Late Cenozoic onset of the latitudinal diversity gradient of North American mammals

Jonathan D. Marcot, David L. Fox, and Spencer R. Niebuhr

*Department of Animal Biology, University of Illinois, Urbana-Champaign, IL 61801; †Department of Earth Sciences, University of Minnesota, Minneapolis, MN 55455; and ‡Polar Geospatial Center, University of Minnesota, St. Paul, MN 55108

Edited by P. David Polly, Indiana University, Bloomington, Indiana, and accepted by Editorial Board Member David Jablonski May 6, 2016 (received for review December 15, 2015)

The decline of species richness from equator to pole, or latitudinal diversity gradient (LDG), is nearly universal among clades of living organisms, yet whether it was such a pervasive pattern in the geologic past remains uncertain. Here, we calculate the strength of the LDG for terrestrial mammals in North America over the past 65 My, using 27,903 fossil occurrences of Cenozoic terrestrial mammals from western North America downloaded from the Paleobiology Database. Accounting for temporal and spatial variation in sampling, the LDG was substantially weaker than it is today for most of the Cenozoic and the robust modern LDG of North American mammals evolved only after the last 4 My. The strength of the LDG correlates negatively with global temperature, suggesting a role of global climate patterns in the establishment and maintenance of the LDG for North American mammals.

Mammalia | latitudinal diversity gradient | Cenozoic | climate | biogeography

The equator-to-pole decline in taxonomic richness, commonly referred to as the latitudinal diversity gradient (LDG) (although other terms, such as the latitudinal richness gradient are also used) is a pervasive biogeographic pattern of nearly every clade of extant organisms for which it has been studied (1–3). Despite its ubiquity and intense study, the ecological and evolutionary mechanisms underlying it remain uncertain, with a large number of potential causes that make similar predictions regarding the distribution of extant organisms (4, 5). The fossil record provides a distinct perspective from which to examine the LDG and can provide insights regarding mechanisms that might be inextricable using only extant taxa (e.g., refs. 6–9). For example, some proposed mechanisms to explain the modern LDG invoke particular climatic conditions that exist today, and the fossil record can be used to evaluate these mechanisms under different climatic conditions in the past (6, 10–15).

Modern North American mammals exhibit a typical LDG, with species richness declining substantially from low to high latitudes (Fig. 1 and refs. 16–20). However, we have shown previously (21) that the LDG for terrestrial North American mammals was absent during the mid-Paleocene Epoch (~63–58 Ma). However, the latitudinal gradient in δ18O of Paleocene mammalian tooth enamel was similar to the modern latitudinal gradients in the oxygen isotopic composition of surface waters (precipitation, rivers, and lakes), which suggests that hydrological and climatic controls on δ18O gradients of surface waters were broadly similar during the Paleocene and today. The lack of a LDG in the Paleocene despite climatic gradients like those of today could have been explained as a lingering result of the K/Pg extinction. In this scenario, the mechanisms underlying the modern LDG might have been operating in the Paleocene as they do today, but at that early stage in the postextinction radiation of mammals (22), sufficient time had not yet elapsed for the modern LDG to manifest at the continental scale. If so, a LDG like the modern should become established subsequent to the Paleocene, when North American mammalian faunas take on a modern taxonomic structure (e.g., refs. 23 and 24), then persist until today.

Alternatively, the lack of a Paleocene LDG might reflect the fact that particular continental-scale aspects of climate (e.g., mean annual temperature, precipitation, seasonality, etc.) that drive the modern LDG (17) were at least qualitatively (e.g., spatially) different during the Paleocene. A recent review of studies of the LDG using the fossil record (6) summarizes the general influence of mean global temperature, showing that richness declines toward the poles only during intervals with relatively low global temperature (i.e., “icehouse worlds”), whereas globally warm intervals (i.e., “greenhouse worlds”) typically are characterized by a flat LDG or peak in taxonomic richness at temperate latitudes. Although the latitudinal gradient of temperature in western North America during the Paleocene was similar to today (21), mean global temperature was higher during the Paleocene than today (25), which might, in part, explain the lack of LDG. Mean global temperature has varied substantially since the Paleocene (25), but its relationship to the strength of the LDG over that interval remains unknown.

In the present study, we expand our perspective on the LDG to the entire Cenozoic history of terrestrial mammals in western North America. The fossil record of mammals in central and western North America is arguably the most densely and continuously sampled fossil record of any terrestrial group at the continental scale over the Cenozoic as a whole (26). As such, it is the most suitable—and probably the only—record with which to test the stability of the LDG of mammals at the continental scale over long expanses geological time. Previous studies of this record found variation in the strength of the relationships between latitude and both community composition and the spatial distribution of richness (i.e., β-diversity), respectively, over the Cenozoic (7). Here, we directly estimate the strength of the LDG in the past and demonstrate its variation over time.

Significance

The number of species declines from the Earth’s equator to the poles—a pattern known as the latitudinal diversity gradient (LDG)—for nearly every group of living organisms in which the LDG has been studied. However, our analyses of fossil occurrences demonstrate that the LDG of terrestrial mammals was substantially weaker than it is today, if not absent entirely, for most of the past 65 My. We also show the strength of the LDG varies with global temperature and is weakest during globally warm periods. We find that the modern LDG became established sometime in only the last 4 My, most likely as a result of declining global temperatures.

Author contributions: J.D.M. and D.L.F. designed research; J.D.M. and D.L.F. performed research; J.D.M. and S.R.N. analyzed data; and J.D.M. and D.L.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. P.P. is a guest editor invited by the Editorial Board.

Data deposition: Scripts and the dataset used for the analyses reported in this paper have been deposited in the Dryad Digital Repository (dx.doi.org/10.5061/dryad.53h6n).

To whom correspondence should be addressed. Email: jmarcot@illinois.edu.

Present address: Department of Geodetic Infrastructure, UNAVCO, Boulder, CO 80301.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1524750113/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1524750113

PNAS | June 28, 2016 | vol. 113 | no. 26 | 7189–7194
To do so, we began by dividing the Cenozoic (2–65.5 Ma) into 32 time intervals of uniform 2-Ma duration. We then calculated the “face-value” pattern of richness across latitudes within each time interval, without accounting for variation in sampling among latitudes or intervals. In each interval, we determined species richness within latitudinal bands and intervals using fossil occurrences from the Paleobiology Database (https://www.paleobiodb.org; see Materials and Methods for details). We regressed richness on the midpoint latitudes of the bands to calculate the slope of the LDG (henceforth referred to as “face-value slope”).

Then, we used a conservative approach to determine and account for the influence of species richness of temporal and geographic variation of sampling (Materials and Methods). Intervals were retained for analysis only if they met two criteria: (i) at least three latitudinal bands had a minimum number (i.e., quota) of occurrences; and (ii) those latitudinal bands spanned at least 10° of latitude (Materials and Methods for full analytical details). Then, in each of these best-sampled time intervals that met these criteria, we used sample-standardization to estimate the relative richness among the included latitudinal bands only (i.e., latitudinal bands with fewer occurrences than the quota for the interval were not included for analysis). We regressed these subsampled richness estimates on the midpoint latitudes of the bands to calculate the slope of the LDG (henceforth referred to as “fossil slope”). We ultimately were able to calculate the fossil slope in 16 of 32 intervals that had sufficient sampling (henceforth referred to as “successful” intervals), including three separate stretches of time: (i) most of the Paleocene (four consecutive intervals), (ii) the Middle through Late Eocene (five consecutive intervals), and (iii) Middle Miocene through the Quaternary (seven consecutive intervals) (Table 1). Although restricting our analysis to just these successful intervals leaves many 2-Ma intervals unanalyzed here, our conservative approach ensures that our conclusions are based only on those intervals for which we could explicitly address sampling variation across latitudes, and the analyzed intervals collectively span the entire Cenozoic.

Results

The slope of the LDG of North American terrestrial mammals varied substantially over the past 65 My (Fig. 2). Time series of face-value and sample-standardized fossil LDG slopes are broadly congruent (Fig. 2). Sample standardization appears to have reduced the overall variation of the estimated fossil LDG slopes across the entire Cenozoic, relative to face value, but has no other obvious pervasive effect. For example, sample standardization resulted in fossil slopes uniformly lower than face value during the Paleocene intervals but greater slopes in all successful Eocene intervals but one. In all successful intervals but one (the earliest Paleocene interval), the median fossil slope (over successful pseudoreplicate analyses) was greater than that of the modern North American mammalian LDG, which is strongly negative (i.e., indicating a decline in species richness from low to high latitudes) (Fig. 1). Furthermore, in all successful intervals but one, the calculated fossil slope was near or greater than 0.0, indicating a flat or reversed LDG throughout most of the Cenozoic.

To quantify the influence of the irregular geographic distribution of fossil collections on our estimates of the LDG, we calculated what the slope of the LDG using modern species would be if our knowledge of the geographic ranges of Recent species were restricted to the geography of fossil sampling within each interval (henceforth referred to as “adjusted modern LDG slope”; see Materials and Methods for details). The time series of adjusted modern LDG slopes indicate that the geographic distribution of fossil sampling does indeed influence the estimate of the LDG slope, but not to a great extent. In particular, the 95% confidence interval (CI) of the adjusted modern slopes were beyond the range of the estimates of the modern slope (see Materials and Methods for details) in only 5 of 16 successful intervals (Table 1) indicating that geographic variation in fossil sampling does not hopelessly obscure the fossil LDG. However, the 95% CI of fossil slopes overlap the range of estimates of the modern slope in only four successful intervals (Fig. 2), further revealing that the terrestrial mammalian LDG substantially differed from the modern LDG for most of the Cenozoic.

In 15 of the 16 successful intervals, the fossil slopes were weaker (i.e., less negative) than the adjusted modern LDG slopes in more than 85% of the successful pseudoreplicate analyses (Table 1 and Materials and Methods). In 14 of those 15 intervals, the median fossil slopes lie beyond the 95% CI of the adjusted modern slopes (when median and CI are calculated over the successful pseudoreplicate analyses). Calculated another way, in each pseudoreplicate analysis, we determined whether the adjusted modern slope fell within the calculated 95% CI of the fossil using the regression of standardized fossil richness on latitude (see Materials and Methods for details). In 12 of the 16 successful intervals, the fossil slope was beyond this 95% CI of the adjusted modern slope in the majority (>50%) of the pseudoreplicate analyses (Table 1).

Discussion

Our results strongly demonstrate that the mammalian LDG of North America significantly differed from the modern LDG throughout most of the Cenozoic. Furthermore, the lack of a LDG previously observed in the mid-Paleocene (21) was not merely a lingering consequence of the K/Pg extinction and subsequent mammalian radiation but rather a persistent feature of mammalian biogeography in North America for most of the Cenozoic. The fossil slope is indistinguishable from the modern (negative) LDG during only three intervals: the earliest Paleocene (63–65 Ma), the late Eocene (36–38 Ma), and the Pliocene–Pleistocene (2–4 Ma). The modern-like LDG slope during the earliest Paleocene interval is surprising given the strong support for a lack of a LDG during the mid-Paleocene shown in this and our previous study (21), which used different approaches than we do here. Our result for the earliest Paleocene is puzzling in terms of the temporal and spatial proximity...
to the K/Pg impact event and associated mass extinction, and the Paleocene as a whole is not an interval of major global or regional climate change. During the late Eocene, the fossil slopes fall within the CI of the adjusted modern slope for one interval, which lies at the end of a gradual global cooling trend that lasted more than 10 My, and culminates in the abrupt climate change associated with the onset of Antarctic glaciation at the Eocene–Oligocene boundary (27). Finally, the fossil and modern slopes become indistinguishable in the most recent interval (2–4 Ma) in the Pliocene and early Pleistocene, suggesting that the modern LDG was established sometime in only the last 4 My.

We do not expect that taphonomic incompleteness would greatly distort the observed latitudinal patterns of taxonomic richness. Undoubtedly, the number of species preserved at a fossil locality is often a subset of those that actually inhabited the area in life. However, we pooled species richness among localities within latitudinal bands, so species that escaped sampling at one locality potentially could be found at another locality within the band. We used our sample-standardization procedure to equalize the number of analyzed occurrences among latitudes, thereby reducing the variation among latitudes of this probability of finding these missing species. Furthermore, we are not aware of any evidence for a persistent, systematic, continental-scale bias in taphonomic incompleteness throughout the Cenozoic that could account for the absent or reversed LDG slopes over the analyzed region in nearly every successful interval in our time series.

We restricted our geographic region of analysis to the most continuously and densely sampled region of western North America between 25°N and 70°N latitudes. There are few pre-Pleistocene collections (and constituent occurrences) beyond these geographic restrictions (e.g., in tropical Central America and eastern North America). Latitudinal bands including these collections beyond our latitudinal bounds would almost always be excluded from our analyses due to their small sample sizes (Materials and Methods). Because we did not consider fossil localities south of 25°N, we cannot rule out the possibility that tropical taxonomic richness was substantially higher than within our midlatitude focal region during the Cenozoic. Of the few fossil mammal localities south of 25°, most have very few (fewer than four) occurrences and/or are younger than the temporal window of this analysis. A few older collections (e.g., Arikareean Galliard Cut Fauna of Panama and the Hemphillian–Blancan Rancho el Ocote and Chadronian Inyoo faunas from Mexico) consist of modest numbers of occurrences. Nevertheless, taxonomic richness of these collections is, at most, comparable with similarly aged faunas further north, and far below that observed for the same latitudes today (17). If there were substantially higher, although unsampled, tropical richness, it could only produce a disjunct latitudinal richness pattern in the Northern Hemisphere, consisting of a tropical peak paired with the observed flat gradient throughout midlatitudes. Such a pattern would still differ substantially from modern pattern, which is characterized by a strong gradient over the entire geographic range analyzed in this study, as demonstrated by the adjusted modern slopes.

Mannion et al. (6) classify hypotheses explaining the modern LDG into geographic, historical, and climatic themes. Geographic hypotheses generally attribute greater richness near the equator to the greater geographic area of the globe that is present at lower latitudes. Such hypotheses are not viable explanations for the modern LDG of terrestrial mammals in North America because land area generally decreases southward from North to Central
America. Furthermore, the area of habitable North America has not changed over the Cenozoic to an extent that could have eliminated or reversed the LDG. Indeed, during the Plio–Pleistocene interval, in which the modern-like LDG appears, the extensive habitable terrestrial area of high-latitude North America would have disappeared episodically during major ice sheet advances.

The hypotheses under the historical theme generally attribute the lower richness at higher latitudes to a greater degree of climatic disruption in the geologic past, which caused frequent ecosystem upheaval, or even complete habitat loss, in the case of widespread high-latitude glaciations. At a global scale, such extreme climatic disruptions are characteristic of only the recent geological past. Northern hemisphere polar glaciations first occurred during the Late Miocene and intensified during the Pliocene (28), culminating in the cyclic Pleistocene glacial advances into central North America that recurred at Milankovitch band periodicities. Such dramatic climatic and environmental modifications coincide with the pronounced decline in LDG slope we observe between the lower richness at higher latitudes to a greater degree of climatic

Therefore, the historical hypotheses cannot account for the variation in LDG slope throughout most of our time series.

The timing of the development of the modern LDG of North American mammals coincides reasonably well with the completion of the Isthmus of Panama and the onset of the Great American Biotic Interchange (GABI), so it is reasonable to consider whether the appearance of the modern LDG was driven by immigration of South American lineages and their subsequent diversification in the context of the preexisting flat latitudinal diversity gradient. However, the fossil record documents only a modest number of immigrations into North America of species with South American origins (e.g., primates, marsupials, bats, xenarthrans, and a few species of rodents) as the isthmus developed during the late Cenozoic (29, 30). Given the standing diversity of endemic North American clades further north throughout the GABI, these immigrants were too few in number to tip a flat gradient into its modern configuration.

Finally, the climatic theme encompasses hypotheses directly or indirectly attributable to the lower seasonality and higher insolation found at lower latitudes. The weaker or even reversed LDG in the geologically recent past suggests that the latitudinal distribution of insolation cannot be its only causal mechanism. More generally, these hypotheses generally appeal to latitudinal temperature gradients, which are obvious today. During intervals of higher global temperature that characterize most of the Cenozoic, it is possible
In other words, when average global temperatures during the Eocene led to subtropical environments and greater species richness, tropical regions are nearer this limit than temperate and polar regions than observed today. The disproportionate taxonomic richness of plants (34, 35), insects (36), and mammals (37) since the Late Miocene.

Variation in bandwidth. Several other factors can bias our estimate of the true LDG slope including those particular to the fossil record (e.g., uncertainty in age of fossil collections, temporal and geographic variation in sampling) and parameters of the analysis (e.g., duration of intervals, widths of latitudinal bands). We attempted to ameliorate these potential biases using by performing 10,000 pseudoreplicate analyses, each of which analyzed identical occurrence data, but varied particular parameters using several procedures described below. Although these sources of bias present an analytical challenge (e.g., ref. 7), we calculate the LDG slope for each of the natural loadings of species richness within a band using the paleolatitudes (i.e., the latitudes at the time of deposition) of the fossil occurrences. We then calculated the LDG slope in each interval as the linear regression of the natural logarithm of species richness within a band against the latitudinal midpoint (i.e., degrees north of the equator) of the band.

Our use of 2-My intervals necessarily bins taxa that were not coeval, thereby time-averaging the actual LDG. Nevertheless, we believe that binning taxa into time intervals makes our analysis more conservative. Our null hypothesis is a temporally constant LDG, with a slope equal to that observed today. Under this null hypothesis, pooling taxa over 2 My should only increase the apparent strength of the LDG. Several other factors can bias our estimate of the true LDG slope including those particular to the fossil record (e.g., uncertainty in age of fossil collections, temporal and geographic variation in sampling) and parameters of the analysis (e.g., duration of intervals, widths of latitudinal bands). We attempted to ameliorate these potential biases using by performing 10,000 pseudoreplicate analyses, each of which analyzed identical occurrence data, but varied particular parameters using several procedures described below. Although these sources of bias present an analytical challenge (e.g., ref. 7), we calculate the LDG slope for each of the natural loadings of species richness within a band using the paleolatitudes (i.e., the latitudes at the time of deposition) of the fossil occurrences. We then calculated the LDG slope in each interval as the linear regression of the natural logarithm of species richness within a band against the latitudinal midpoint (i.e., degrees north of the equator) of the band.

Uncertainty of collection ages. Collections in the Paleobiology Database typically are assigned maximum and minimum absolute age estimates. (Any rare occurrences lacking minimum or maximum absolute age estimates were excluded from analysis.) In each of the 10,000 pseudoreplicate analyses, we assigned each collection an absolute age drawn from a random uniform distribution bounded by that collection’s designated minimum and maximum ages.

Variation in bandwidth. Because of the irregular geographic distribution of fossil collections, the choice of bandwidth affects the calculated LDG slope. Accordingly, each pseudoreplicate analysis used a different bandwidth between 1° and 5° of


7. Fraser D, Hassall C, Gorelick R, Rybczynski N (2014) Mean annual precipitation explains the modern latitudinal diversity gradient of North American mammals, irrespective of the distribution of fossil collections. As above, we used the overf function in R package sp to determine which species geographic range maps overlapped the same bands described above, only here, not limiting the longitudinal range. We then pooled these to determine the extant species richness in each band. As above, we calculated the extant LDG slopes and the linear regression of the natural logarithms of species richness within a band against the latitudinal midpoint (i.e., degrees north of the equator) of the band. Bandwidth had little influence on the resulting face-value slopes, which ranged between -0.0454 and -0.0378 (Fig. 2).


Supporting Information

Marcot et al. 10.1073/pnas.1524750113

SI Materials and Methods

The reader understandably may be concerned that the strength of the observed correlation between first-differences of fossil slope and $\delta^{18}$O may be disproportionately influenced by what appears to be an outlier on the abscissa of Fig. 3 (figure repeated in Fig. S1, Left). First of all, we emphasize that we determined the strength of this correlation using Spearman’s nonparametric rank correlation. Therefore, the absolute value of the point showing the highest mean change in median fossil slope does not influence the strength of the correlation. In other words, if the value were 0.02 or 0.20—any value, so long as that point remained the highest observed value among all of the points—the strength of the nonparametric correlation would remain unchanged.

Nevertheless, to demonstrate that the apparent outlier does not solely determine the strength of the observed correlation between first-differences of fossil slope and $\delta^{18}$O, we repeatedly recalculated the correlation, each time excluding one pair of coordinates.

The results of this analysis are shown in Fig. S1. The removal of any point, with the exception of “61,” increases the $P$ value greater than 0.05. Generally, this is simply the predictable result of losing statistical power with smaller sample sizes. Three points in particular (“3,” “13,” and “63”) seem to increase the $P$ value substantially more than the others, one of these being the apparent outlier (“63”). Therefore, point “63” is not the only point that has a great influence on the strength of the correlation. Furthermore, removing the apparent outlier (“63”) results in only the third-greatest increase in $P$ value, indicating that point “63” is not even the greatest contributor to the strength of the correlation. Notably, one point (“61”) actually decreases the $P$ value substantially, and removal of this point would actually increase the strength of the correlation.

In summary, given our rigorous vetting of intervals for analysis (Materials and Methods), we prefer to use the data from all of available intervals and believe the use of nonparametric correlation helps ameliorate the distorting effects of potential outliers on the correlation.
The dependence of the correlation between first-differences on each coordinate pair of δ¹⁸O and fossil slopes. (Left) The same as Fig. 3 but with the addition of labels adjacent to each point. This correlation is between first-differences: changes between consecutive intervals, so the point labels refer to the younger of the two intervals. For example, the point labeled “63” is the change in slope and δ¹⁸O between the interval beginning at 65 Ma and the subsequent interval beginning 63 Ma. (Right) Diagram showing the P value of the Spearman correlation when the indicated point is removed.
Calculated LDG slopes for all band widths for each analyzed interval. The calculated fossil (squares) and adjusted modern slopes (circles) are shown for each successful pseudoreplicate analysis, each of which had a different bandwidth (Materials and Methods). Pairs of points are colored red for fossil slopes, blue for adjusted modern slopes. Pairs of points are colored gray for pseudoreplicates in which the fossil slope was significantly greater than the adjusted modern slope (i.e., a weaker LDG). Pairs of points are both colored gray for pseudoreplicates and pairs in which the slopes were not significantly different. Note that in some intervals, slopes could only be calculated for a small subset of band widths (e.g., 11 Ma). Nevertheless, although the choice of bandwidth did influence the absolute value of the slopes calculated, it very rarely influenced the relationship between the fossil and adjusted modern slopes.