP)	0.450	0.101	187
range size (k-LOCOH)	0.296	0.249	121
duration	0.205	0.179	211
occupancy	0.284	0.030	244

PGLS regression ($p_{\text{Brownian}} < 0.001$, $p_{\text{Pagel}} < 0.001$) (tables 2 and 3). This result is expected from the ruminant digestive physiology [17].

Our results suggest that hypsodonty resulted in increased speciation rate in hypsodont clades. Yet, the evolutionary success accrued by hypsodonty was not ubiquitous within the clades: we estimated the phylogenetic signal by Blomberg *et al.*'s [46] K statistic (table 4 and electronic supplementary material, figure S2). K is significant for hypsodonty, body size and duration but not for estimates of occupancy or geographical range size. This means that closely related species inherited body size, hypsodont molars and, to a lesser extent, a propensity for short duration in the fossil record from their ancestors, but neither commonness nor geographical range size. Since range size generally grows during the existence of a species [22–24,47,48], a possible explanation might be that hypsodont lineages have tended to originate daughter species before their full potential range size was occupied, forming clades containing mostly short-lived species of either very large or very small range sizes. This agrees with the notion that rapidly diversifying lineages should split frequently into peripheral isolates with small range size [49]. It would also account for the negative relationship between hypsodonty and both geographical range size and occupancy, and for the opposite signs of the relationships between hypsodonty and range size when common species are compared with rare species (figure 2). A similar explanation was advanced to account for the great difference in species richness between closely related clades of specialist versus generalist species, for example, Vrba's effect hypothesis explaining the high species richness of alcelaphine bovids relative to the low richness of the impala clade [50].

In a related study, Feranec also found that hypsodonty works as a key innovation by promoting speciation [19]. He contended that the geographical ranges of hypsodont taxa are no larger than those of non-hypsodont taxa, and argued instead that their niche breadths are larger. Feranec's methodology was different and his sample included horses, tapirs, camels and proboscideans besides ruminants, which may account for his different result. Despite the difference, both studies provide evidence that hypsodonty increased diversification rate, as originally proposed by Simpson [9]. In terms of species richness, the net balance shifted in time significantly in favour of hypsodont taxa and most modern ungulate-rich ecosystems are accordingly dominated by hypsodont species.

The big picture that emerges from our study is that hypsodonty is strongly heritable, and that it spawned increased speciation rate. Species within hypsodont clades were large, and very different from each other in terms of commonness and geographical range size. Some were successful and attained large geographical ranges and occupancy, whereas others were geographically restricted. Overall, hypsodont clades clearly had a more dynamic evolutionary history than non-hypsodont clades. As expected for an evolutionary key trait, the acquisition of hypsodonty produced a great number of species, leading to the current dominance of hypsodont over non-hypsodont taxa in the fossil record.

Alan Gentry kindly helped us in correcting the taxonomy of bovids in the tree and gave us important insights as for the phylogenetic position of some clades. David Polly and one anonymous reviewer kindly helped us to make this paper more robust and our conclusions more sound. We are grateful to NOW contributors for their continuous support and work. This study grew out of a study visit by P.R. to Helsinki, supported by a grant (to M.F.) from the Academy of Finland.

REFERENCES

- 1 Erwin, D. H. 1992 A preliminary classification of evolutionary radiations. *Hist. Biol.* **6**, 133–147. (doi:10.1080/10292389209380423)
- 2 Hunter, J. P. 1998 Key innovations and the ecology of macroevolution. *Trends Ecol. Evol.* **13**, 31–36. (doi:10.1016/S0169-5347(97)01273-1)
- 3 Donoghue, M. J. 2005 Key innovations, convergence, and success: macroevolutionary lessons from plant

- phylogeny. *Paleobiology* **31**, 77–93. (doi:10.1666/0094-8373(2005)031[0077:KICASM]2.0.CO;2)
- 4 Heard, S. B. & Hauser, D. L. 1995 Key evolutionary innovations and their ecological mechanisms. *Hist. Biol.* **10**, 151–173. (doi:10.1080/10292389509380518)
 - 5 de Queiroz, A. 2002 Contingent predictability in evolution: key traits and diversification. *Syst. Biol.* **51**, 917–929. (doi:10.1080/10635150290155980)
 - 6 Ricklefs, R. E. 2003 Global diversification rates of passerine birds. *Proc. R. Soc. Lond. B* **270**, 2285–2291. (doi:10.1098/rspb.2003.2489)
 - 7 Ree, R. H. 2005 Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution* **59**, 257–265. (doi:10.1111/j.0014-3820.2005.tb00986.x)
 - 8 Vermeij, G. J. 2006 Historical contingency and the purported uniqueness of evolutionary innovations. *Proc. Natl Acad. Sci. USA* **103**, 1804–1809. (doi:10.1073/pnas.0508724103)
 - 9 Simpson, G. G. 1953 *The major features of evolution*. New York, NY: Columbia University Press. (doi:10.1126/science.119.3098.699-a).
 - 10 MacFadden, B. J. 2000 Cenozoic mammalian faunas from the Americas: reconstructing ancient diets and terrestrial communities. *Ann. Rev. Ecol. Syst.* **31**, 33–59. (doi:10.1146/annurev.ecolsys.31.1.33)
 - 11 Jernvall, J. & Fortelius, M. 2002 Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature* **417**, 538–540. (doi:10.1038/417538a)
 - 12 Bouchenak-Khelladi, Y., Verboom, G. A., Hodkinson, T. R., Salamin, N., Francois, O., Chonghaile, G. N. & Savolainen, V. 2009 The origins and diversification of C4 grasses and savanna-adapted ungulates. *Global Change Biol.* **15**, 2397–2417. (doi:10.1111/j.1365-2486.2009.01860.x)
 - 13 Janis, C. M. 1989 A climatic explanation for patterns of evolutionary diversity in ungulate mammals. *Palaeontology* **32**, 463–481.
 - 14 Eronen, J. T., Evans, A. R., Fortelius, M. & Jernvall, J. 2010 The impact of regional climate on the evolution of mammals: a case study using fossil horses. *Evolution* **64**, 398–408. (doi:10.1111/j.1558-5646.2009.00830.x)
 - 15 Janis, C. M. 2008 An evolutionary history of browsing and grazing ungulates. In *The ecology of browsing and grazing* (eds I. J. Gordon & H. H. T. Prins), pp. 21–45. Berlin, Germany: Springer.
 - 16 Demment, M. W. & Van Soest, P. J. 1985 A nutritional explanation for body-size patterns of ruminant and non-ruminant herbivores. *Am. Nat.* **125**, 641–672. (doi:10.1086/284369)
 - 17 Clauss, M., Frey, R., Kiefer, B., Lechner-Doll, M., Loehlein, W., Polster, C., Rössner, G. E. & Streich, W. J. 2003 The maximum attainable body size of herbivorous mammals: morphophysiological constraints on foregut, and adaptations of hindgut fermenters. *Oecologia* **136**, 14–27. (doi:10.1007/s00442-003-1254-z)
 - 18 Clauss, M. & Hummel, J. 2005 The digestive performance of mammalian herbivores: why big may not be that much better. *Mammal. Rev.* **35**, 174–187. (doi:10.1111/j.1365-2907.2005.00062.x)
 - 19 Feranec, R. S. 2007 Ecological generalization during adaptive radiation: evidence from Neogene mammals. *Evol. Ecol. Res.* **9**, 555–577.
 - 20 Schluter, D. 2000 *The ecology of adaptive radiation*. New York, NY: Oxford University Press.
 - 21 Fortelius, M. 2010 Neogene of the Old World Database of fossil mammals (NOW). University of Helsinki (coordinator). See <http://www.helsinki.fi/science/now/>.
 - 22 Jernvall, J. & Fortelius, M. 2004 Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *Am. Nat.* **164**, 614–624. (doi:10.1086/424967)
 - 23 Carotenuto, F., Barbera, C. & Raia, P. 2010 Occupancy, range size and phylogeny in Eurasian Pliocene to Recent large mammals. *Paleobiology* **36**, 399–414. (doi:10.1666/09059.1)
 - 24 Raia, P., Meloro, C., Loy, A. & Barbera, C. 2006 Species occupancy and its course in the past: macroecological patterns in extinct communities. *Evol. Ecol. Res.* **8**, 181–194.
 - 25 Lyons, S. K. & Smith, F. A. 2010 Using macroecological approach to study geographic range, abundance and body size in the fossil record. In *Quantitative methods for studying paleobiology* (eds J. Alroy & G. Hunt), pp. 117–141. Pittsburgh, PA: The Paleontological Society.
 - 26 Lavin, S. R., Karasov, W. H., Ives, A. R., Middleton, K. M. & Garland, T. 2008 Morphometrics of the avian small intestine compared with that of non flying mammals: a phylogenetic approach. *Physiol. Biochem. Zool.* **81**, 526–550. (doi:10.1086/590395)
 - 27 Pagel, M. 1999 Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884. (doi:10.1038/44766)
 - 28 Freckleton, R. P., Harvey, P. H. & Pagel, M. 2002 Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **60**, 712–726. (doi:10.1086/343873)
 - 29 Getz, W. M. & Wilmers, C. C. 2004 A local nearest neighbor convex-hull construction of home ranges and utilization distributions. *Ecography* **27**, 489–505. (doi:10.1111/j.0906-7590.2004.03835.x)
 - 30 Getz, W. M., Fortmann-Roe, S., Cross, P. C., Lyons, A. J., Ryan, S. J. & Wilmers, C. C. 2007 LoCoH: non-parametric kernel methods for constructing home ranges and utilization distributions. *PLoS ONE* **2**, e207. (doi:10.1371/journal.pone.0000207)
 - 31 Fortelius, M., Gionis, A., Jernvall, J. & Mannila, H. 2006 Spectral ordering and biochronology of European fossil mammals. *Paleobiology* **32**, 206–214. (doi:10.1666/04087.1)
 - 32 Pagel, M. 1997 Inferring evolutionary processes from phylogenies. *Zool. Scr.* **26**, 331–348. (doi:10.1111/j.1463-6409.1997.tb00423.x)
 - 33 Paradis, E. 2005 Statistical analysis of diversification with species traits. *Evolution* **59**, 1–12. (doi:10.1111/j.0014-3820.2005.tb00889.x)
 - 34 Maddison, W. P., Midford, P. E. & Otto, S. P. 2007 Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701–710. (doi:10.1080/10635150701607033)
 - 35 Goldberg, E. E. & Igić, B. 2008 On phylogenetic tests of irreversible evolution. *Evolution* **62**, 2727–2741. (doi:10.1111/j.1558-5646.2008.00505.x)
 - 36 FitzJohn, R. G., Maddison, W. P. & Otto, S. P. 2009 Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.* **58**, 595–611. (doi:10.1093/sysbio/syp067)
 - 37 Damuth, J. & McFadden, B. J. 1990 *Body size in mammalian paleobiology: estimation and biological implications*. New York, NY: Cambridge University Press.
 - 38 Raia, P., Carotenuto, F., Meloro, C., Piras, P., Barbera, C. & Kotsakis, T. 2009 More than three million years of community evolution. The temporal and geographical resolution of the Plio-Pleistocene Western Eurasia mammal faunas. *Palaeog. Palaeoclim. Palaeoecol.* **276**, 15–23. (doi:10.1016/j.palaeo.2009.02.005)
 - 39 Barnosky, A. D., Carrasco, M. A. & Davis, E. B. 2005 The impact of the species-area relationship on estimates of paleodiversity. *PLoS Biol.* **3**, e266. (doi:10.1371/journal.pbio.0030266)

- 40 Raia, P., Carotenuto, F., Meloro, C., Piras, P. & Barbera, C. 2011 Species accumulation over space and time in European Plio-Holocene mammals. *Evol. Ecol.* **25**, 171–188. (doi:10.1007/s10682-010-9392-3)
- 41 Van Valen, L. 1960 A functional index of hypsodonty. *Evolution* **14**, 531–532. (doi:10.2307/2406003)
- 42 Calenge, C. 2006 The package adehabitat for the R software: a tool for the analysis of space and habitat use by animals. *Ecol. Model.* **197**, 516–519. (doi:10.1016/j.ecolmodel.2006.03.017)
- 43 Blomberg, S. P. & Garland, T. 2002 Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. *J. Evol. Biol.* **15**, 899–909. (doi:10.1046/j.1420-9101.2002.00472.x)
- 44 Felsenstein, J. 1985 Phylogenies and the comparative method. *Am. Nat.* **125**, 1–15. (doi:10.1086/284325)
- 45 Lavin, S. R., Karasov, W. H., Ives, A. R., Middleton, K. M. & Garland, T. 2008 Morphometrics of the avian small intestine compared with that of nonflying mammals: a phylogenetic approach. *Physiol. Biochem. Zool.* **81**, 526–550. (doi:10.1086/590395)
- 46 Blomberg, S. P., Garland, T. & Ives, A. R. 2003 Testing for phylogenetic signal in comparative data: behavioural traits are more labile. *Evolution* **47**, 717–745. (doi:10.1111/j.0014-3820.2003.tb00285.x)
- 47 Foote, M. 2007 Symmetric waxing and waning of marine invertebrate genera. *Paleobiology* **33**, 517–529. (doi:10.1666/06084.1)
- 48 Foote, M., Crampton, J. S., Beu, A. G., Marshall, B. A., Cooper, R. A. & Matcham, I. 2007 Rise and fall of species occupancy in Cenozoic marine molluscs. *Science* **318**, 1131–1134. (doi:10.1126/science.1146303)
- 49 Barraclough, T. G. & Vogler, A. P. 2000 Detecting the geographical pattern of speciation from species-level phylogenies. *Am. Nat.* **155**, 419–434. (doi:10.1086/303332)
- 50 Vrba, E. S. 1980 Evolution, species and fossils: how does life evolve? *S. Afr. J. Sci.* **76**, 61–84.