Supplementary Information (SI) for article: Herbivore teeth predict climatic limits in Kenyan ecosystems

Indrė Žliobaitė[a,b,1], Janne Rinne[a,c], Anikó B. Tóth[d,e], Michael Mechenich[a], Liping Liu[f], Anna K. Behrensmeyer[d], and Mikael Fortelius[a,g,h]

[a]Department of Geosciences and Geography, University of Helsinki, P.O. Box 64, FI-00014University of Helsinki, Finland; [b]Helsinki Institute for Information Technology HIIT, Aalto University, PO Box 15600, FI-00076 Aalto, Finland; [c]Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, S-22362 Lund, Sweden; [d]Evolution of Terrestrial Ecosystems Program, Department of Paleobiology, MRC 121, National Museum of Natural History, Smithsonian Institute, Washington, D.C. 20013 USA; [e]Department of Biological Sciences, Macquarie University, NSW 2109 Sydney, New South Wales, Australia; [f]Department of Paleobiology, Swedish Museum of Natural History, P.O. Box 50007, SE-11418, Stockholm, Sweden; [g]Centre for Ecological and Evolutionary Synthesis, Dept. of Biosciences, University of Oslo, P.O. Box 1066 Blindern, NO-0316 Oslo, Norway; [h]Museum of Natural History, Invalidenstraße 43, 10115, Berlin, Germany

This manuscript was compiled on September 30, 2016

Materials and methods

Study area and environmental characteristics. We use data spanning the past 60 years from 13 Kenyan National parks. The protected areas of the national parks and their basic environmental characteristics are listed in Table 1. The protected areas, or sites, cover a range of climatic conditions, from hot and dry, e.g., Tsavo, to moist and cool, e.g. Aberdares. While rainfall in Kenya is mostly dependent on elevation, there are other factors influencing it such as proximity to the Indian Ocean, rain shadow effects on lee sides of major mountains, and the generally higher rainfall rates in the west, close to the Congo basin. In our dataset the NDVI, i.e. the greenness of the vegetation, depends mostly on rainfall.

The climate variables included in this study are temperature (TEMP) and precipitation (PREC). These data were obtained from WorldClim [1] and represent interpolated and averaged observations from year 1950 to 2000, roughly the same period (1950-2012) represented by the mammal occurrence data.

Net primary productivity (NPP) is computed from temperature and precipitation using the classic formula from [2], as cited in [3]. The formula computes NPP components from temperature and precipitation separately, and then takes the minimum of the two. As our study domain lies in low latitudes, the NPP varies mostly due to precipitation, except for high (above 2500 m) elevation sites (ABD, EL, MKE), where NPP is due to temperature in accordance with the findings in [4], which demonstrate that in colder areas the NPP depends mostly on temperature while in warmer areas it depends on precipitation.

Normalized difference vegetation index (NDVI) indicates the amount of live green vegetation as observed via remote sensing. NDVI is calculated from satellite observations as difference between reflected near-infrared and visible radiation, normalized by dividing with their sum [5]. In this study the spatial and temporal average observed index for each national park is calculated over nine years (2001-2009) using observations available at every 16 days.

In addition to average TEMP, PREC, NPP and NDVI we compute extremities of these variables (minimums and maximums) expecting to capture short and longer term limiting conditions for productivity, that is, the availability of high-quality plant foods. Due to nature of the data sources not all variants are possible for all variables, but where possible we at computing extreme values over different time spans as well as spatially. Over time we look at monthly, monthly, yearly and 9 years (only possible for NDVI) extremes with the goal to analyze dental traits in relation to variability of the environment.

Table 1. National park sites and their characteristics. For productivity NPP (P) indicates that the value is mostly precipitation controlled, and (T) indicates that it is mostly temperature controlled.

Site name	Area km2	Elev. m	Temp. C	Prec. mm	NPP gC/m2	NDVI	Spec. count
Aberdares	1930	2712	11.9	1656	1576 (T)	0.71	23
Elgon	733	2504	15	1606	1846 (T)	0.68	17
Kakamega	540	1793	18.8	1779	2079 (P)	0.75	20
Masai Mara	1520	1601	19.5	1038	1494 (P)	0.53	35
Meru	5090	421	25.6	430	745 (P)	0.37	38
Mt Kenya	2600	2686	12.4	1594	1620 (T)	0.63	23
Nairobi	66	1628	19	742	1167 (P)	0.42	31
Naivasha	119	1974	16.7	709	1126 (P)	0.46	31
Samburu	424	883	23.6	504	853 (P)	0.3	29
Shimba Hills	224	242	24.5	1102	1557 (P)	0.62	20
Shompole	183	1120	22.6	567	941 (P)	0.54	29
Tana	560	77	27.3	508	859 (P)	0.39	37
Tsavo	20210	544	24.5	714	1133 (P)	0.34	35

Data sources. Ecometric modeling requires three types of data: environmental characteristics of sites, dental traits of species, and occurrence-of-species at sites. We use environmental and occurrence-of-species data covering the last 60 years from 13 protected areas (sites) of national parks in Kenya. Our study area covers a relatively small spatial scale (34 thousand km^2), which is not meant to generalize on a continental scale, but it offers very precise lists of species occurrence, which allows for very accurate analysis as compared to global scale possibilities.

Data used in our analysis originate from several sources: presence-absence of species with body masses, as well as temperature and elevation of the sites, are from [6], updated presence-absence data were received from A. Tóth, temperature originally obtained from WorldClim [1], body mass obtained from Pantheria [7] and the MOM database [8]. Precipitation data are from WorldClim [1]. Dental traits (functional

[1]To whom correspondence should be addressed. E-mail: indre.zliobaite@helsinki.fi

crown types) are newly developed for this study. Hypsodonty values are obtained from [3], completed and amended by one of the authors (MF). The Net Primary Productivity (NPP) was calculated from WorldClim data. The Normalized Difference Vegetation Index (NDVI) is a remote sensing product downloaded from NASA Earth Observations website[1].

The climate data have been processed as follows. First, WorldClim global temperature and precipitation data were obtained in grid format. These grids were then converted to vector polygon format, and intersected with polygons representing the parks included in this study. Finally, from these intersections, summary climatic values, including spatial minimums and area-weighted averages, were calculated for each park.

From the original species distribution reported in [6] we include only orders Proboscidea, Primates, Perissodactyla, and Artiodactyla, following [9] and subsequent publications on ecometrics (see [10] for a recent review), the main orders of large herbivores. We exclude domestic species (*Bos Taurus*, *Capra hircus*, *Ovis aries*, *Camelus dromedaries*, *Equus asinus*, *Equus caballus*) from the analysis. Table 5 reports the numbers of species found in each protected area.

The dental traits of large herbivorous mammal species occurring in Kenya national parks used in this study have been scored (by MF) according to the new FCT scheme. The scores are given in Table S1 in Supporting Information.

Regression models. We build upon work by Liu et al [3], where an ecometric regression model was developed for estimating Net Primary Productivity (NPP) using dental functional traits (molar height and the number of longitudinal lophs). The model built as Ordinary Least Squares (OLS) regression explained about 73% of the global variance in the land productivity.

Figure 1 illustrates our modeling procedure. From species occurrence data and dental trait data we compute average traits for each site. Then we relate these average traits to the climate and environmental characteristics of each site by models fit on modern day data. The resulting regression model can be applied to fossil record to predict climate and environmental conditions in the past.

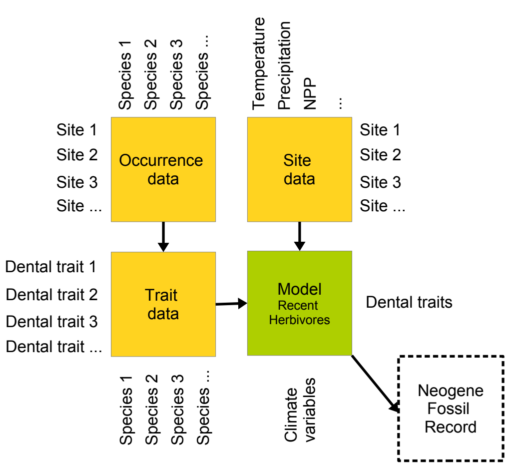

Fig. 1. Modeling procedure.

[1] http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD13A2_M_NDVI

We have a small sample size (13 sites) and many dental traits to consider, therefore, to control the risk of rather than capturing generic patterns to capture once-off deviations due to models being to flexible, we restrict our analysis to simple linear models. Instead of the standard OLS regression for model fitting, we use Least Angle Regression (LARS) [11], which is better suited for high dimensional data (relative to the number of observations) with correlated predictors. We use LARS with dental traits as inputs to predict each environmental variable separately. The resulting models from LARS look like normal linear regression equations, but the procedure for obtaining these equations is special.

LARS constructs a model iteratively. First a predictor is selected that is the most correlated with the target variable. A one-variable regression model is fitted. Instead of using the regression coefficient in full, a small fraction of the coefficient is used, estimates are made using this regression model and residuals are computed. The procedure is repeated and the utilization of the regression coefficient is increased until the next predictor becomes more correlated with the residual. The procedure continues until a stopping criterion is met, in our case - a user defined number of selection steps is reached. The optimal number of steps is either fixed a priori, or selected via leave-one-out cross validation (LOOCV).

We use leave-one-out cross-validation (LOOCV) procedure for optimizing model parameters and analyzing the performance on unseen data. The procedure is as follows. Having a dataset of n observations (sites), one observation is removed from the dataset, a model is built on the remaining n-1 observations, and a prediction is made on the omitted observation. The same procedure is repeated for all n observations. This way testing is done on all the observations in the dataset while not using those observations in model building. This allows more objective assessment of the performance.

In addition to LARS, for comparison purposes we report the standard OLS regression performance on log-mass alone, and the number of species alone, and the seven dental traits plus log-mass.

We use the coefficient of determination (R^2) for assessing the predictive performance, since it gives a standardized score that is easy to interpret. R^2 takes values between minus infinity and one, the higher – the better. If tested on the same data as the model was fitted, zero means that the performance is the same as a naive baseline always predicting a constant. When using R^2 with cross-validation, we need to normalize the score as otherwise the naive baseline is below zero (for details see [12]). The normalization depends only on the number of data points in cross-validation, and is $R^{2\star} = (R^2 - R^2_{naive})/(1 - R^2_{naive})$, where $R^2_{naive} = 1 - n^2/(n-1)^2$.

Data accessibility and implementation of computational experiments. The data and the code for reproducing the computational experiments reported in this study are available online[2]. In addition, for visual reference the datasets are given in Supplementary Information. Computational experiments are implemented and executed in R[3]. We use the implementations of Least Angle Regression (LARS) provided in R package lars[4].

[2] https://github.com/zliobaite/paper-Kenya-parks
[3] https://www.r-project.org/
[4] https://cran.r-project.org/web/packages/lars/index.html

Functional Crown Type (FCT) scoring scheme

Jernvall [13] introduced a modular system called crown types to describe the topography of all developing and unworn mammalian molar teeth. While the original crown type scheme has proven useful for capturing the functional evolution of the teeth of mammalian herbivores (e.g., [3, 13, 14], it is not ideal for capturing the functional competence of herbivore teeth, created and maintained by dental wear. Here we introduce a system of functional crowntypes explicitly designed for capturing molar shape and main functional traits of worn occlusal surfaces of the molar dentition. The system has been designed to be generally applicable to all living and fossil herbivorous mammals, regardless of phylogenetic relationships, size or design of the chewing apparatus.

Scoring scheme. Our scheme has seven variables, two ordinal and five binary:

- hypsodonty (HYP)
- horizodonty (HOD)
- presence of acute lophs (AL)
- presence of obtuse or basin-like lophs (OL)
- structural fortification of cusps (SF)
- occlusal topography (OT)
- coronal cementum (CM)

The scoring system is as follows:

- HYP: brachydont (1), mesodont (2), hypsodont (3)
- HOD: brachyhorizodont (1), mesohorizodont (2), hypsohorizodont (3)
- AL: absent (0), present (1)
- OL: absent (0), present (1)
- SF: absent (0), present (1)
- OT: has raised elements (0), is flat (1)
- CM: absent or very thin (0), thick coating (1)

Hypsodonty definition = a tooth lower than long is brachydont, teeth of approximately equal length and height are mesodont, teeth distinctly taller that long are hypsodont. Non-binding guidelines based on this ratio: brachydont <=0.8, mesodont 0.8-1.2, hypsodont >1.2

Horizodonty definition = a hypsohorizodont tooth is distinctly elongated in the mesiodistal (anteroposterior) direction. The number of cusp pairs in this dimension is a good guideline. 1-2 cusp pairs indicates brachyhorizodonty, 3 cusp pairs indicates mesohorizodonty and 3 or more cusp pairs indicates hypsohorizodonty. Cusplets distinctly smaller than the main cusp pairs are not counted.

Loph definition = a loph is a continuous, straight, zigzagged, curved or undulating continuous structure consisting of one or two enamel crests or a row of separate cusps with elongated features closely spaced. A loph has occlusal enamel-enamel contact but planar facets are not required. A loph is at least half the size of the tooth along the main axis of the loph.

Loph shape = acute has a single, distinct edge, obtuse does not. A loph is only acute if the leading enamel itself forms a sharp edge, or supports a distinct, planar wear facet. If in doubt, go for obtuse. Acute lophs are seen in tapirs, chalicotheres, brachydont horses, ectolophodont rhinos, deinotheres, some monkeys. Selenodont lophs are scored as obtuse and the differences are recorded under occlusal topography.

Fortified cusps = a structure amplifying the protrusion of a cusp, as in reduncine bovids. There must be some clear sign of fortification, such as Furchen (e.g., most suids) or locally thickened enamel loops (e.g. hippos and various bovids). Regular selenodont cusps are not fortified.

Occlusal topography = either with raised cusps/lophs or fused into a single surface essentially in one plane, often with cusps/lophs embedded in thick cementum. Examples include plagiolophodont horses and rhinos, elephantids, warthogs, many rodents. Modern Bovini, and Alcelaphini have a flat occlusal topography but most other bovids have raised elements. Occlusal topography is a property determined mainly by dental structure, rather than the diet of the individual. In contrast to the relief variable of traditional mesowear analysis, it is assessed perpendicular to the direction of occlusal movement (the direction least sensitive to dietary effects). Where occlusal movement is minor, any direction will do.

Cementum = presence of cementum on the crown (as well as on the roots). Very thin cementum is scored as absence, mainly because it is often impossible to tell these states apart on fossil specimens. Thick cementum means that cementum thickness is of the same order of magnitude as enamel thickness.

Examples are given in Figure 2.

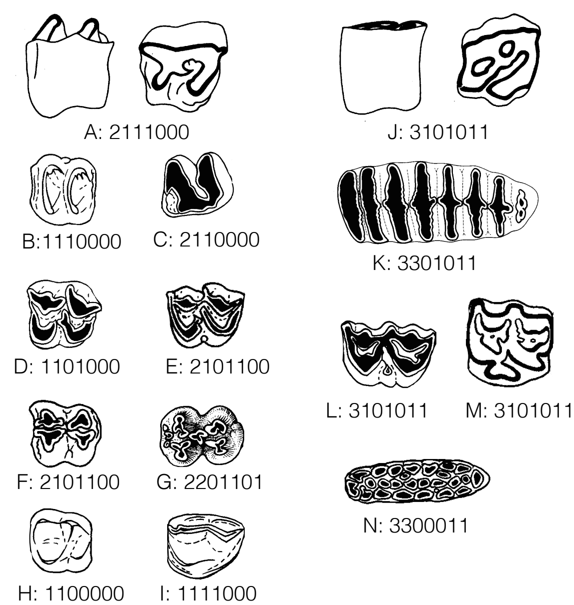

Fig. 2. Drawings of selected molars to illustrate the scoring system of functional crown types. The numbers indicate the sequence of crown type scores (HYP, HOD, AL, SF, OT, CM). A and J show the principal difference between an occlusal surface with raised elements. The left column shows additional examples where the occlusal surface is scored as having raised elements, the right one shows examples where it is scored as flat. Note that selenodont teeth can have both states (E, D versus L). All teeth shown are brachyhorizodont except K and N, which are hypsohorizodont. The genera shown are: A *Diceros*, B *Listriodon*, C *Arsinoitherium*, D *Giraffa*, E *Cervus*, F *Hippopotamus*, G *Hylochoerus*, H *Pan*, I *Megaladapis*, J *Ceratotherium*, K *Loxodonta*, L *Bos*, M *Equus*, N *Phacochoerus*. A and J are from [15], all other examples from [16].

Scoring guidelines. The general spirit of this scoring system is to concentrate on the obvious structures and ignore minor structures and nuances.

The default tooth is upper M2, for practical as well as developmental reasons [17], but when some other tooth is clearly more characteristic the other tooth is used (P4 for most carnivores, M3 for many derived pigs). The philosophical principle is that M2 should always be used when it doesn't matter. When it makes a significant difference, the dominant tooth should be used. The purpose is always to capture the significant traits of the entire molar dentition. Whenever possible, multiple specimens should be investigated.

Wear may influence the scoring and teeth in early middle wear should be used if possible. Unworn or little worn teeth and very worn teeth are disqualified.

Analysis of indicator species

From the ecometric analysis it looks like distribution of the functional dental traits over sites is capturing two characteristic environmental conditions in Kenya: (1) precipitation related wetlands, which are species-rich, while have relatively low precipitation, with other water sources, such as rivers or lakes, available, and (2) temperature related galley forests, which are cooler spots within extremely hot and dry surroundings.

To find out what is the dental signature for wetlands we select two indicator species and investigate in what kind of sites they occur, and in what kind of sites they do not occur. We look at hippos and reduncines, which are known to inhabit wetlands. Our dataset includes only one species of hippos (*Hippopotamus amphibius*) and three species of reduncines (*Redunca fulvorufula*, *Redunca redunca*, and *Kobus ellipsiprymnus*). Out of reduncines *Redunca fulvorufula* occurs only in two sites, and *Kobus ellipsiprymnus* occurs in all but one sites. Lack of variability in occurence would mask the pattern, thus we pick *Redunca redunca*, which occurs in about half of the parks.

Hippo is found in 8 parks out of 13, redunca in 7 out of 13. Among those 5 parks overlap. Even though hippo and redunca are not found in exactly the same parks, the signatures of the parks where they are found vs. not found are highly similar, and how these characteristics deviate from the mean is also highly similar, as Figure 3 shows.

Environment-wise hippos and reduncas are clearly in places with high above-average species counts, with low precipitation, hence low NPP, also low NDVI places. With hotter temperatures than average, especially for hippos. Both occur in below average elevations, especially hippos.

Trait-wise hippos and reduncas occur in high above average hypsodonty (and cementum) sites, with high above average obtuse lophs (OL), below average acute lophs (AL), and with indeed lower than average structural fortification (SF), which is due to species count, as discussed in the main manuscript text.

To find out what is the dental signature for forests we select three indicator species expected to be found in forests: *Hylochoerus meinertzhageni*, *Galago senegalensis*, and *Tragelaphus imberbis*. Interestingly, all three are nocturnal animals that tend to hide from daily heat in trees.

Hylochoerus meinertzhageni is native to wooded habitats. The giant forest hog is mainly a herbivore, but also scavenges. *Galago senegalensis* tend to live in dry woodland regions and savannah regions. They are omnivores, eating birds and insects, fruit, seeds, flowers, eggs, nuts, and tree gums. *Tragelaphus imberbis* - is a forest antelope, it inhabits dry, flat, and heavily forested regions, and eats foliage from trees and shrubs.

Hylochoerus present in 6, Galago in 9, out of those 5 overlap. *Tragelaphus* occur in 5 sites, out of which only 2 overlap with *Hylochoerus*, and 4 overlap with *Galago*. Based on inspection of the site characteristics, *Hylochoerus* tend to occur in classical forests, while *Tragelaphus* occurs in galley forests near water sources at extremely hot and dry areas. Figure 4 summarizes site characteristics for these species occurrence.

Hylochoerus present in very humid places, *Galago* in a bit drier than average, *Tragelaphus* in very dry. *Hylochoerus* in colder than average, *Galago* and *Tragelaphus* in highly hotter than average. Out of those, *Tragelaphus* has hotter maximums, and higher difference between max hot and average hot. Species count does not matter for *Hylochoerus*, *Galago* is in high species, *Tragelaphus* in very high species. Elevation-wise *Hylochoerus* is high, *Galago* and *Tragelaphus* are low.

In terms of the dental signature *Hylochoerus* and *Galago* are in slightly low hypsodonty (HYP) sites, but not very distinctly. For *Tragelaphus* HYP does not matter much. *Hylochoerus* is in sites with high acute loph (AL) scores, for *Galago* and *Tragelaphus* AL do not seem to matter much. Both *Hylochoerus* and *Galago* are in low obtuse loph (OL) sites, for *Tragelaphus* OL do not matter much. *Hylochoerus* is in high structural fortification (SF) sites, *Tragelaphus* is in low SF, while *Galago* is middle range SF. Cementum does not matter for either.

It seems that for *Tragelaphus* only HOD and SF matters - high HOD, low SF. They live in extremely hot and extremely low rain. Looking at the map, these parks: MR, SB, SHP, TV, TV. These look mostly like galley forests, rivers in deserts, see the map in Figure 5, as well as most of the seasonal rivers are there.

Datasets

The data and the code for reproducing the computational experiments reported in this study is available online[5]. For visual reference the datasets are given in Tables 3, 4, 5.

Abbreviations used

Table 2 presents abbreviations used. All the climate and environmental variables are given in Table 6.

Table 2. Abbreviations and acronyms.

FCT	Functional Crown Types
HYP	Hypsodonty
AL	Acute Lophs
OL	Obtuse, or basin-like Lophs
SF	Structural Fortification
OT	Occlusal Topography
CM	Coronal Cementum
OLS	Ordinary Least Squares
LARS	Least Angle Regression
PREC	Precipitation
NPP	Net Primary Productivity
NDVI	Normalized Difference Vegetation Index
TEMP	Temperature

[5] https://github.com/zliobaite/paper-Kenya-parks

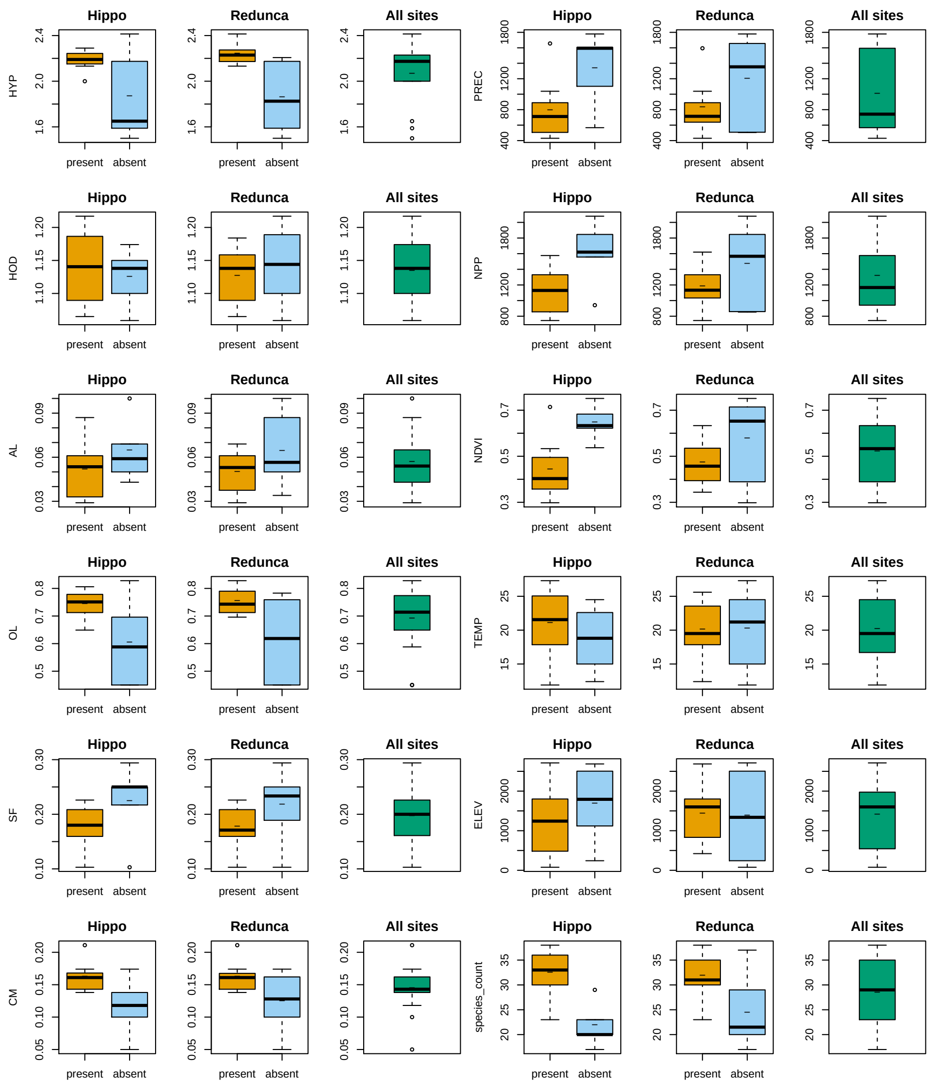

Fig. 3. Environment characteristics and dental trait characteristics of sites where emphHippopotamus amphibius, and *Redunca redunca* occur.

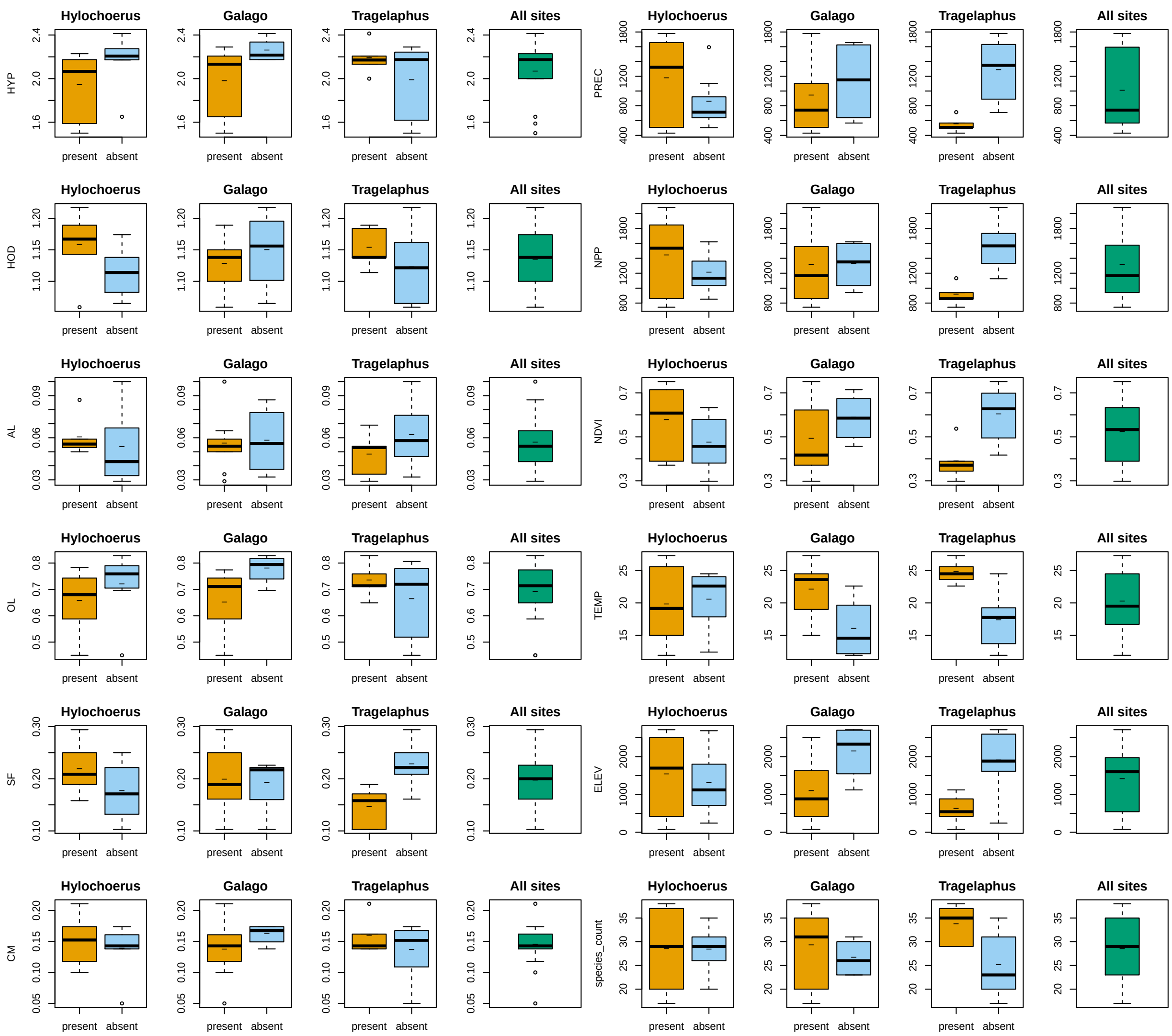

Fig. 4. Environment characteristics and dental trait characteristics of sites where *Hylochoerus meinertzhageni*, *Galago senegalensis*, and *Tragelaphus imberbis* occur.

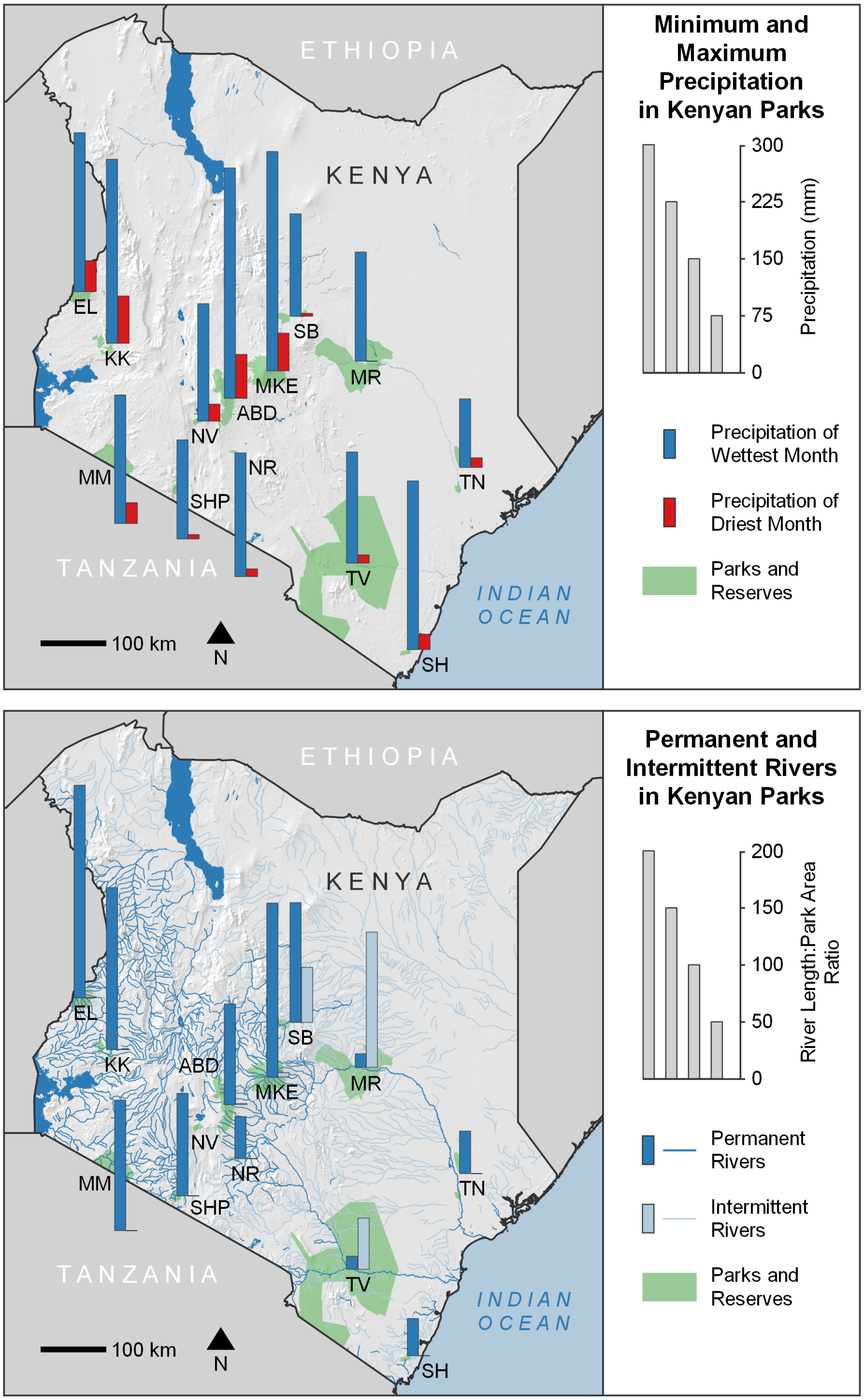

Fig. 5. Permanent and seasonal rivers and precipitation on sites. The vertical bars indicate the length of rivers relative to the park area, and relative amount of annual precipitation for comparison purposes.

Table 3. Dental traits and mass data.

TAXON	ORDER	FAMILY	MASS_KG	HYP	HOD	AL	OL	SF	OT	CM
Aepyceros melampus	Artiodactyla	Bovidae	52.5	3	1	0	1	0	0	0
Alcelaphus buselaphus	Artiodactyla	Bovidae	171	3	1	0	1	0	1	0
Beatragus hunteri	Artiodactyla	Bovidae	80	3	1	0	1	0	1	0
Cephalophus adersi	Artiodactyla	Bovidae	9.2	2	1	0	1	1	0	0
Cephalophus harveyi	Artiodactyla	Bovidae	14.5	2	1	0	1	1	0	0
Cephalophus nigrifrons	Artiodactyla	Bovidae	13.9	2	1	0	1	1	0	0
Cephalophus silvicultor	Artiodactyla	Bovidae	72.5	2	1	0	1	1	0	0
Cephalophus weynsi	Artiodactyla	Bovidae	17	2	1	0	1	1	0	0
Connochaetes taurinus	Artiodactyla	Bovidae	180	3	1	0	1	0	1	0
Damaliscus korrigum	Artiodactyla	Bovidae	112	3	1	0	1	0	1	0
Eudorcas thomsonii	Artiodactyla	Bovidae	20.5	3	1	0	1	0	0	0
Hippotragus equinus	Artiodactyla	Bovidae	270	3	1	0	1	1	0	0
Hippotragus niger	Artiodactyla	Bovidae	227.5	3	1	0	1	1	0	0
Kobus ellipsiprymnus	Artiodactyla	Bovidae	210	3	1	0	1	1	0	0
Litocranius walleri	Artiodactyla	Bovidae	38	2	1	0	1	0	0	0
Madoqua guentheri	Artiodactyla	Bovidae	7.5	2	1	0	1	0	0	0
Madoqua kirkii	Artiodactyla	Bovidae	5.3	2	1	0	1	0	0	0
Nanger granti	Artiodactyla	Bovidae	55	3	1	0	1	0	0	0
Neotragus moschatus	Artiodactyla	Bovidae	6.5	2	1	0	1	0	0	0
Oreotragus oreotragus	Artiodactyla	Bovidae	13	3	1	0	1	0	0	0
Oryx beisa	Artiodactyla	Bovidae	169	3	1	0	1	1	1	0
Ourebia ourebi	Artiodactyla	Bovidae	17.2	3	1	0	1	0	0	0
Philantomba monticola	Artiodactyla	Bovidae	6.2	1	1	0	1	0	0	0
Raphicerus campestris	Artiodactyla	Bovidae	10.5	3	1	0	1	0	0	0
Redunca fulvorufula	Artiodactyla	Bovidae	29.5	3	1	0	1	1	0	1
Redunca redunca	Artiodactyla	Bovidae	44.1	3	1	0	1	1	0	1
Sylvicapra grimmia	Artiodactyla	Bovidae	19.5	2	1	0	1	0	0	0
Syncerus caffer	Artiodactyla	Bovidae	580	3	1	0	1	0	1	0
Taurotragus oryx	Artiodactyla	Bovidae	570	3	1	0	1	0	0	0
Tragelaphus eurycerus	Artiodactyla	Bovidae	329	2	1	0	1	0	0	0
Tragelaphus imberbis	Artiodactyla	Bovidae	70	2	1	0	1	0	0	0
Tragelaphus scriptus	Artiodactyla	Bovidae	43.3	2	1	0	1	0	0	0
Tragelaphus spekii	Artiodactyla	Bovidae	78	2	1	0	1	0	0	0
Tragelaphus strepsiceros	Artiodactyla	Bovidae	213.5	2	1	0	1	0	0	0
Giraffa camelopardalis	Artiodactyla	Giraffidae	850	1	1	0	1	0	0	0
Hippopotamus amphibius	Artiodactyla	Hippopotamidae	1417.5	2	1	0	1	1	0	0
Hylochoerus meinertzhageni	Artiodactyla	Suidae	188.5	2	2	0	1	1	0	1
Phacochoerus aethiopicus	Artiodactyla	Suidae	50	3	3	0	0	0	1	1
Phacochoerus africanus	Artiodactyla	Suidae	82.5	3	3	0	0	0	1	1
Potamochoerus larvatus	Artiodactyla	Suidae	97.5	1	1	0	0	1	0	0
Equus burchellii	Perissodactyla	Equidae	276	3	1	0	1	0	1	1
Equus grevyi	Perissodactyla	Equidae	408	3	1	0	1	0	1	1
Diceros bicornis	Perissodactyla	Rhinoceratidae	1180.5	2	1	1	1	0	0	0
Ceratotherium simum	Perissodactyla	Rhinocerotidae	2177.6	3	1	0	1	0	1	1
Cercocebus galeritus	Primates	Cercopithecidae	7.8	1	1	0	0	0	0	0
Cercopithecus ascanius	Primates	Cercopithecidae	3.5	1	1	0	0	0	0	0
Cercopithecus mitis	Primates	Cercopithecidae	5	1	1	0	0	0	0	0
Cercopithecus neglectus	Primates	Cercopithecidae	5.5	1	1	0	0	0	0	0
Chlorocebus pygerythrus	Primates	Cercopithecidae	5.2	1	1	0	0	0	0	0
Chlorocebus tantalus	Primates	Cercopithecidae	4.2	1	1	0	0	0	0	0
Colobus angolensis	Primates	Cercopithecidae	9.8	1	1	1	0	0	0	0
Colobus guereza	Primates	Cercopithecidae	10.6	1	1	1	0	0	0	0
Erythrocebus patas	Primates	Cercopithecidae	8.1	1	1	0	0	0	0	0
Papio anubis	Primates	Cercopithecidae	19.4	1	1	0	0	0	0	0
Papio cynocephalus	Primates	Cercopithecidae	21	1	1	0	0	0	0	0
Piliocolobus rufomitratus	Primates	Cercopithecidae	8	1	1	1	0	0	0	0
Galago gallarum	Primates	Galagidae	0.2	1	1	0	0	0	0	0
Galago senegalensis	Primates	Galagidae	0.2	1	1	0	0	0	0	0
Galago zanzibaricus	Primates	Galagidae	0.1	1	1	0	0	0	0	0
Otolemur crassicaudatus	Primates	Galagidae	1.5	1	1	0	0	0	0	0
Otolemur garnettii	Primates	Galagidae	0.8	1	1	0	0	0	0	0
Perodicticus potto	Primates	Lorisidae	1.1	1	1	0	0	0	0	0
Loxodonta africana	Proboscidea	Elephantidae	3940	3	3	0	1	0	1	1

Table 4. Species occurrences.

TAXON	ABD	EL	KK	MM	MR	MKE	NR	NV	SB	SH	SHP	TN	TV
Aepyceros melampus	1	0	0	1	1	0	1	1	1	0	1	0	1
Alcelaphus buselaphus	0	0	0	1	1	1	1	1	0	0	1	0	1
Beatragus hunteri	0	0	0	0	0	0	0	0	0	0	0	1	1
Cephalophus adersi	0	0	0	0	0	0	0	0	0	1	0	1	0
Cephalophus harveyi	1	1	1	0	0	1	0	1	0	1	0	1	1
Cephalophus nigrifrons	1	1	0	0	0	1	0	0	0	0	0	0	0
Cephalophus silvicultor	0	0	1	0	0	0	0	0	0	0	0	0	0
Cephalophus weynsi	0	1	1	0	0	0	0	1	0	0	0	0	0
Connochaetes taurinus	0	0	0	1	0	0	1	1	0	0	1	0	0
Damaliscus korrigum	0	0	0	1	0	0	0	0	0	0	1	1	0
Eudorcas thomsonii	0	0	0	1	0	1	1	1	1	0	1	0	1
Hippotragus equinus	0	0	0	1	0	0	0	0	0	0	0	0	0
Hippotragus niger	0	0	0	0	0	0	0	0	0	1	0	0	0
Kobus ellipsiprymnus	1	1	0	1	1	1	1	1	1	1	1	1	1
Litocranius walleri	0	0	0	0	1	0	0	1	1	0	1	1	1
Madoqua guentheri	0	0	0	0	1	0	0	0	1	0	0	1	0
Madoqua kirkii	0	0	0	1	1	0	1	1	1	0	1	1	1
Nanger granti	1	0	0	1	1	1	1	1	1	0	1	1	1
Neotragus moschatus	1	0	1	0	1	1	1	1	0	1	0	1	0
Oreotragus oreotragus	1	0	0	1	0	0	1	1	1	0	1	0	1
Oryx beisa	0	0	0	1	1	0	1	0	1	0	1	1	1
Ourebia ourebi	0	0	0	1	1	0	1	0	0	0	1	0	0
Philantomba monticola	1	1	1	1	1	0	1	1	0	1	0	1	0
Raphicerus campestris	0	0	0	1	1	1	1	1	1	0	1	1	1
Redunca fulvorufula	0	0	0	0	0	0	1	1	0	0	0	0	0
Redunca redunca	0	0	0	1	1	1	1	1	0	0	1	0	1
Sylvicapra grimmia	1	1	1	1	1	1	1	1	1	1	1	0	1
Syncerus caffer	1	0	0	1	1	1	1	1	1	1	1	1	1
Taurotragus oryx	1	0	0	1	1	1	1	1	1	0	1	1	1
Tragelaphus eurycerus	1	0	0	0	0	0	0	0	0	0	0	0	0
Tragelaphus imberbis	0	0	0	0	1	0	0	0	1	0	1	1	1
Tragelaphus scriptus	1	1	1	1	1	1	1	1	1	1	1	1	1
Tragelaphus spekii	0	0	1	0	0	0	0	0	0	0	0	0	0
Tragelaphus strepsiceros	0	0	0	1	1	0	0	1	1	0	1	0	1
Giraffa camelopardalis	0	1	0	1	1	1	1	1	1	0	1	1	1
Hippopotamus amphibius	1	0	0	1	1	0	1	1	1	0	0	1	1
Hylochoerus meinertzhageni	1	1	1	1	1	0	0	0	0	0	0	1	0
Phacochoerus aethiopicus	0	0	0	0	1	1	0	0	0	0	0	1	0
Phacochoerus africanus	1	0	1	1	1	0	1	1	1	1	1	1	1
Potamochoerus larvatus	0	0	1	1	1	1	0	1	0	1	0	1	1
Equus burchellii	1	1	0	1	1	1	1	1	1	0	1	1	1
Equus grevyi	0	0	0	0	1	0	0	0	1	0	0	1	1
Diceros bicornis	1	0	0	1	1	0	1	0	1	0	1	1	1
Ceratotherium simum	0	0	0	0	1	0	1	1	0	0	0	0	0
Cercocebus galeritus	0	0	0	0	0	0	0	0	0	0	0	1	0
Cercopithecus ascanius	0	0	1	0	0	0	0	0	0	0	0	0	0
Cercopithecus mitis	1	1	1	1	1	1	1	1	1	1	1	1	1
Cercopithecus neglectus	0	1	1	0	0	0	0	0	0	0	0	0	0
Chlorocebus pygerythrus	1	1	1	1	1	1	1	1	1	1	1	1	1
Chlorocebus tantalus	0	0	0	0	0	0	0	0	0	0	0	0	0
Colobus angolensis	0	0	0	0	0	0	0	0	0	1	0	0	0
Colobus guereza	1	1	1	1	1	1	1	1	0	1	1	0	0
Erythrocebus patas	0	0	0	1	1	0	0	0	1	0	0	0	0
Papio anubis	1	1	1	1	1	1	1	1	1	1	1	1	1
Papio cynocephalus	0	0	0	0	0	0	0	0	0	1	0	1	1
Piliocolobus rufomitratus	0	0	0	0	0	0	0	0	0	0	0	1	0
Galago gallarum	0	0	0	0	1	0	0	0	0	0	0	1	0
Galago senegalensis	0	1	1	1	1	0	1	0	1	1	0	1	1
Galago zanzibaricus	0	0	0	0	0	0	0	0	0	1	0	1	1
Otolemur crassicaudatus	0	0	1	1	1	0	1	0	1	1	0	1	1
Otolemur garnettii	0	0	0	0	0	0	0	0	0	0	0	0	1
Perodicticus potto	0	1	1	0	0	1	0	0	0	0	0	0	0
Loxodonta africana	1	0	0	1	1	1	0	0	1	0	1	1	1

Table 5. Values of the climate and environmental variables. In order to be comparable, all precipitation variables are reported in mm per annum (for example, PREC_MIN is precipitation of the driest month in mm times 12 months), temperature is reported in deg. Celsius, NPP is reported in gC per m2, and NDVI is an index between 0 and 1.

SITE	ABD	EL	KK	MM	MR	MKE	NR	NV	SB	SH	SHP	TN	TV
ELEV	2712	2504	1793	1601	421	2686	1628	1974	883	242	1120	77	544
TEMP	11.9	15.0	18.8	19.5	25.6	12.4	19.0	16.7	23.6	24.5	22.6	27.3	24.5
TEMP_MIN	10.4	14.0	17.7	18.2	23.8	11.1	16.7	14.8	22.6	22.3	20.5	25.2	22.1
TEMP_low	5.5	8.3	10.7	12.7	18.4	5.9	12.9	9.6	15.4	20.3	16.2	22.0	19.0
TEMP_low_MIN	4.3	8.0	9.9	11.6	16.7	4.8	11.0	8.4	14.6	18.3	14.5	19.9	17.1
TEMP_MAX	13.1	16.1	19.8	20.6	27.2	13.6	20.7	18.2	24.8	26.6	24.2	29.1	26.5
TEMP_high	18.4	21.8	27.0	26.4	32.9	19.1	25.3	23.9	32.0	28.8	29.2	32.6	30.1
TEMP_high_MAX	20.7	23.8	28.5	27.8	34.8	21.2	27.8	26.5	33.5	31.3	31.3	34.8	32.6
PREC	1656	1606	1779	1038	430	1594	742	709	504	1102	567	508	714
PREC_MIN	700	491	751	334	4	596	127	273	47	251	71	153	133
PREC_low	937	1127	1461	899	306	807	635	669	432	1003	474	451	504
PREC_MAX	3637	2514	2910	2034	1728	3461	1953	1859	1619	2665	1563	1086	1754
PREC_high	2325	1868	1980	1352	656	2174	840	826	689	1185	798	650	1117
NPP	1576	1846	2079	1494	745	1620	1167	1126	853	1557	941	859	1133
NPP_MIN_MIN	1115	835	1178	597	8	980	243	497	92	461	138	290	254
NPP_low_low	1022	1257	1469	1349	552	1054	1032	1076	748	1459	810	776	853
NPP_low_MIN	1022	835	1178	597	8	980	243	497	92	461	138	290	254
NPP_MIN_low	1390	1581	1863	1349	552	1244	1032	1076	748	1459	810	776	853
NDVI	0.714	0.683	0.751	0.533	0.371	0.633	0.417	0.457	0.298	0.622	0.537	0.389	0.344
NDVI_MIN9y	0.391	0.298	0.561	0.276	0.227	0.268	0.198	0.257	0.181	0.431	0.338	0.235	0.202
NDVI_low1y	0.484	0.537	0.642	0.341	0.245	0.435	0.258	0.333	0.203	0.487	0.389	0.277	0.231
NDVI_low1y_MIN	0.692	0.654	0.734	0.483	0.329	0.604	0.332	0.357	0.228	0.551	0.450	0.331	0.272
NDVI_low	0.690	0.659	0.741	0.502	0.345	0.602	0.390	0.436	0.277	0.596	0.509	0.368	0.322
NDVI_low_MIN	0.576	0.592	0.666	0.376	0.246	0.526	0.294	0.382	0.208	0.537	0.396	0.298	0.236
species_count	23	17	20	35	38	23	31	31	29	20	29	37	35

Table 6. Climate and environmental variables.

PREC	Annual precipitation
PREC_MIN	Minimum over monthly precipitation (driest month of the year)
PREC_low	Minimum annual precipitation spatially (driest place)
PREC_MAX	Maximum over monthly precipitation (wettest month of the year)
PREC_high	Maximum annual precipitation spatially (wettest place)
NPP	Standard computed Net Primary Productivity
NPP_MIN_MIN	Coldest month with driest month
NPP_low_low	Coldest days with driest place
NPP_low_MIN	Coldest days with driest month
NPP_MIN_low	Coldest month with driest place
NDVI	Observed annual average Normalized Difference Vegetation Index
NDVI_MIN9y	Global minimum over 9 years
NDVI_low1y	Average over yearly minimums
NDVI_low1y_MIN	Minimum over yearly averages
NDVI_low	Average over monthly minimums
NDVI_low_MIN	Minimum over monthly minimums
TEMP	Average temperature
TEMP_MIN	Minimum over monthly means (coldest month of the year)
TEMP_low	Average over monthly minimums (average coldest day of the month over year)
TEMP_low_MIN	Minimum over monthly minimums (coldest day of the coldest month)
TEMP_MAX	Maximum over monthly means (hottest month of the year)
TEMP_high	Average over monthly maximums (average hottest day of the month over year)
TEMP_high_MAX	Maximum over monthly maximums (hottest day of the hottest month)

1. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
2. Leith H (1972) Modeling the primary productivity of the world in *Primary productivity of the biosphere*, eds. Leith H, Whittaker RH. (Springer), pp. 237–263.
3. Liu L et al. (2012) Dental functional traits of mammals resolve productivity in terrestrial ecosystems past and present. *Proceedings of the Royal Society B* 279(1739):2793–2799.
4. Yi C et al. (2010) Climate control of terrestrial carbon exchange across biomes and continents. *Environmental Research Letters* 5(3):034007.
5. Huete A, Justice C, Liu H (1994) Development of vegetation and soil indices for MODIS-EOS. *Remote Sensing of the Environment* 49:224–234.
6. Tóth AB, Lyons SK, Behrensmeyer AK (2014) Mammals of Kenya's protected areas from 1888 to 2013. *Ecology* 95(6):1711–1711.
7. Jones KE et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* 90(2648).
8. Smith FA et al. (2003) Body mass of late quaternary mammals. *Ecology* 84:3402 (updated version obtained from the authors).
9. Fortelius M et al. (2002) Fossil mammals resolve regional patterns of eurasian climate change during 20 million years. *Evolutionary Ecology Research* 4:1005–1016.
10. Fortelius M et al. (2014) Evolution of neogene mammals in eurasia: environmental forcing and biotic interactions. *Annual Reviews of Earth and Planetary Science* 42:579–604.
11. Efron B, Hastie T, Johnstone I, Tibshirani R (2004) Least angle regression. *Annals of Statistics* 32(2):407–499.
12. Žliobaitė I, Tatti N (2016) A note on adjusting R2 for using with cross-validation. *arXiv* 1605.01703.
13. Jernvall J (1995) Mammalian molar cusp patterns: Developmental mechanisms of diversity. *Acta Zoologica Fennica* 198:1–61.
14. Jernvall J, Fortelius M (2002) Common mammals drive the evolutionary increase of hypsodonty in the neogene. *Nature* 417:538–540.
15. Fortelius M (1981) Functional aspects of occlusal cheek-tooth morphology in hypsodont non-ruminant ungulates in *Proc. of International Symposium on Concept and Method in Palaeontology, Barcelona*. pp. 153–162.
16. Thenius E (1989) Z ahne und gebiß der s augetiere in *Handbook of Zoology*, Mammalia, eds. Niethammer J, Schliemann H, Starck D. (Walter de Gruyter, Berlin) Vol. VIII, pp. 1–513.
17. Kavanagh K, Evans A, Jernvall J (2007) Predicting evolutionary patterns of mammalian teeth from development. *Nature* 449:427–432.