Taxon free analysis of mammalian communities in relation to evolutionary survivorship and environmental factors

Lintulaakso, Kari-Erik Johannes

2018-10-24

Lintulaakso, K-E J 2018, 'Taxon free analysis of mammalian communities in relation to evolutionary survivorship and environmental factors', Doctoral Programme in Geosciences, Helsinki.

http://hdl.handle.net/10138/337125

acceptedVersion

Taxon free analysis of mammalian communities in relation to evolutionary survivorship and environmental factors

KARI LINTULAAKSO

ACADEMIC DISSERTATION
To be presented with the permission of the Faculty of Science of
the University of Helsinki, for public examination in auditorium A 110
Chemicum, A. I. Virtasen aukio 1, on 24th October 2018, at 12 noon.

DEPARTMENT OF GEOSCIENCES AND GEOGRAPHY A65 / HELSINKI 2018

Author's address: Kari Lintulaakso
Finnish Museum of Natural History LUOMUS
P.O. Box 44 (Kumpula Manor, Jyrängöntie 2)
00014 University of Helsinki, Finland
kari.lintulaakso@helsinki.fi

Supervised by: Professor Mikael Fortelius
Department of Geosciences and Geography
University of Helsinki, Finland

Docent Jussi Eronen
Department of Geosciences and Geography
University of Helsinki, Finland

Reviewed by: Professor Pasquale Raia
University of Naples Federico II
Department of Earth Sciences, Environment and Resources, Italy

Professor Lars Werdelin
Swedish Museum of Natural History
Department of Paleobiology, Sweden

Opponent: Professor Jordi Agustí
University Rovira Virgili
Institute of Human Paleoecology and Social Evolution, Spain

ISSN-L 1798-7911
ISSN 1798-7911 (print)
ISBN 978-951-51-3982-5 (paperback)
ISBN 978-951-51-3983-2 (PDF)
http://ethesis.helsinki.fi

Helsinki University Print
Helsinki 2018

Lintulaakso K., 2018. *Taxon free analysis of mammalian communities in relation to evolutionary survivorship and environmental factors.* Helsinki University Press. Helsinki. 26 pages.

Abstract

The overall aims of this thesis study were to examine how different species traits are related to the evolutionary survivorship of land mammal taxa and how the taxon-free trait structure of mammalian communities is related to the climate, habitats and biogeography at the continental level.

The evolutionary survivorship of taxa was studied using a large fossil dataset of Cenozoic land mammals (New and Old Worlds, NOW, Fortelius, 2009) by controlling sampling biases, calculating the per capita origination and extinction rates of boundary crossers (genera that were recorded as present or absent during time bins that span 1, 1.5 or 2 M.y.) and estimating survival probabilities using capture–mark–recapture methods. The results demonstrate a recurring pattern whereby large mammal genera and species have higher origination and extinction rates and therefore shorter durations. This observed pattern can be explained because of behaviours such as hibernation, torpor and burrowing, collectively termed "sleep-or-hide" (SLOH). SLOH behaviour is more common in small mammals and they are therefore better protected from environmental extremes than are large mammals, resulting in higher average survivorship and lower origination probabilities.

The relationship between species traits and climate was investigated using global distribution data on large mammals and their ecomorphology, utilizing non-linear regression-tree analysis and linear regression. The resulting regression trees provide a reasonably accurate estimate of precipitation values for today's world, having the strongest correlations between annual precipitation and absolute and relative numbers of species, diet, tooth crown height, and diet and tooth crown height combined. These results give confidence that this methodology can be applied to palaeocommunities to estimate the past environmental conditions.

The taxon-free trait structure of mammalian communities and their relationships with tropical habitats across the world were examined using data for 169 localities. These were assigned a priori to hierarchical Olson (1983) vegetation categories, and the species were classified into dietary, locomotion and body mass groups. The resulting group structures were analysed using community structure analysis. The results indicate that the community structure significantly differs between all of the studied vegetation categories, being the highest at major and minor ecosystem levels. The results demonstrate that community structure defined by both dietary and locomotor adaptations is a powerful discriminator of tropical ecosystems and habitats across the continents, but body mass should be interpreted with caution when the research question pertains to multiple continents. As general ecosystem categories are broken down into more precisely defined habitats, more detailed knowledge of species adaptations is required to distinguish between them.

The biogeography of mammals was studied by identifying the smallest spatial scales in which the mammalian faunas are both climatically and functionally distinct in North America and Europe. These faunas were produced using k-means clustered point-sampled species occurrences of land mammals. They were tested for significant differences in climate (mean annual precipitation, mean annual temperature) and functional traits (body mass, locomotion and diet). In North America climatic differentiation exists at the scale where mammals are divided into 11 distinct faunas and, in Europe, at the scale where there are five faunas. Functional trait differentiation in body mass occurs at a larger spatial scale in North America (eight distinct faunas), but locomotor differentiation is present at all spatial scales, and dietary differentiation is not present at any scale. No significant differentiation in any functional trait at any scale is present in Europe. Since the functional traits (body mass, locomotion) are observable in the fossil record, community-level functional trait analysis has the potential to be used to reconstruct past climatic gradients.

Tiivistelmä

Tämän väitöskirjan tarkoituksena on tutkia kuinka erilaiset lajien ominaisuudet liittyvät maanisäkkäiden evolutiiviseen selviytymiseen ja kuinka taksoneista vapaa, lajien ominaisuuksiin perustuva nisäkäsyhteisöjen rakenne liittyy ilmastoon, elinympäristöihin ja eliömaantieteeseen.

Taksonien evolutiivista selviytymistä tutkittiin käyttämällä laajaa tietokantaa kenotsooisen maailmankauden fossiilisista maanisäkkäistä (New and Old Worlds - NOW, Fortelius, 2009). Analyysit tehtiin kontrolloimalla näytteiden otantaharhaa laskemalla nk. rajanylittäjien lukumäärät ja arvioimalla niiden selviytymistodennäköisyyksiä merkitse-vapauta-pyydystä uudelleen –menetelmällä (CMR). Rajanylittäjiä ovat ne suvut, jotka ovat joko olleet olemassa tai puuttuvat eri pituisilla aikaväleillä, jotka olivat 1, 1.5 tai 2 miljoonaa vuotta pitkiä. Tulokset osoittavat säännöllisen ilmiön, jossa isommilla nisäkässuvuilla ja -lajeilla on suurempi lajiutumisen ja sukupuuttoon kuolemisen tahti ja siten lyhyempi ajallinen kesto. Tätä voidaan selittää niiden lajien avulla, jotka viettävät talviunta, horrostavat tai kaivautuvat maahan, laajemmin kutsuttuna “sleep-or-hide” (SLOH) -lajeilla. Koska SLOH-käyttäytyminen on yleisempää pienillä nisäkkäillä, ovat ne siten keskimäärin enemmän suojattuja ympäristön ääriolosuhteilta kuin isommat nisäkkäät ja näin ollen kokonaisuudessaan vaikuttavat pienempien lajien korkeampaan keskimääräiseen selviytymiseen ja matalampaan lajiutumistodennököisyyteen.

Lajien ominaisuuksien suhdetta ilmastoon tutkittiin analysoimalla suurten maanisäkkäiden maailmanlaajuista levinneisyystietoa ja lajien ekomorfologiaa hyödyntämällä ei-lineaarisia regressiopuuanalyyseja ja lineaarista regressiota. Tuloksena saadut regressiopuut tuottavat kohtuullisen tarkat arviot nykyajan sadannasta, suurimpien korrelaatioiden ollessa vuotuisen sadannan ja lajien absoluuttisen ja suhteellisen lukumäärän, ruokavalion, hampaan kruunun korkeuden sekä ruokavalion ja hampaan kruunun korkeuden yhdistelmän välillä. Nämä tulokset vahvistavat, että tätä menetelmää voidaan soveltaa muinaisiin nisäkäsyhteisöihin arvioitaessa niiden ympäristöolosuhteita.

Taksoneista vapaan, lajien ominaisuuksien perusteella tehtyjen nisäkäsyhteisöjen rakennetta ja niiden suhdetta trooppisiin elinympäristöihin tutkittiin käyttämällä lajilistoja 169 yksittäisestä paikasta. Näille paikoille määriteltiin etukäteen hierarkinen Olson (1983) kasvillisuusluokittelu ja niissä olleet lajit luokiteltiin ruokavalion, liikkumismuodon ja ruumiinpainon mukaisiin ryhmiin. Näiden ryhmien rakenteita suhteessa elinympäristöön analysoitiin käyttämällä yhteisörakenneanalyyseja. Tulosten mukaan yhteisörakenteet ovat tilastollisesti merkitsevästi erilaisia tutkittujen kasvillisuusluokkien välillä, ollen suurin ekosysteemitasolla. Tulosten perusteella voidaan todeta, että nisäkäsyhteisöt jotka perustuvat ruokavalio- tai liikkumismuotoryhmiin, erottelevat hyvin mantereiden trooppiset ekosysteemit sekä elinympäristöt, mutta ruumiinkokoon perustuvaa ryhmittelyä kannattaa tulkita varovaisemmin, mikäli useita eri mantereita käsitellään tutkimuksessa. Mitä tarkempiin määriteltyihin elinympäristöihin yleiset ekosysteemiryhmät halutaan pilkota, sitä yksityiskohtaisempaa tietoa lajien ominaisuuksista tarvitaan, jotta eri elinympäristöt voidaan erottaa toisistaan.

Nisäkkäiden eliömaantiedettä tutkittiin erottamalla ne Pohjois-Amerikan ja Euroopan pienimmät mahdolliset maantieteelliset alueet, joissa niiden lajistot ovat sekä ilmastollisesti että lajiominaisuuksiltaan erilaiset. Nämä alueet luotiin jakamalla maanisäkkäiden levinneisyystiedot K-

keskiarvo klustereiksi – lajistoiksi. Tuloksena saatujen lajistojen tilastollista eroavaisuutta testattiin käyttämällä levinneisyystiedoista riippumattomia ilmaston (vuotuinen sadanta, vuotuinen lämpötila) ja lajiominaisuuksien (ruumiinpaino, liikkumismuoto, ruokavalio) välisiä eroavaisuuksia. Tulosten perusteella Pohjois-Amerikassa lajistot ovat ilmastoltaan erilaisia mittakaavassa, jossa nisäkkäät on jaettu 11 erilliseen lajistoon kun taasen Euroopassa ne jakautuvat alueellisesti viiteen erilaiseen lajistoon. Ruumiinpainon erojen perusteella Pohjois-Amerikan lajisto jakautuu maantieteellisesti laajempiin alueisiin, kahdeksaan kokonaisuuteen, kun taasen liikkumismuodon perusteella lajistot eroavat kaikissa jakoluokissa. Ruokavalion perusteella ei lajistojen välille muodostunut tilastollisia eroja millään jaolla. Euroopassa yksikään lajisto ei eronnut tilastollisesti toisistaan minkään lajiominaisuuden perusteella millään jaolla. Tutkimuksessa saatujen tulosten perusteella voidaan todeta, että niitä havaittavia lajiominaisuuksia, joita voidaan erottaa myös fossiiliaineistosta (ruumiinpaino, liikkumismuoto) voidaan käyttää nisäkäsyhteisöjen analysointiin ja muinaisten ilmasto-olosuhteiden mallintamiseen.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Mikael Fortelius and Jussi Eronen, who guided me through this journey. Mikael gave me the inspiration to study mammalian communities at broader scales when in 1997 I had my first glance at the NOW database. Now, 21 years later, I am still learning new methods to study this subject. Mikael taught me that ignorance is something one should not consider as an obstacle but an opportunity to learn new things. He also gave me many good pieces of advice and provided me great opportunities to collaborate with other researchers in the field – too many people to list here. I thank Jussi for his patience while supervising my writing and giving instructions. From Jussi, I learnt new and practical ways to conduct a scientific study that results in a published article.

I am grateful to my main co-authors, Lee Hsiang Liow, Kai Puolamäki, Fire Kovarovic, Jussi Eronen and David Polly. During our work, Lee Hsiang showed how well-formulated ideas can be put into the format of an article quite swiftly. She also provided helpful comments on my tropics manuscript. From Kai, I gained ideas on how to analyse and study mammalian community data, while from David I received good support in interpreting the results of our article. From Jussi, I received valuable help while writing our final manuscript. During the years, I have visited Fire in Durham several times while working on manuscripts we have already published or are going to publish. I am looking forward to our forthcoming co-operation.

I also want to thank my other co-authors: John Damuth, Christine Janis and Liu Liping for the precipitation articles and Mikael Fortelius, Ella Bingham, Larry Flynn, Heikki Mannila and Nils Christian Stenseth for the sleep-or-hide articles. During the writing and preparation of these articles, I learned a lot about scientific teamwork, writing and reasoning.

During my studies, I had the privilege to collaborate with the IMPPS (Integrating Macroecological Patterns and Processes across Scales) Research Coordination Network in Sante Fe. Funded by the US National Academy of Sciences and hosted by Felisa Smith, Morgan Ernest and Kathleen Lyons, the IMPPS project resulted in many significant publications not included in this thesis. For the many memorable occasions, I thank Felisa, Morgan and Kathleen.

I thank Richard Sibly and James H. Brown for the opportunity to give a talk on our sleep-or-hide results at the British Ecological Society’s Annual Meeting, and Terry Harrison and Mikael Fortelius for the chance to give a presentation at an international conference in University College London. In addition, I wish to thank the Björn Kurtén club and Janina Rannikko, the latest chair of the club, for the possibility to present our published work.

Thank you, Juha Saarinen and Pierre Mauriès, ‘Junior’ and ‘Mr. Pierre’, for being such great travel companions. I really enjoyed sharing thoughts and the adventures we had. I thank Tomi Luoto, Elina Sahlstedt and Ilona Romu, my co-students while we were sharing the same office at the department. I also thank the fellow palaeopeople: Aleksis Karme for his innovative ideas and support as a peer during this long process; Anu Kaakinen, Liu Liping, Majid Mirzaie Ataabadi, Jussi Eronen, Ferhat Kaya, Diana Pushkina, Juha Saarinen and Leena Sukselainen for the enjoyable moments at the department and for their encouraging example of how to finalise a PhD dissertation. Hannes Heikinheimo introduced me to data clustering methods, which proved to be valuable knowledge in many of my works, while Indrė Žliobaitė gave me general advice on analytical methods.

I thank the personnel of the Department of Geosciences and Geography for the great working environment and companionship. I am grateful to Mia Kotilainen, Seija Kultti and Tino Johansson, who patiently helped me with study-related matters. Thank you, Mikko Haaramo, for the support with technological issues.

I am thankful to Björn Kröger, who gave me the opportunity to do my internship at the Finnish Museum of Natural History at Kumpula. I enjoyed our co-work and the beautiful atmosphere at Kumpula Manor.

I am thankful to Tapani Rämö, who recommended me to participate in the GeoDoc Writing Workshop held at the Iloranta resort, Hauho. I gained a significant boost in finalising my work during those days. Tapani also provided valuable comments on the thesis. I am grateful to my pre-examiners, Professor Lars Werdelin from the Swedish Museum of Natural History and Associate Professor Pasquale Raia of the University of Naples, for their valuable input in improving this manuscript. Julia Lintulaakso finalised the layout of this thesis, for which I am grateful.

Funding for my research came from the Emil Aaltonen Foundation, Waldemar von Frenckell Foundation and Nordenskiöld Society in Finland, for which I am grateful. I also received a Chancellor's travel grant, which helped me to collaborate with the scientific community abroad.

I thank all my friends and relatives who on regular basis have asked after my progress. Last, but not least, I want to thank my family, Arja, Roni, Julia, Joonas and Oskari, for their time and support during all these years.

"Put it in your database!"

Christoph Zollikofer

Contents

List of original publications

This thesis is based on the following publications:

I. Liow, L.H., Fortelius, M., Bingham, E., Lintulaakso, K., Mannila, H., Flynn, L. and Stenseth, N.C., 2008. Higher origination and extinction rates in larger mammals. *Proceedings of the National Academy of Sciences,* 105(16), pp.6097-6102.

II. Eronen, J.T., Puolamäki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C. and Fortelius, M., 2010. Precipitation and large herbivorous mammals I: estimates from present-day communities. *Evolutionary Ecology Research,* 12(2), pp.217-233.

III. Lintulaakso, K. and Kovarovic, K., 2016. Diet and locomotion, but not body size, differentiate mammal communities in worldwide tropical ecosystems. *Palaeogeography, Palaeoclimatology, Palaeoecology,* 454, pp.20-29.

IV. Lintulaakso, K., Polly, P.D. and, Eronen, J.T. 2018. Land mammals form eight distinct biogeographical clusters in North America but only one in Europe. *Journal of Biogeography (*submitted).

The publications are referred to in the text by their roman numerals.

Author's contribution to the publications

K. Lintulaakso, M. Fortelius and J.T. Eronen planned the research project.

I. L.H.L., M.F., H.M., and N.C.S. designed the research; L.H.L., M.F., E.B., **K.L.**, and L.F. performed the research; L.H.L., E.B., **K.L.**, H.M., and L.F. analysed the data; and L.H.L., M.F., L.F., and N.C.S. wrote the paper.

II. The study was planned by J.T.E. and K.P. The data were collected by **K.L.** and J.T.E., and analysed by K.P. The results were interpreted by all the authors. All the authors wrote the article, with the main responsibility held by J.T.E.

III. The original idea for the study came from M.F., and it was planned by **K.L.** and K.K. The data were collected and analysed by **K.L.** and both authors interpreted results. Both authors wrote the article, with the main responsibility held by **K.L.**

IV. The study was planned by P.D.P and J.T.E. The data were collected by **K.L.** and J.T.E., and analysed by P.D.P. and **K.L.** All authors interpreted the results. The article was written by all the authors, with the main responsibility held by **K.L.**

1. Introduction

For palaeoecological analyses, the relationship between an animal and its habitat can be divided into two distinct niches: spatial and trophic. The spatial niche relates to the physical space in which a species sleeps, moves, reproduces and eats in during its life cycle, while the trophic niche relates to the feeding ecology a species exploits to meet its nutritional and energy requirements. During evolution, each species has adapted both physically and behaviourally to its habitat and available resources: it fills a particular niche. Some of these adaptations can be identified and traced back to niche exploitation. Thus, they can be used for habitat reconstruction (Kovarovic et al., 2017).

From a palaeontological point of view, behavioural adaptations are seldom identifiable in fossil material. However, the preservation of skeletal and dental remains offers a vast resource of study material from the past that can be related back to spatial and trophic niche exploitation. Of the postcranial remains, particularly elements of the fore and hind limbs provide information on the locomotion method and substrate to which the species is adapted. Skeletal limbs show correspondence between the locomotor type and morphological indices (Van Valkenburgh, 1987) and can be used to assess the association between limb morphology and locomotor habits (e.g., Brown and Yalden, 1973; Van Valkenburgh, 1987; Kappelman, 1988; MacLeod and Rose, 1993; Samuels and Van Valkenburgh, 2008). Teeth, in comparison to postcranial elements, provide information on the diet of a species and the trophic niche to which it is adapted and for which dental morphologies vary. Carnivorous, insectivorous, frugivorous and herbivorous diets require different dental morphologies. A good example is the tooth crown height (hypsodonty) of large herbivorous mammals. This is used as a measure of dental durability. As mammals have only a single adult dentition, abrasive material such as grass or dusty and gritty plants would wear low-crowned teeth down before the end of the natural lifespan of a species. To extend the useful life of teeth, the height of the tooth crowns has evolved for species that have a more abrasive diet. Usually, species that mostly eat soft leafy material have relatively low tooth crowns (brachyodont), while species eating abrasive material have higher molars (hypsodont). Generally speaking, hypsodonty is informative ecologically because it is well developed in mammals eating fibrous and abrasive foods, which are most abundant in open and generally seasonally dry environments, while species that have low cheek teeth live in more humid, closed environments (van Valen, 1960; Janis and Fortelius, 1988; Jernvall and Fortelius, 2002). A second example is quite intuitive: tropical and subtropical rainforests or similar biomes support more purely frugivorous mammals. These areas have a high annual average temperature and precipitation, where fruit resources are available year-round (Kay and Madden, 1997; Hanya et al., 2011; Pineda-Munoz et al., 2016). As for postcranial remains, many studies have been conducted to infer the relationship between dental morphologies and diets (e.g., Janis, 1990; Kaiser et al., 2013).

Besides locomotor adaptation and substrate utilisation or the dietary preferences of a species, body mass has been considered as an ecologically relevant variable for palaeoecological analyses. Body mass is often said to be one of the most fundamental biological characteristics of mammals, correlating with many other variables such as the metabolic rate (Kleiber 1932, 1947), home range (Lindstedt et al. 1986) and longevity (Schmidt-Nielsen 1984). It also imposes physiological restrictions, which influence the space utilization, locomotor adaptation, diet or the geographical distribution of the species. Skeletal and dental remains provide a good source for body size estimations of fossil species (see Damuth and MacFadden 1990 for several examples).

Palaeontologists have long used the abovementioned species traits for autecological studies, dealing with an individual species in relation to its environment (e.g., Hutchinson, 1959). However, synecological studies, which deal with the structure, development and distribution of ecological communities, have only been carried out since the late 1970s, when the work of Andrews et al. (1979) brought the taxon-free approach to the field of palaeontological community structure analysis. These authors published a series of histograms organized according to five broad habitat types of three continents and showed that habitats with a similar array of available niches have similar community structure profiles, regardless of their geographical location. By using specific dietary, locomotion and body mass groups, there is no longer a need to have a full taxonomy of the fossil species. Instead, there is a taxon-free classification, which is freed from geographical location and geological time. By having these abstract groups, different time series and locations can be examined at the same analytical level, allowing the investigation of how different environments evolve or how similar they are.

Using this theoretical framework to reconstruct environments from the mammalian community structure, several aspects of the environments and the community structure itself, as well as species traits, can be studied. The goal of this thesis study was to address the following questions: how does body size relate to the evolutionary survivorship of taxa in a changing environment, and are there other species traits that might explain these survivorship patterns? (Paper I); can mammalian communities be used to form boundary values for palaeoclimatic variables such as precipitation? (Paper II); can tropical habitats across the world be differentiated by using the taxon-free mammalian community structure, and what are the implications of this type of analysis for palaeoecological community studies? (Paper III); and at which level the mammalian faunas are both climatically and functionally distinct in North America and Europe, and can they provide community-level functional trait analysis that has the potential to be used to reconstruct past climatic gradients (Paper IV)?

2. Material

For the four papers that constitute this thesis, the author compiled a comprehensive database of mammals (MammalBase; Lintulaakso, 2013). This is based on hundreds of published sources of ecomorphological attributes, diets, species distributions and locality-based occurrences of mammals. The main taxonomy is from Wilson and Reeder (2005), which is linked to other data sources by managing the different nomenclatures, synonyms and spelling variations between them (e.g. Lawing et al., 2017).

The temporal coverage in these papers extends from the distant past (Paper I) to recent time (Papers II, III, IV), and spatially the species occurrence data covered large areas: Europe in Paper I, the global scale in Paper II (excluding the Indo-Pacific area and Australasia), the tropics in Paper III and continental North America and Europe in Paper IV. The species occurrence data were in point data format (Papers I, III and VI) and in grid format (Paper II).

Data on species occurrences were obtained from several sources:

- WWF WildFinder, the World Wildlife Fund's species distribution dataset (World Wildlife Fund, 2006), Paper II;

- the NOW database (New and Old Worlds, Fortelius, 2009), Paper I;
- the database compiled by the National Center for Ecological Analysis and Synthesis (NCEAS) Workshop on Mammalian Communities (Badgley et al., 2001; van Dam et al., 2001; Damuth et al., 2002), Paper III;
- previously published studies, including Andrews et al. (1979), Andrews and Humphrey (1999), Reed (1997) and Reed (2008), Paper III;
- information from WDPA, the World Database on Protected Areas (IUCN and UNEP, 2009), was used to locate more specific coordinates for some species occurrence localities, Paper III;
- the distribution ranges for mammals in North America came from NatureServe (Patterson et al., 2005), Paper IV;
- the distribution ranges for mammals in Europe came from Societas Europaea Mammalogica (the Atlas of European mammals, EMMA; Mitchell-Jones et al., 1999), Paper IV.

The taxonomic setting consists of mammals mainly from the orders Artiodactyla, Carnivora, Creodonta, Hyracoidea, Insectivora (and the new orders previously included in it), Lagomorpha, Perissodactyla, Primates, Proboscidea, Rodentia and Tubulidentata. In addition, the following taxa and species groups were excluded: marine species from the orders Cetacea and Sirenia, the families Odobenidae, Phocidae and Otariidae, as well as those species coded as aquatic in the database. The order Chiroptera is also excluded. This follows a well-established practice of excluding these groups in palaeocommunity analyses due to the practical difficulties with including them as fossil species; they do not frequently fossilise, or they are fossilised under very different taphonomic conditions compared to other land-dwelling or larger mammals. Given this, and their rarity in the fossil record, most workers thus exclude them (e.g. Andrews et al., 1979; Reed, 1998; Kovarovic et al., 2002; Mendoza et al., 2004; Louys et al., 2011). In addition, some body size restrictions were set (large herbivores in Paper II, >500 g in Paper III).

For ecomorphology, several species traits were used in the analyses: body mass (Papers I, II, III, and IV), diet (Papers II, III, and IV), tooth crown height (Paper II), locomotion (Papers III and IV) and behaviour such as hibernation, torpor and burrowing, collectively termed "sleep-or-hide" (SLOH) behaviour (Paper I).

Data for these ecomorphological traits come from the NOW database (New and Old Worlds, Fortelius, 2009), MammalBase (Lintulaakso, 2013) and from previously published datasets from NCEAS (Badgley et al., 2001; van Dam et al., 2001; Damuth et al., 2002) and PanTHERIA (Jones et al., 2009).

Besides species occurrence and trait data, climatological and environmental information was obtained: for Papers II and IV, climatological data from Hijmans et al. (2005; www.worldclim.org) was used. In Paper III, a 50-km grid data was assigned to the vegetation classification system from the "Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation" GIS dataset (Olson et al., 1983, 1985), and in Paper IV to Bailey's ecoregion categorization (https://www.fs.fed.us/rm/ecoregions/products/map-ecoregions-continents/; Bailey & Hogg, 1986; Bailey, 1989).

3. Methods

The methods varied according to the study settings for each paper. Generally, the methods can be divided into three categories: classical statistical analyses, community analyses and data mining.

Several "classical" analyses were used in the papers. In Paper I, the Kolmogorov–Smirnov test and Mann–Whitney U-test were used to assess whether small genera have greater mean and median durations than larger ones. In Paper II, a linear regression was used to predict the annual precipitation (mm) using absolute and relative numbers of species, body size, body mass, diet and tooth crown height as covariates or predictor variables. In Paper IV, one-way ANOVA (Krzanowski, 1988; Rencher, 2002) was used to analyse whether the generated clusters differed significantly ($p < 0.01$) in the values of the climate variables.

The community analyses included capture–mark–recapture (CMR, Connolly & Miller, 2001), which was used in Paper I to control for sampling biases instead of classical estimation of the population size per se. In Paper III, several methods were used to examine whether there are differences between communities and their environmental settings. These included permutational multivariate analysis of variance (PERMANOVA or NPMANOVA; Anderson, 2001, 2005), canonical analysis of principal coordinates (CAP, Anderson and Willis, 2003) and similarity percentage analysis (SIMPER, Clarke, 1993).

Some of the methods that were used are common in data mining. Quite often, the data are multivariate, and in some cases the method does not, for example, assume that the covariates are independent. In Paper II, nonlinear regression-tree analysis was used to estimate the relationships between the environment (precipitation) and the characteristics of mammalian communities. In Paper III, non-metric multidimensional scaling (NMDS) was used for the preliminary analyses, while in Paper IV, k-means clustering was used to group mammalian distribution data into coherent entities.

4. Summary of the original publications

4.1 Paper I

Higher origination and extinction rates in larger mammals

Paper I focuses on the question of how body size is related to the evolutionary survivorship of taxa in a changing environment, and whether there are other species traits that may explain these survivorship patterns. The study used a large fossil dataset of mammals from the Neogene of the Old World (NOW, Fortelius, 2009). By controlling for sampling biases, calculating the per capita origination and extinction rates of boundary crossers (genera that were recorded as present or absent during time bins that span 1, 1.5 or 2 M.y.) and estimating survival probabilities using capture–mark–recapture (CMR) methods, this study found the recurring pattern that large mammal genera and species have higher origination and extinction rates, and therefore shorter durations. This pattern is surprising in the light of molecular studies. Although the shorter generation times and higher metabolic rates of small mammals may increase the molecular rates of evolution, and even if molecular changes translate substantially to phenotypic changes, these shorter-term, generational changes do not appear to translate to higher turnover rates. To explain the observed pattern, a nearest living relative approach was used to code behaviour such as hibernation, torpor and burrowing, collectively termed "sleep-or-hide" (SLOH) behaviour, for the fossil taxa. The patterns in the data indicate that the ability of mammals to shield themselves from environmental fluctuations lowers their turnover rates, and that the ability to "sleep or hide" is constrained by body size. SLOH behaviour is more common in some small mammals, and, as a result, SLOH small mammals contribute to the higher average survivorship and lower origination probabilities among small mammals.

4.2 Paper II

Precipitation and large herbivorous mammals I: estimates from present-day communities

Paper II presents a study on the relationship between climate (precipitation) and mammalian community characteristics. Previous studies reconstructing past environmental conditions have mainly been based on plant data (e.g. Leaf Margin Analysis, Wolfe, 1979; Climate Leaf Analysis Multivariate Program, Wolfe, 1993; Coexistence Approach, Mosbrugger and Utescher, 1997; Climate Amplitude Method, Fauquette et al., 1998), while data from large fossil mammals have not extensively been used for quantitative estimates of past environmental conditions. Fortelius et al. (2002) used the fossil teeth of terrestrial plant-eating mammals as a quasi-quantitative proxy for environmental aridity. The mean hypsodonty value was calculated for each locality by averaging the ordinated scores for brachyodont (1), mesodont (2) and hypsodont (3) teeth, and the mean values were plotted on maps using interpolated colours to indicate regional differences. Paper II follows the idea of a connection between tooth hypsodonty and environmental aridity. It examines how mammalian community characteristics can be used to estimate regional precipitation with global distribution data for large mammals and their ecomorphology (tooth hypsodonty) by using nonlinear regression-tree analysis and linear regression.

The resulting regression trees provide a reasonably accurate estimate of precipitation values for today's world, having the strongest correlations between annual precipitation and absolute and relative numbers of species, diet, tooth crown height, and diet and tooth crown height combined. These results give confidence that this methodology can be applied to palaeocommunities to estimate past environmental conditions.

4.3 Paper III

Diet and locomotion, but not body size, differentiate mammal communities in worldwide tropical ecosystems

Paper III focuses on the relationships between the taxon-free mammalian community structure and the physical environment in which they exist. Previous studies have demonstrated that there is a relationship between the mammalian community composition and the tropical environment, especially with the amount of tree cover (Harrison, 1962; Andrews et al., 1979; Reed, 1997, 1998; Louys et al., 2011; Andrews and Hixson, 2014); this also relates to the amount of precipitation. The aim of this paper was to investigate whether tropical habitat groups of the world can be differentiated by using the taxon-free mammalian community structure and to discuss the implications of the analysis for palaeoecological community studies. Mammalian community data were used on 169 modern tropical localities from four continents: the Americas, Africa, Asia and Australia. These localities were assigned to a standardised vegetation classification (Olson, 1983) and 607 species over 500 g were assigned to four ecological categories: body mass, locomotor pattern and two hierarchical dietary categories. These data were obtained from MammalBase, a database of species attributes and diets based on hundreds of published sources, compiled by K. Lintulaakso (Lintulaakso, 2013). Additional data came from NCEAS (Badgley et al., 2001; van Dam et al., 2001; Damuth et al., 2002) and PanTHERIA (Jones et al., 2009). The mass categories were modified from those in Andrews et al. (1979): 0.5–8 kg (A), 8–45 kg (B), 45–90 kg (C), 90–180 kg (D), 180–360 kg (E) and >360 kg (F). The species were divided into six locomotor groups that describe the substrate(s) in which the species moves (Reed, 1998; Miljutin, 2009): arboreal (A), arboreal–terrestrial (AT), subterranean (S), subterranean–terrestrial (ST), terrestrial (T) and terrestrial–aquatic (TA). The dietary variable was divided into two levels of food resource preference. Diet Level 1 represented the broadest categories of feeding strategy: animalivorous, frugivorous and herbivorous (Miljutin, 2009). Diet Level 2 followed that of Eisenberg (1981) with some modifications: mixed-feeders were added to the herbivore group, and aerial insectivores and foliage-gleaning insectivores were combined into one category: insectivores. The resulting communities were analysed using community structure analyses. According to the results, the mammalian community structure significantly differed between all of Olson's categories, being highest at Olson's major and minor ecosystem levels. As general ecosystem categories are broken down into more precisely defined habitats, more detailed knowledge of the species adaptations is required to distinguish between them. The results also demonstrate that community structure defined by both dietary and locomotor adaptations is a powerful discriminator of tropical ecosystems and habitats across the continents, but body mass should be interpreted with caution when the research question pertains to multiple continents.

4.4 Paper IV

Land mammals form eight distinct biogeographical units in North America, but only one in Europe

In Paper IV, the mammalian faunas that are both climatically and functionally distinct in North America and Europe were studied. The point-sampled species occurrences of land mammals were clustered using the k-means algorithm, ranging from three clusters at the largest spatial scale and 21 at the smallest. After clustering, the faunas were tested for significant differences in climate (mean annual precipitation, mean annual temperature) and functional traits (body mass, locomotion and diet). As a result, in North America climatic differentiation exists at the scale where mammals are divided into 11 distinct faunas and, in Europe, at the scale where there are five faunas. Functional trait differentiation in body mass occurs at a larger spatial scale in North America (eight distinct faunas), but locomotor differentiation is present at all spatial scales, and dietary differentiation is not present at any scale. No significant differentiation in any functional trait at any scale is present in Europe. Faunal clusters can be constructed at any spatial scale, but clusters are climatically and functionally meaningful only at larger scales, producing entities that form regional species pools for community assembly processes. Since the functional traits (body mass, locomotion) are observable in the fossil record, community-level functional trait analysis has the potential to be used to reconstruct past climatic gradients.

5. Discussion

Paper I shows that among mammal genera with shorter taxon durations, there is a predominance of large mammals (Artiodactyla, Carnivora, Creodonta, Hyracoidea, Perissodactyla, Primates, Proboscidea and Tubulidentata), whereas among those with longer durations, there is an excess of small mammals (Insectivora, Rodentia and Lagomorpha). Large mammal genera have higher per capita rates of both origination and extinction than small mammals. By using a capture–mark–recapture (CMR) approach, the results indicate that large animals have comparatively lower survivorship. The best models in each case from CMR analyses always include a body size effect, indicating that large mammals have a lower probability of survival. Although the shorter generation times and higher metabolic rates of small mammals may increase the molecular rates of evolution, and even if molecular changes translate substantially to phenotypic changes, these shorter-term, generational changes do not appear to translate into higher turnover rates, as demonstrated here. Body size is indeed a general predictor of mammal taxon duration (this study), even though there are certainly small short-duration and large long-duration mammal taxa.

In this study, a hypothesis was presented that it is not just being smaller per se that allows greater survivorship (many small mammals also have short species and genus durations). Rather, a greater proportion of small mammals may be more apt at hiding (e.g., in burrows or tree holes) or going into low metabolic rate modes (e.g., hibernation and torpor). Taxa that demonstrate the ability to buffer themselves from environmental stress in this manner are here termed SLOH (“sleep-or-hide”) taxa. A nearest living relative approach was used to infer the presence/absence of SLOH behaviour in fossil genera. According to the results, genera having SLOH behaviour in the dataset generally have smaller body mass estimates. There is also a group of small mammals that does not display SLOH behaviour. Therefore, some, but not all, small mammals may be better buffered against environmental changes than larger mammals. Based on the analyses in this study, there is good evidence that the longest-lived small mammal genera have SLOH traits. These results suggest that SLOH genera are less extinction prone and experience less selective pressure to evolve to cope with environmental change. Whether species are seasonal or facultative in the hibernation–torpor continuum, this behaviour is associated with decreased food resources, lower ambient temperatures and lowered oxygen availability. Because this behaviour is in part plastic, species that display it are expected to cope better with (un)expected climate variation.

In Paper II, a method was developed to estimate regional precipitation from mammalian community characters using non-linear regression-tree analysis and linear regression between annual precipitation and ecomorphologies. According to the results, a non-linear algorithm better estimates precipitation than a linear correlation. The strongest correlations were observed for annual precipitation versus diet (R2 = 0.665), precipitation versus tooth crown height (R2 = 0.658), and precipitation versus diet and tooth crown height combined (R2 = 0.742). Many of the regional and all of the global-scale patterns are well represented in the predictions. Africa in particular, and both the Eurasian continent and the New World are relatively well estimated.

Mammal communities reflect the vegetation habitat, which in turn reflects the patterns of rainfall. The causal chain between rainfall and mammal abundances goes through grass and foliage production, which is approximately linearly related to rainfall (Huxman et al., 2004). The annual net primary production (ANPP) in world biomes is related to rainfall. However, at some stage, rainfall is not the main limiting factor. At high northern latitudes, primary production is controlled by temperature and light more than by precipitation (Churkina and Running, 1998; Nemani et al., 2003), and there, the relationship between precipitation and mammalian community traits become less clear. The production of arid environments is more rainfall limited, whereas in environments that are more humid, production is temperature or nutrient limited. This can also be seen from the model. Mammalian communities cannot be confidently used to estimate rainfall above ~2000 mm, representing the present-day habitats of tropical forests (with rainfall exceeding 1500 mm) and tropical rainforests (over 2000 mm). For example, in South America, tropical rainforest is especially troubling for hypsodonty estimation. However, combined hypsodonty and diet estimation captures much of the precipitation variation in South America. In addition to certain rainfall thresholds, some other areas had poor estimates for rainfall: Madagascar and areas having orographically driven changes in precipitation, such as Alpine regions of Europe, the highlands of central India and the South Himalayan region. The likely reason for the poor estimates for Madagascar is that it has almost no terrestrial herbivores. During the last 10,000 years, large terrestrial and arboreal species have gone extinct in Madagascar due to excessive human hunting and human-induced environmental change (Burney et al., 2004). For the high-altitude areas, a likely reason is that mammals are mobile and track the seasonal variation in mountains, which causes mammal species to be recorded in multiple adjacent ecoregions. Perhaps the method is too coarse to detect orographically driven changes in precipitation.

Generally, the estimated annual precipitation values based on regression trees using absolute and relative numbers of species, hypsodonty, diet, and hypsodonty and diet depict the actual annual precipitation pattern remarkably well. The correlation values are larger when the full set of absolute and relative numbers are used in comparison with relative numbers only. Africa shows the best correspondence between predicted and actual rainfall patterns. This is significant for the potential use of this method to estimate climatic conditions of the past. As Africa was the continent least affected by end-Pleistocene megafauna extinctions (Koch and Barnosky, 2006), it has the best available approximation of the mammal community structure that could be thought to represent pre-anthropogenic conditions.

Since the publication of Paper II, species trait-based models for estimating climatic variables have been developed by other researchers (e.g., Liu et al., 2012; Žliobaitė et al., 2016; Fortelius et al., 2016) using other functional traits of herbivore molar surfaces than hypsodonty. The longitudinal loph count has been used to extract estimates of rainfall and temperature (Liu et al., 2012), FCT (Functional Crown Types, which includes seven different dental traits) has been used for modelling local precipitation, temperature, primary productivity and a vegetation index as functions of these traits (Žliobaitė et al., 2016), while hypsodonty and longitudinal lophedness have been used in regression models for separately estimating temperature and precipitation (Fortelius et al., 2016).

Paper III presents an investigation into whether tropical habitats across the world can be differentiated on the basis of their mammalian community structure. Olson's classification was used to categorise the habitats. The results demonstrate that the diet and locomotion categories differentiate the communities best, while body mass performed poorly.

The habitat differences using dietary categories are largely accounted for by the number of frugivorous mammals (frugivore–granivores and frugivore–herbivores). These are mainly primates, but there are also other frugivores, such as ungulates. The number of herbivores is only important at Olson's minor ecosystem and vegetation level, where "interrupted woods" and treeless areas such as grasslands and tundra are differentiated from each other. Animalivorous groups do not impact on the differences between major or minor ecosystems. Herbivores and frugivores are mainly primary or secondary consumers, while animalivores are higher in the trophic hierarchy. This may be the reason why habitats are reflected more by their primary and secondary consumers and the 'signal' from primary production weakens towards the top of the trophic hierarchy. The results show that the broader dietary classification (Diet Level 1, animalivores, herbivores and frugivores) may make it a more robust approach to worldwide ecosystem and vegetation comparisons, although clearly less detail is captured. More nuanced differences between habitats can be obtained using Diet Level 2 (Eisenberg, 1981, see summary of Paper III). This requires between one and six dietary variables.

In terms of the locomotion groups, the number of arboreally adapted species differentiates closed habitats from open ones. This is a conclusion borne out by others (Reed, 1997; Louys et al., 2011), who have also demonstrated that an abundance of arboreal species indicates the presence of well-developed tree cover. These taxa are generally primates and tree-dwelling rodents. Interestingly, the number of terrestrial species contributed very little to the differences between community structures. Terrestriality may be too general a category for describing species that move about on the land surface. By adding more detailed locomotor classes (i.e. cursorial), the separation between open and closed habitats could be clearer. Subterranean–terrestrial and terrestrial species distinguished humid and forested localities from seasonal or less forested areas. Seasonality and the inconsistent availability of food throughout the year can be dealt with by migration, which requires terrestrial movement, or by many of the "sleep-or-hide" (SLOH) behaviours described in Paper I. "Sleeping" (hibernation, torpor, estivation) provides a means for escaping harsh conditions, such as seasonality, for up to several months, while "hiding" (living underground, taking shelter in caves or tree holes) can provide a more secure environment to live in, stabilising some of the climatic seasonality. Food caching is also one of the traits that is regularly used by SLOH species in order to cope with seasonal changes in food availability.

Body mass categories did not clearly differentiate community structure, particularly in vegetation categories. However, further analysis of the data showed that the size ranges of mammals differed on each continent. Africa has a higher proportion of larger mammals, whilst America and Australia have smaller species, with Asia somewhere in between. This pattern is also observable in the palaeontological record of the Pleistocene and is likely to have roots deep in time (Janis, 1993; De Vivo and Carmignotto, 2004; Louys et al., 2011; Fortelius, 2013; Owen-Smith, 2013). Although a recent study (Rodríguez et al., 2006) has shown that body size can be useful, this is only clear at higher latitudes, suggesting that the warmer the annual temperature, the less likely body size is to be helpful in distinguishing habitats.

The study in Paper III confirmed that palaeoecological reconstructions of tropical fossil localities based on dietary and locomotor variables can be obtained. The results indicate only relatively small numbers of similar mammalian community structures for worldwide tropical vegetation classes. This indicates that historical and ecological factors shape the mammalian community structure slightly differently in relatively similar primary production environments (Endler, 1982). Some vegetation groups represent a continuum of cover that is, at least in the worldwide comparison, too detailed to differentiate when broad generalities are sought. However, when the use of well-defined broad habitat categories is pursued in palaeoecological community studies, three worldwide tropical major ecosystems could be used: "humid, closed forests", "seasonal or interrupted forests and grasslands" and "seasonal, open drylands".

In Paper IV, the study investigated the findings of Heikinheimo et al. (2007 and 2012), who demonstrated that mammalian distribution data for Europe naturally divide into clusters that are spatially highly connected. These findings were investigated in Paper IV between two continents: North America and Europe. The mammalian presence/absence data were clustered using the k-means algorithm. The resulting clusters, which were solely based on the mammalian distribution data, formed coherent groups — termed as faunas. Using climatic variables that were independent of the clustering itself showed that these groups were significantly different in their climate for up to eleven faunas in North America and five in Europe. When these faunas were analysed using their functional trait data (body mass, locomotion, diet), they differed significantly in North America. The functional trait differentiation in body mass occurred at a broader spatial scale in North America (eight distinct faunas), while the locomotor differentiation was present at all spatial scales. Dietary differentiation was not present at any scale. No significant differentiation in any functional trait at any scale was present in Europe.

The lack of trait differentiation among European faunas is an interesting result. Europe's narrower climate range probably explains why its faunas do not show statistically significant differentiation in body mass. North America has a wider range of mean annual temperature (-26°C to 29°C) and annual precipitation (54 mm to 4860 mm) and fills a broader climate space than Europe (-9.7°C to 18.2°C, 242 mm to 2331 mm, respectively). Only three North American faunas overlap climatically with European ones: the Eastern US, which overlaps with France and Southern Scandinavia–UK, and Southern Canada and British Columbia, which overlaps with Northern Scandinavia–Finland. The remaining eight North American faunas lie outside the climatic boundaries of the European faunas, forming three unique groups: warm and moist, dry, and cold. The greater climate gradient of North America may be necessary to sort species into faunas based on body mass. In Europe, the reduced climate range and the smaller number of species decrease statistical power to detect body mass differences. Similarly, a narrower variety of habitats in Europe probably does not facilitate locomotor sorting among the faunas. Tropical, desert, and basin and range environments are missing entirely from Europe. The dispersion of faunas in climate space is also greater in North America, suggesting that even though the climatic differences in the European faunas are statistically significant, they are smaller. Therefore, the strong sorting effects imposed by extremely different North American habitats like dense tropical forests, open desert and scrubland, grasslands, and large expanses of taiga may be absent in Europe, thus explaining the lack of locomotor diversification among faunas in the latter.

In North America, locomotor categories differ at comparatively small spatial scales, forming a hierarchy that creates significant differences between faunas at small scales from k=21 to large scales at k=4. Body mass differed at a comparatively large spatial scale where k=8. The proportion of large (> 8,000 g) species varied substantially among these functionally distinct faunas, being lowest (5% of the faunal composition) in the south (Mesoamerica) and highest in the north (High Arctic Canada, Eastern Beringia, and Northern High Canada, >24%). Diet did not differentiate faunas at any scale. The lack of differences may be because the dietary categories were too detailed (Lintulaakso & Kovarovic, 2016), but is more likely because all types of diet are likely to be mixed within local communities whereas body mass and locomotor specialisations have a functional relationship to climate or landscape conditions that vary geographically. These findings suggest a hierarchy of processes involved in the formation of regional species pools and local community assembly. Functional locomotor traits associated with mobility and thus the ability to colonise local communities are differentiated at a smaller scale that is subequal to physiographic differences. However, body mass, which is associated more with temperature and openness of habitat, is differentiated at a larger scale. The lack of body mass differentiation in Europe is consistent with this hypothesis because the scale of climatic differentiation is less there. However, the absence of locomotor differentiation among European faunas is puzzling since in North America that differentiation is found at almost all scales.

The aim of this study was not to create formulae of any kind between taxon-free species composition of a locality and the climate, as in Paper I. This should be and is in some part already being investigated in other studies (see Žliobaitė et al., 2016). However, the hierarchical distribution of faunas, climate, and functional traits provides a framework for interpreting paleontological faunas in terms of climate. Faunal clusterings based on a combination of species occurrences, body size, and locomotor traits should correspond to some level of climatic and environmental differentiation. The results suggest that the frequency of locomotor types in a fauna may be a guide to palaeoenvironmental interpretation. Purely terrestrial locomotion dominates the northern faunas (38%-46% of the fauna), while subterranean-terrestrial species dominate the mid-latitude and southern faunas (35-51%). Subterranean species are found in faunas that have varied soils associated with high topographic relief and variable conditions, both diurnally and seasonally. This combination of conditions may favour subterranean and subterranean-terrestrial species that look for shelter and food storage underground (see SLOH species, Paper I). Mesoamerica, with its tropical and subtropical forests, has a high proportion of arboreal and arboreal-terrestrial species (12% and 22% respectively). Arboreality is associated with to dense tree cover, while arboreal-terrestrial species are associated with savanna and woodland environments. The results confirm previous studies that showed that standing diversity and body size distributions are related to climate and could thus be useful for palaeoclimatic reconstruction. Cold regions (mean annual temperatures < -5 °C) have a low number of mammals (between 49 to 58 species). The proportion of large (> 8000 g) species is over 24% of the community composition, and the proportion of small species (< 500 g) is less than 51%. Relatively wet and mild regions (700 - 1050 mm per year; 0 - 11 °C MAT) have a moderate number of species (≈ 80) with large species making up between 10 - 21% of the fauna and small species between 53 - 61%. Dry areas with low precipitation, moderate temperatures, and high elevations (300 - 500 mm per year; 6 - 20 °C; > 1400 m) have a high number of species (140 - 150) with few large species (9%) and many small ones (65 - 68%). Warm and humid areas (> 23 °C, > 1700 mm per year) have a very high number of species (>240) with fewer than 5% large species and more than 70% small species.

6. Conclusions

This thesis study examined the relationships between mammalian traits, community structure and their relationship with the environment and climate. Paper I investigated the relationship between body mass and the origination and extinction rates of taxa, presenting a new hypothesis that taxa with the ability to buffer themselves from environmental stress last longer. In Paper II, a method was developed to estimate precipitation from a set of ecomorphological characteristics of modern large herbivorous mammal communities. Paper III analysed the modern mammalian community structure to determine whether these structures similarly relate to the vegetation in tropical localities. Paper IV concentrated on the mammal distribution of North America and Europe to examine whether they construct similar types of climate-related communities.

In Paper I, the results showed that there are higher origination and extinction rates in larger mammals, i.e. the smaller taxa last longer. By analysing the results, a new hypothesis for this pattern was presented: it is not just being smaller per se that allows greater survivorship. Rather, a greater proportion of small mammals may be more apt at hiding (e.g., in burrows or tree holes) or going into low metabolic rate modes (e.g., hibernation and torpor). Taxa displaying such behaviour generally have a small body size (this study), and demonstrate the ability to buffer themselves from environmental stress, suggesting that SLOH (“sleep-or-hide”) genera are less extinction prone and experience less selective pressure to evolve to cope with environmental change. Whether species are seasonal or facultative in the hibernation–torpor continuum, this behaviour is associated with decreased food resources, lower ambient temperatures and lowered oxygen availability. Because this behaviour is in part plastic, species that display it are expected to cope better with (un)expected climate variation.

The method of Paper II was developed to estimate precipitation from a set of ecomorphological characteristics of large herbivorous mammal communities. These estimations were based on regression trees and used the characteristics of mammalian communities to define the relationships with the environment. These regression trees gave reasonably accurate estimates of precipitation values for today’s world, and thus can be applied to palaeocommunities to estimate past environmental conditions.

In Paper III, the tropical mammalian community structure was analysed to determine whether there are any dependencies in the communities and habitats. According to the results, there is wide applicability in distinguishing between modern habitats and, by extension, fossil localities where relatively complete species lists with species traits can be obtained. While defining palaeospecies traits is not a straightforward task, the findings of Paper III encourage palaeontological research to continue to gather information on individual taxa. This will enable both broad ecosystem discriminations and more precise habitat reconstructions, as well as the construction of palaeocommunity structure with habitat affiliations.

While the study in Paper III concentrated on locality-based data, the approach in Paper IV was to detect broader entities based on mammal distribution data. The results in Paper IV concur with earlier studies: clustered land mammal community compositions express both climatic and physical environments modified by biogeographical history. By using species traits (especially locomotion), the connection of fossil-based studies with living ecosystems may be used to reconstruct past environmental conditions.

To conclude, the results of Papers I–IV have demonstrated that many types of relationships exist between the community structure of mammals and their environment, ranging from continental-level climatic relationships to local vegetation models. The studies have also shown that more precise information on species can provide tools to interpret more nuanced information about past environments and how different climatic conditions have evolved during the past million years. The more we know about past environments, the better we can interpret how the future climatic conditions are going to change.

References

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46.

Anderson, M.J., 2005. PERMANOVA. Permutational Multivariate Analysis of Variance. Department of Statistics, University of Auckland, Auckland.

Anderson, M.J., Willis, T.J., 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84, 511–525.

Andrews, P., Lord, J.M. and Evans, E.M.N., 1979. Patterns of ecological diversity in fossil and modern mammalian faunas. Biological Journal of the Linnean Society, 11(2), pp.177-205

Andrews, P. and Humphrey, L., 1999. African Miocene environments and the transition to early hominines. African Biogeography, Climate Change, and Human Evolution. Oxford University Press, Oxford, pp.282-300.

Andrews, P. and Hixson, S., 2014, April. Taxon-free methods of palaeoecology. In Annales Zoologici Fennici (pp. 269-284). Finnish Zoological and Botanical Publishing.

Badgley, C., Damuth, J., Theodor, J., Werdelin, L., 2001. Global variation in ecological structure of mammalian faunas and paleoenvironmental inferences. J. Vertebr. Paleontol. 21 (suppl.), 30A.

Bailey, R. G. (1989). Explanatory supplement to ecoregions map of the continents. Environmental Conservation, 16(04), 307–309.

Bailey, R. G., and Hogg, H. C. (1986). A world ecoregions map for resource reporting. Environmental Conservation, 13(03), 195–202.

Brown, J. and Yalden, D.W., 1973. The description of mammals–2 Limbs and locomotion of terrestrial mammals. Mammal Review, 3(4), pp.107-134.

Burney, D.A., Burney, L.P., Godfrey, L.R., Jungers, W.L., Goodman, S.M., Wright, H.T. and Jull, A.T., 2004. A chronology for late prehistoric Madagascar. Journal of Human Evolution, 47(1-2), pp.25-63.

Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. Austral ecology, 18(1), pp.117-143.

Connolly, S.R. and Miller, A.I., 2001. Joint estimation of sampling and turnover rates from fossil databases: capture-mark-recapture methods revisited. Paleobiology, 27(4), pp.751-767.

Damuth, J.D. and MacFadden, B.J. eds., 1990. Body size in mammalian paleobiology: estimation and biological implications. Cambridge University Press.

Damuth, J., Fortelius, M., Andrews, P., Badgley, C., Hadly, E.A., Hixon, S., Janis, C., Madden, R.H., Reed, K., Smith, F.A., Theodor, J., Van Dam, J.A., Van Valkenburgh, B.,Werdelin, L., 2002. Reconstructing mean annual precipitation based on mammalian dental morphology and local species richness. J. Vertebr. Paleontol. 22 (suppl.), 48A.

De Vivo, M. and Carmignotto, A.P., 2004. Holocene vegetation change and the mammal faunas of South America and Africa. Journal of Biogeography, 31(6), pp.943-957.

Eisenberg, J.F., 1981. The mammalian radiations: an analysis of trends in evolution, adaptation, and behaviour (No. 591.5: 599 EIS).

Endler, J.A., 1982. Problems in distinguishing historical from ecological factors in biogeography. American Zoologist, 22(2), pp.441-452.

Fauquette, S., Guiot, J. and Suc, J.P., 1998. A method for climatic reconstruction of the Mediterranean Pliocene using pollen data. Palaeogeography, Palaeoclimatology, Palaeoecology, 144(1), pp.183-201.

Fortelius, M., coordinator. 2009. Neogene of the Old World Database of Fossil Mammals (NOW). Helsinki: University of Helsinki (http://www.helsinki.fi/science/now/).

Fortelius, M., 2013. The grassiness of all flesh. Journal of Biogeography, 40(7), pp.1213-1214.

Fortelius, M., Eronen, J., Jernvall, J., Liu, L., Pushkina, D., Rinne, J., Tesakov, A., Vislobokova, I., Zhang, Z. and Zhou, L., 2002. Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. Evolutionary Ecology Research, 4(7), pp.1005-1016.

Fortelius, M., Žliobaitė, I., Kaya, F., Bibi, F., Bobe, R., Leakey, L., Leakey, M., Patterson, D., Rannikko, J. and Werdelin, L., 2016. An ecometric analysis of the fossil mammal record of the Turkana Basin. Phil. Trans. R. Soc. B, 371(1698), p.20150232.

Hanya, G., Stevenson, P., van Noordwijk, M., Te Wong, S., Kanamori, T., Kuze, N., Aiba, S.I., Chapman, C.A. and van Schaik, C., 2011. Seasonality in fruit availability affects frugivorous primate biomass and species richness. Ecography, 34(6), pp.1009-1017.

Harrison, J.L., 1962. The distribution of feeding habits among animals in a tropical rain forest. The Journal of Animal Ecology, pp.53-63.

Heikinheimo, H., Fortelius, M., Eronen, J. & Mannila, H. (2007). Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. Journal of Biogeography, 34(6), 1053–1064.

Heikinheimo, H., Eronen, J.T., Sennikov, A., Preston, C.D., Oikarinen, E., Uotila, P., Mannila, H. & Fortelius, M. (2012). Convergence in the distribution patterns of Europe's plants and mammals is due to environmental forcing. Journal of Biogeography, 39(9), 1633–1644.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. International journal of climatology, 25(15), 1965–1978

Hutchinson, G.E., 1959. Homage to Santa Rosalia or why are there so many kinds of animals?. The American Naturalist, 93(870), pp.145-159.

IUCN, UNEP, 2009. The World Database on Protected Areas (WDPA). UNEP-WCMC, Cambridge, UK (Available from http://www.wdpa.org).

Janis, C.M., 1990. Correlation of cranial and dental variables with dietary preferences in mammals: a comparison of macropodoids and ungulates. Memoirs of the Queensland Museum, 28(1), pp.349-366.

Janis, C.M., 1993. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. Annual Review of Ecology and Systematics, 24(1), pp.467-500.

Janis, C.M. and Fortelius, M., 1988. On the means whereby mammals achieve increased functional durability of their dentitions, with special reference to limiting factors. Biological Reviews, 63(2), pp.197-230.

Jernvall, J. and Fortelius, M., 2002. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. Nature, 417(6888), pp.538-540.

Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., Safi, K., Sechrest, W., Boakes, E.H., Carbone, C. and Connolly, C., 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology, 90(9), pp.2648-2648.

Kaiser, T.M., Müller, D.W., Fortelius, M., Schulz, E., Codron, D. and Clauss, M., 2013. Hypsodonty and tooth facet development in relation to diet and habitat in herbivorous ungulates: implications for understanding tooth wear. Mammal Review, 43(1), pp.34-46.

Kappelman, J., 1988. Morphology and locomotor adaptations of the bovid femur in relation to habitat. Journal of Morphology, 198(1), pp.119-130.

Kay, R.F. and Madden, R.H., 1997. Mammals and rainfall: paleoecology of the middle Miocene at La Venta (Colombia, South America). Journal of Human Evolution, 32(2-3), pp.161-199.

Kleiber M. 1932: Body size and metabolism. Hilgardia, 6(11), pp. 315—353

Kleiber, M., 1947. Body size and metabolic rate. Physiological reviews, 27(4), pp.511-541.

Koch, P.L. and Barnosky, A.D., 2006. Late Quaternary extinctions: state of the debate. Annual Review of Ecology, Evolution, and Systematics, 37.

Kovarovic, K., Andrews, P. and Aiello, L., 2002. The palaeoecology of the Upper Ndolanya Beds at Laetoli, Tanzania. Journal of Human Evolution, 43(3), pp.395-418.

Kovarovic, K., Su, D.F. & Lintulaakso, K. (Accepted). Mammal community structure analysis. In Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities. Croft, D.A., Simpson, S.W. & Su, D.F. Springer.

Krzanowski, W. J. (1988) Principles of multivariate analysis. Oxford University Press, New York.

Lawing, A.M., Eronen, J.T., Blois, J.L., Graham, C.H. and Polly, P.D., 2017. Community functional trait composition at the continental scale: the effects of non-ecological processes. Ecography, 40(5), pp.651-663.

Linder, H.P., de Klerk, H.M., Born, J., Burgess, N.D., Fjeldså, J. & Rahbek, C. (2012). The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. Journal of Biogeography, 39(7), 1189–1205.

Lindstedt, S.L., Miller, B.J. and Buskirk, S.W., 1986. Home range, time, and body size in mammals. Ecology, 67(2), pp.413-418.

Lintulaakso, K., 2013. MammalBase — Database of recent mammals. http://www.mammalbase.net.

Lintulaakso, K. and Kovarovic, K., 2016. Diet and locomotion, but not body size, differentiate mammal communities in worldwide tropical ecosystems. Palaeogeography, Palaeoclimatology, Palaeoecology, 454, pp.20-29.

Liu, L., Puolamäki, K., Eronen, J.T., Ataabadi, M.M., Hernesniemi, E. and Fortelius, M., 2012. Dental functional traits of mammals resolve productivity in terrestrial ecosystems past and present. Proceedings of the Royal Society of London B: Biological Sciences, p.rspb20120211.

Louys, J., Meloro, C., Elton, S., Ditchfield, P. and Bishop, L.C., 2011. Mammal community structure correlates with arboreal heterogeneity in faunally and geographically diverse habitats: implications for community convergence. Global Ecology and Biogeography, 20(5), pp.717-729.

MacLeod, N. and Rose, K.D., 1993. Inferring locomotor behavior in Paleogene mammals via eigenshape analysis. American Journal of Science, 293(A), p.300.

Miljutin, A., 2009. Substrate utilization and feeding strategies of mammals: description and classification. Estonian Journal of Ecology, 58(1).

Mitchell-Jones, A.J., Mitchell, J., Amori, G., Bogdanowicz, W., Spitzenberger, F., Krystufek, B., Vohralík, V., Thissen, J., Reijnders, P., Ziman, J.M. & Stubbe, C.M. (1999). The atlas of European mammals (Vol. 3). London: Academic Press. (http://www.european-mammals.org/)

Mendoza, M., Goodwin, B. and Criado, C., 2004. Emergence of community structure in terrestrial mammal-dominated ecosystems. Journal of theoretical biology, 230(2), pp.203-214.

Mosbrugger, V. and Utescher, T., 1997. The coexistence approach—a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. Palaeogeography, Palaeoclimatology, Palaeoecology, 134(1-4), pp.61-86.

Olson, J.S.,Watts, J.A., Allison, L.J., 1983. Carbon in Live Vegetation of Major World Ecosystems.

Oak Ridge National Laboratory, Oak Ridge TN.

Olson, J.S., Watts, J.A., Allison, L.J., 1985. Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation (NDP-017).

Owen-Smith, N., 2013. Contrasts in the large herbivore faunas of the southern continents in the late Pleistocene and the ecological implications for human origins. Journal of Biogeography, 40(7), pp.1215-1224.

Patterson, B., Ceballos, G., Sechrest, W., Tognelli, M., Brooks, T., Luna, L., Ortega, P., Salazar, I. & Young, B. (2005). Digital distribution maps of the mammals of the western hemisphere, ver. 2.0 Arlington. Virginia in The NatureServe on line: http://www.natureserve.org.

Pineda-Munoz, S. and Alroy, J., 2014. Dietary characterization of terrestrial mammals. Proceedings of the Royal Society of London B: Biological Sciences, 281(1789), p.20141173.

Polly, P.D. (2010). Tiptoeing through the trophics: geographic variation in carnivoran locomotor ecomorphology in relation to environment. Carnivoran evolution: new views on phylogeny, form, and function, (ed. by A. Goswami, and A. Friscia), pp. 374–401. Cambridge University Press, Cambridge

Reed, K.E., 1997. Early hominid evolution and ecological change through the African Plio-Pleistocene. Journal of human evolution, 32(2-3), pp.289-322.

Reed, K.E., 1998. Using large mammal communities to examine ecological and taxonomic structure and predict vegetation in extant and extinct assemblages. Paleobiology, 24(3), pp.384-408.

Reed, K.E., 2008. Paleoecological patterns at the Hadar hominin site, Afar regional state, Ethiopia. Journal of Human Evolution, 54(6), pp.743-768.

Rencher, A.C. (2002) Methods of multivariate analysis, 2nd edn. John Wiley & Sons, New York.

Rodríguez, M.Á., López-Sañudo, I.L. and Hawkins, B.A., 2006. The geographic distribution of mammal body size in Europe. Global Ecology and Biogeography, 15(2), pp.173-181.

Rodríguez, M.Á., Olalla-Tárraga, M.Á. and Hawkins, B.A., 2008. Bergmann's rule and the geography of mammal body size in the Western Hemisphere. Global Ecology and Biogeography, 17(2), pp.274-283.

Samuels, J.X. and Van Valkenburgh, B., 2008. Skeletal indicators of locomotor adaptations in living and extinct rodents. Journal of morphology, 269(11), pp.1387-1411.

Schmidt-Nielsen, K., 1984. Scaling: why is animal size so important? Cambridge University Press.

van Dam, J.A., Andrews, P., Badgley, C., Damuth, J., Fortelius, M., Hadly, E.A., Hixson, S., Janis, C., Madden, R.H., Reed, K., Smith, F.A., Theodor, J., Van Valkenburgh, B., Werdelin, L., 2001. Within-habitat mammal diversity and productivity and their recent patterns across latitude. J. Vertebr. Paleontol. 21 (suppl), 43A.

van Valen, L., 1960. A functional index of hypsodonty. Evolution, 14(4), pp.531-532.

van Valkenburgh, B., 1987. Skeletal indicators of locomotor behavior in living and extinct carnivores. Journal of Vertebrate Paleontology, 7(2), pp.162-182.

Wilson, D.E. and Reeder, D.M. eds., 2005. Mammal species of the world: a taxonomic and geographic reference. JHU Press.

Wolfe, J.A., 1979. Temperature parameters of humid to mesic forests of Eastern Asia and relation to forests of other regions of the Northern Hemisphere and Australasia: analysis of temperature data from more than 400 stations in Eastern Asia. United States. Geological Survey. Professional paper (USA).

Wolfe, J.A., 1993. A method of obtaining climatic parameters from leaf assemblages (No. 2040-2041). US Government Printing Office.

World Wildlife Fund, 2006. WildFinder: Online Database of Species Distributions, January 2006. http://www.worldwildlife.org/wildfinder

Žliobaitė, I., Rinne, J., Tóth, A., Mechenich, M., Liping, L., Behrensmeyer, A. and Fortelius, M., 2016. Herbivore teeth predict climatic limits in Kenyan ecosystems. Proceedings of the National Academy of Science of the United States of America, 113, pp.12751-12756.

Paper I

Liow, L.H., Fortelius, M., Bingham, E., Lintulaakso,
K., Mannila, H., Flynn, L., Stenseth, N.C.
Higher origination and extinction rates in larger mammals.
Proceedings of the National Academy of Sciences, 105(16), 2008.

Higher origination and extinction rates in larger mammals

Lee Hsiang Liow*, Mikael Fortelius[†‡], Ella Bingham[§], Kari Lintulaakso[†], Heikki Mannila[§¶], Larry Flynn[‖], and Nils Chr. Stenseth*,**

*Center for Ecological & Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, P.O. Box 1066 Blindern, N-0316 Oslo, Norway; [†]Department of Geology, University of Helsinki, P.O. Box 64, FIN-00014 Helsinki, Finland; [‡]Institute of Biotechnology, University of Helsinki, P.O. Box 56, FIN-00014 Helsinki, Finland; [§]Helsinki Institute for Information Technology, Department of Computer Science, University of Helsinki, P.O. Box 68, FIN-00014 Helsinki, Finland; [¶]Helsinki Institute for Information Technology, Department of Computer and Information Science, Helsinki University of Technology, P.O. Box 5400, FIN-02015 Helsinki, Finland; and [‖]Peabody Museum, Harvard University, Cambridge, MA 02138

Edited by David Pilbeam, Harvard University, Cambridge, MA, and approved February 5, 2008 (received for review October 14, 2007)

Do large mammals evolve faster than small mammals or vice versa? Because the answer to this question contributes to our understanding of how life-history affects long-term and large-scale evolutionary patterns, and how microevolutionary rates scale-up to macroevolutionary rates, it has received much attention. A satisfactory or consistent answer to this question is lacking, however. Here, we take a fresh look at this problem using a large fossil dataset of mammals from the Neogene of the Old World (NOW). Controlling for sampling biases, calculating per capita origination and extinction rates of boundary-crossers and estimating survival probabilities using capture-mark-recapture (CMR) methods, we found the recurring pattern that large mammal genera and species have higher origination and extinction rates, and therefore shorter durations. This pattern is surprising in the light of molecular studies, which show that smaller animals, with their shorter generation times and higher metabolic rates, have greater absolute rates of evolution. However, higher molecular rates do not necessarily translate to higher taxon rates because both the biotic and physical environments interact with phenotypic variation, in part fueled by mutations, to affect origination and extinction rates. To explain the observed pattern, we propose that the ability to evolve and maintain behavior such as hibernation, torpor and burrowing, collectively termed ''sleep-or-hide'' (SLOH) behavior, serves as a means of environmental buffering during expected and unexpected environmental change. SLOH behavior is more common in some small mammals, and, as a result, SLOH small mammals contribute to higher average survivorship and lower origination probabilities among small mammals.

body size | environmental buffering | metabolism | Neogene mammals | turnover

Evolution operates at different scales of time and levels of the biological hierarchy (1). Body size covaries with many individual and species level traits (2), each of which could influence the tempo of evolution at population, species and clade levels. Multiple studies have shown that smaller sized mammals have higher molecular rates of evolution in absolute time, possibly because of a generation time effect and/or metabolic rate effect (3–8). Higher molecular rates may translate to higher rates of phenotypic changes (9 but see 10, 11) and a greater chance of reproductive isolation, which could ultimately lead to higher speciation rates (12) and higher rates of pseudoextinction, which could be observed as higher extinction rates among fossil taxa. As such, small mammals could be expected to have higher origination and extinction rates as observed in the fossil record. However, empirical studies on historical extinctions show that large mammals are at higher risk of extinction and have been selectively removed (13–17), as is also indicated by the Pleistocene megafauna extinction literature (18, 19), even though body size *per se* may not always be a good predictor of extinction risk (20–22). Neither molecular studies nor research on historical extinctions and extinction in the shallow fossil record gives us direct insight into the distribution of taxon durations and long-term turnover rates. Previous studies show mixed results on mammal size-biased selectivity. Three possible scenarios have been described, namely, as no size bias (23, 24), greater survivorship in large mammals versus small (25, 26), and greater survivorship in small mammals versus large (27, 28). However, little is known about mammal body size biases in origination rates in the fossil record. The uncertainty in both the simple existence of a body size bias in origination and extinction rates, as well as the direction in which the bias should manifest itself, is considerable. On average, compared within groups, species with larger body sizes often have wider geographic ranges (24, 29), need greater patch sizes (30), have smaller population sizes/lower densities/abundances (31), have longer generation times (32), have lower fecundity (33), have lower specific production rates (34), and have longer individual life spans and lower metabolic rates (35). These traits potentially push survivorship and the propensity to generate new species and genera in opposing directions: larger mammals have wider geographic ranges and potentially greater dispersal abilities that could lower their extinction risk (36). Conversely, the longer generation times and smaller population sizes of larger mammals might increase extinction risk (37). Similarly, the smaller effective population size of larger mammals may confer higher speciation and origination rates, but, simultaneously, their ability to disperse and encourage gene flow may dampen these rates.

Here, we use a large Neogene Old Word (NOW) fossil mammal dataset (38) to study whether there are body size biases in taxon durations, survivorship, and originations. Common problems encountered in studying fossil lineages are the heterogeneous quality of localities and the uncertain endpoints of taxon durations. In addition, although taphonomic differences among mammals of different body sizes have long been recognized (39, 40), the bias introduced into preservation rates has not been incorporated in analyses in a comprehensive fashion. We embrace these issues and clearly demonstrate, using various analytical approaches, that large mammals have both higher extinction and origination rates and therefore a duration distribution that has a more truncated right tail.

Results and Discussion

Among mammal genera with shorter durations, there is a predominance of large mammals (Artiodactyla, Carnivora, Creodonta,

---

Author contributions: L.H.L., M.F., H.M., and N.C.S. designed research; L.H.L., M.F., E.B., K.L., and L.F. performed research; L.H.L., E.B., K.L., H.M., and L.F. analyzed data; and L.H.L., M.F., L.F., and N.C.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

**To whom correspondence should be addressed. E-mail: n.c.stenseth@bio.uio.no.

This article contains supporting information online at www.pnas.org/cgi/content/full/0709763105/DCSupplemental.

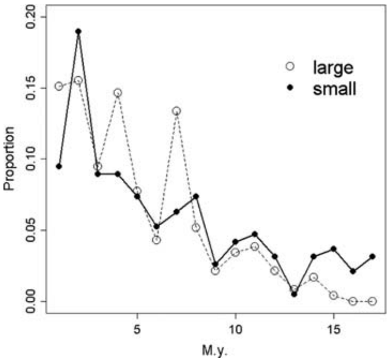


Fig. 1. Histogram of durations of extinct small and large mammal genera. This histogram represents the full dataset (All); for subsets, see Fig. S1. Proportions are calculated for small (●) and large (○) mammals separately. One-million-year bins are plotted.

Hyracoidea, Perissodactyla, Primates, Proboscidea, and Tubulidentata) whereas, among those with longer durations, there is an excess of small mammals (Insectivora, Rodentia, and Lagomorpha) (Fig. 1 and supporting information (SI) Fig. S1). Mean and median durations of large and small mammal genera reflect the same tendency: small genera have greater mean and median durations than large ones (Kolmogorov–Smirnov test, $P < 0.05$, Mann–Whitney U test < 0.05; Table S1 in Dataset S1). At species level, the same difference is present in duration means, although it is not statistically significant (Table S1 in Dataset S1). This difference in durations could be due to only a lower rate of extinction, but we find that large mammal genera have both higher per capita rates of origination and extinction than small mammals (Table 1 and Table S2 in Dataset S1; Mann–Whitney U test < 0.05). The differences in extinction rates are greater than in origination rates, in general (Table S2 in Dataset S1). Using a capture-mark-recapture (CMR) approach, we conclude similarly that large animals have comparatively lower survivorship (Table S3 in Dataset S1). The best models in each case from CMR analyses always include a body size effect, and the weights for these top models are high (Table S3 in Dataset S1), regardless of the sub-setting or temporal binning of the data. The estimates for body size effect, and their lower confidence limits, are positive (Table 2), indicating that large mammals survive with a lower probability. Species results are statistically weaker or insignificant, because species occurrences are sparse, but nevertheless point in the same direction (CMR species results not presented).

Table 1. Per capita rates of origination and extinction

Rates	Data subsets	Large	Small	M–W	*N*
p	All	0.40	0.25	0.008	18
	5_occ	0.23	0.13	0.003	17
	5_taxa	0.35	0.27	0.094	18
	10_occ_10_taxa	0.17	0.16	0.078	17
q	All	0.30	0.15	0.103	18
	5_occ	0.35	0.15	0.059	17
	5_taxa	0.23	0.13	0.103	18
	10_occ_10_taxa	0.34	0.13	0.045	17

Mean per capita rates of origination, p, and extinction, q, per million years for large and small mammal genera. All, the entire dataset; 5_occ, data subset where only genera with at least five occurrences were included; 5_taxa, subset where only localities with at least five genera were included 10_occ_10_taxa, subset where first all genera with fewer than 10 occurrences were excluded and then localities with fewer than 10 genera excluded. M–W, *P* values for paired, two-tailed Mann–Whitney *U* tests. *N*, number of 1-M.y. time bins for which comparisons were possible. See Table S2 in Dataset S1 for alternative binning schemes.

Large mammal genera from NOW (Artiodactyla, Carnivora, Creodonta, Hyracoidea, Perissodactyla, Primates, Proboscidea, and Tubulidentata), for which we have body mass estimates via a nearest living relative approach, truly have greater body mass (Fig. 2; Kolmogorov–Smirnov test $P < 0.0001$) compared with small mammals (Insectivora, Rodentia, and Lagomorpha). Unfortunately, because the duration distributions of these genera are based on truncated durations, we cannot distinguish their duration distribution from that of small mammals for which we have similar body mass estimates (Kolmogorov–Smirnov test $P = 0.43$). Body size is indeed a general predictor of mammal taxon durations (this study) even though there are certainly short-duration small and long-duration large mammal species and genera (see *SI Text* for a discussion).

How much of this detected difference is due to preservation differences? Small and large mammals have different modes of preservation and are often collected in the field and processed in the laboratory with different methods (25, 41). Specifically, large mammals are often preferentially preserved and recorded (39); therefore, they should have greater apparent durations. Despite the prediction of this preservation bias, we find that large mammals have on average shorter durations. However, taxonomic practices could have an influence in the opposite direction. For example, large mammal species might be preferentially described as new (e.g., more "splitters" among large mammal researchers), which could potentially bias large mammal durations toward being shorter. Unfortunately, no available data allow us to address such potential factors. We were, however, able to model body size as a covariate in preservation rate and found that its effect is inconsistent (Table 2). In cases where body size had an effect on preservation rate, the effect of body size on survivorship remained strong (Table 2). In other words, large mammals may have a somewhat higher preservation rate, but their survival probability is lower than that of small mammals even after accounting for differential preservation.

Why do our results differ from other mammalian body size turnover rate studies mentioned in the introduction? Muñoz-Durán (23) and Viranta (24) found no size bias in survivorship, but their studies involved only carnivores and these are all considered large mammals in our analyses; therefore, there is no direct contradiction (see *SI Text* for a discussion on carnivores). Van Valen (26) demonstrated that large mammal taxa have a longer half life than mammals in general in a semi-global genus dataset, but mammal taxonomy and chronology have improved over time. An independent lower latitude fossil mammal dataset from Miocene deposits in Pakistan, known as the Siwaliks (25, 41), also shows opposing results. The Siwaliks dataset has a broad taxonomic coverage, but a more limited temporal and geographic coverage. Also, differences in collecting strategies not taken into account could have played a part in opposite conclusions. A real biological signal may also be present: the Siwaliks assemblages represent a subtropical fauna and differ from NOW taxa, which were derived from a largely temperate fauna. The disadvantages of being a large mammal may not be felt under conditions of a more dependable (tropical or subtropical) environment. Consequently, the effects of a shorter generation time and species competition (42) as local small mammal diversity increases could lead to more rapid turnover of small mammals in the (sub)tropics (43).

Although shorter generation times and higher metabolic rates of small mammals may increase molecular rates of evolution (8, 11), and even if we were to assume that molecular changes translate substantially to phenotypic changes (12), these shorter-term, generational changes do not seem to translate to higher turnover rates as we have shown here (but see ref. 44). The study of evolutionary rates should therefore embrace different levels of the biological hierarchy. We emphasize that this discrepancy does not imply

Table 2. Body size effects on survivorship and preservation

Bin data subsets		1-M.y. bins				1.5-M.y. bins				2-M.y. bins			
	Estimate	Beta	SD	L	U	Beta	SD	L	U	Beta	SD	L	U
All	Φ	0.664	0.181	0.309	1.018	0.600	0.170	0.267	0.932	1.208	0.000	1.208	1.208
5_occ	Φ	0.783	0.191	0.409	1.157	0.678	0.161	0.363	0.993	1.931	0.000	1.931	1.931
	pr	−0.449	0.211	−0.864	−0.035	−0.434	0.244	−0.913	0.045	−0.542	0.260	−1.052	−0.031
5_taxa	Φ	0.824	0.233	0.368	1.281	0.669	0.218	0.243	1.096	0.817	0.304	0.220	1.413
	pr	−0.473	0.212	−0.889	−0.058	−0.544	0.296	−1.124	0.036	−0.685	0.405	−1.478	0.108
10_occ_10_taxa	Φ	0.969	0.280	0.420	1.518	0.617	0.252	0.123	1.111	NA	NA	NA	NA

Group effect for each of the best models from the four data subsets, using 1-, 1.5-, and 2-M.y. bins. Estimates presented are survivorship (Φ) and preservation (pr). Beta, estimates for body size effect; SD, standard deviations; L, lower of the 95% confidence interval; U, upper of the 95% confidence interval.

micro- and macroevolution discontinuity, but that using different observational scales and units reveals basal and emergent processes in turn.

Small mammals have on average lower turnover rates at both species and genus levels, even though noisier species data did not allow us to infer species level results with statistical confidence. We take this result one step further and present a hypothesis that to our knowledge has never been analyzed: it is not just being smaller *per se* that allows greater survivorship (many small mammals also have short species and genus durations) (Fig. 1 and Fig. S1). Rather, a greater proportion of small mammals may be more apt at hiding (e.g., in burrows or tree holes) or going into low metabolic rate modes (e.g., hibernation and torpor). We call such behavior "sleep-or-hide" (SLOH). Taxa that demonstrate the ability to buffer themselves from environmental stress in this manner are here termed SLOH taxa.

We used living species to infer the presence/absence of SLOH behavior in fossil genera. Of the 50 large NOW mammal genera for which we have the nearest living relative estimates, 15 exhibit SLOH behavior whereas, for 67 small NOW mammals genera (i.e., genera from the orders Insectivora, Lagomorpha, and Rodentia), 41 exhibit SLOH behavior (see Table S4 in Dataset S1). This bias is significant (χ^2 test $P = 0.0001$).

More genera inferred to have SLOH behavior in our dataset have smaller body mass estimates (Fig. 3; Kolmogorov–Smirnov test $P < 0.001$). There is also a group of small mammals that does not display SLOH behavior (Fig. 3). Therefore, some, but not all, small mammals may be better buffered against environmental changes than larger mammals.

As discussed in more detail in *SI Text*, there is good evidence that the longest-lived small mammal genera have SLOH traits, as inferred from their nearest living relatives. Thus, for the most inclusive dataset (All) we find in the longest living 16-million year (M.y.) class a mole, two gliding squirrels, and two dormice. In the 15-M.y class, there is a shrew, a hedgehog, a ground squirrel, a hamster, and a burrowing field mouse. The 14-M.y class comprises two shrews, a mole, a gliding arboreal squirrel, and a burrowing mole rat. Apart from two eomyid rodents with unknown lifestyles, all genera in these highest range classes are likely to exhibit one or more SLOH traits. The sole large mammal in these duration classes is the tapir, genus *Tapirus* (extinct in the study area and therefore included in the analysis).

Body masses of large mammals increased over the Neogene (Fig. S2; see also ref. 45), and this temporal trend corresponds to a decrease in the ability to sleep and/or hide (Fig. S2). In contrast, there are no such trends over the Neogene for small mammals (Fig. S2). Incidentally, this body size constraint on SLOH behavior could explain an old paradox: the European Miocene–Pliocene boundary is well defined by a turnover event in the large mammal fauna but not that of small mammals (46).

We interpret the patterns in our data to indicate that the ability of a species or genus to shield itself from environmental fluctuations lowers turnover rates, and that the ability to "sleep-or-hide" is constrained by body size (47). We suggest that SLOH genera are less extinction prone and experience less selective pressure to evolve to cope with environmental change. Whether species are seasonal or facultative in the hibernation–torpor continuum (48), this behavior is associated with decreased food resources, lower ambient temperatures, and lowered oxygen availability (49). Because this behavior is in part plastic (48, 50–52), species that display it are expected to cope better with (un)expected climate variation. It is typical for mammals to take shelter in burrows, holes, or caves,

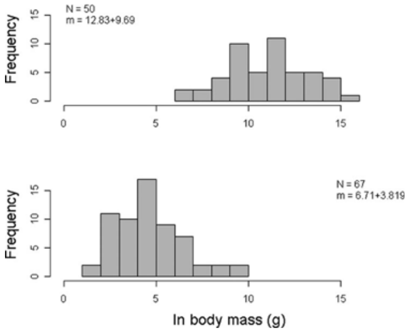


Fig. 2. Histograms of ln body mass of NOW mammal genera designated large (*Upper*) and small (*Lower*). Body masses are inferred by using a nearest relative approach. *N*, sample size;*m*, mean natural log (ln) body mass (in grams).

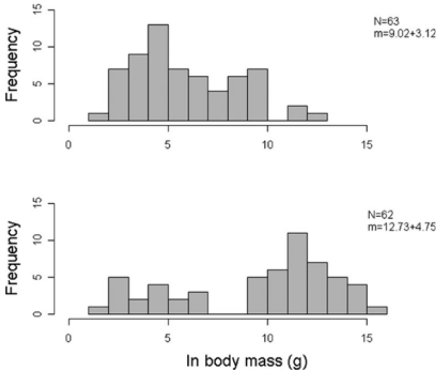


Fig. 3. Histograms of ln body mass of SLOH (*Upper*) and non-SLOH (*Lower*) genera. *N*, sample size; *m*, mean ln body mass (in grams). See Fig. S3 for an alternative plot of this figure.

during hibernation or torpor, in part to maintain a more constant ambient environment. In fact, burrowing could have evolved as an adaptation to deteriorating conditions (53, 54), in combination with the exploitation of underground resources. In addition to providing a more constant ambient environment, such hiding places can also double as refuges from predators (17, 55, 56), including human hunters (21).

Climate and environment both have a large part to play in the evolutionary theater, even though the play must also depend on life-history and biotic interactions. Which taxa survive may be strongly influenced by the number and scale of climate reversals (57). However, in a similar environmental backdrop (Neogene of Europe), large mammals survive at lower rates than small mammals. Large mammals suffer from compound disadvantages. They have more deleterious mutations (37), smaller population sizes (31), and longer generation times (32) and are, in addition, probably constrained by their size from being apt at hiding or shutting down or lowering their metabolic rates temporarily. They may also require more energy to survive during periods of environmental stress (58, 59). However, we show that some small mammals are at an advantage because of "intrinsic traits," in contrast to claims that small mammal extinctions are driven only by environmental factors (16). But it is not just bad news for large mammals: they enjoy higher origination rates, and therefore their overall diversity is maintained. Even though community energy use is independent of body size (60, 61), large and small mammals, with their different specific production rates and life styles (34), respond differently to the same general environment. Paleobiological analyses (this study), combined with analyses clarifying how life history affects evolutionary and ecological patterns (34) inform us with greater confidence which kinds of species are more vulnerable to climatic fluctuations and trends, not the least of which are seen in the current climate crisis.

Materials and Methods

Data. The data we used are species occurrences of Neogene Old World land mammals from the NOW database (38) accessed on June 26, 2007. These data are based on published accounts of species occurrences at named sites (localities), vetted for taxonomic inconsistencies. We exclude localities east of 60°E, localities in Afro-Arabian countries, for the purpose of including only European Neogene localities for which we have greater confidence in both age estimates of localities and taxonomic identifications of samples. The occurrence data we used include specimens unidentified at the species level (i.e., with the suffixes sp., cf., and indet.), but any discussion of species results reflects only species that are fully identified. Our resulting dataset covers the longitudes 9.13°W to 58.9°E and latitudes 30.86°N to 55.0°N. Volant and marine mammals were excluded, but all trophic levels are represented.

Locality ages in the NOW database are recorded as minimum and maximum age estimates based on various combinations of direct dates (e.g., radiometric) and time units (62–66). In our analyses, locality ages are the means of minimum and maximum age estimates. However, we remove all localities with maximum minus minimum ages greater than the duration of the MN-unit into which the mean age falls, according to the chronology of Steininger *et al*. (62).

We have used both the full dataset as described in the previous paragraphs (All) and various subsets of it in our analyses. In the subset 5_occ, we include only genera that appear at least five times in the data, in which case we lose some localities that are potentially taxon rich. Conversely, in 5_taxa, we exclude localities that have fewer than five genera. Therefore, some genera could have been removed from analyses simply by association with those poorer localities. Similarly, 10_occ_10_taxa is the data subset where, first, only genera that appear at least 10 times in the data are retained, and then those sites with at least 10 genera. We ran analogous analyses for species but do not present all results for economy and clarity.

Durations. For the calculation and comparisons of untruncated taxon durations (in Fig. 1, Fig. S1, and Table 1), we used the following approach to remove potentially extant taxa in our fossil data. We first excluded taxa that appear at localities that are 0.5 M.y. old and younger and then further excluded taxa known to be living today (see *SLOH*) within the geographic bounds of the NOW data we are using. This is a highly conservative approach to removing extant or very recently extinct taxa. However, less conservative approaches (e.g., doing only the second sweep of exclusion) do not change results. Note that the analyses described in the next sections do not require any removal of extant taxa because one-side truncation of stratigraphic ranges is not a problem in rate calculations, as it is in duration calculations.

Per Capita Rates. For each data subset, we tabulated whether a genus was recorded as present or absent during time bins that span 1, 1.5, or 2 M.y. These alternative binning schemes were done to check for the robustness of our results with respect to the variability in the uncertainty of locality ages. If a genus was absent during one or more time bins, but was present in some time bin before and after those absences, it was assumed to have survived those time intervals. We calculated per capita extinction (q) and origination (p) rates (67) for large and small mammals separately in each dataset, where

$$p = \ln(N_{\mathrm{t}}/N_{\mathrm{bt}})/\Delta t,$$

$$q = \ln(N_{\mathrm{b}}/N_{\mathrm{bt}})/\Delta t$$

and N_{bt} is the number of taxa that cross both the earlier (bottom) and the later (top) boundary of a time bin, N_{t} is the number of genera that cross only the later of the time bins, N_{b} the number that cross only the earlier boundary of the time bins, and Δt is the width of the time bins, following ref. 67. We use per capita rates, based on boundary crossers, because they are relatively insensitive to secular trends in the quality of preservation (67). Even though the observed first and last occurrences of taxa may be temporally quite far removed from the true time of origination and extinction, even in densely sampled datasets (68), there is a strong positive correlation between the ranks of observed ranges and estimates of true durations (see ref. 69).

CMR Estimates. Capture-mark-recapture (CMR) methods have many powerful applications in ecology (70, 71) and are gradually finding their way into the paleobiological literature (70–72). Our purpose of employing CMR methods is to estimate both survivorship and preservation probabilities with the aim of investigating the effect of body size on both probabilities. By using a CMR approach, we also control for differences in preservation among large and small mammals. We compared time-varying estimates of survival (Φ) and preservation (pr) probabilities with either no body size effect, or additive or multiplicative effects of body size using a model selection approach (73) (see *SI Text* for more details).

Body Size Groups. Body size estimates for individual genera are not available for a substantial portion of fossil taxa in the NOW database. Therefore, we assigned to each genus one of two size classes, based on their taxonomic identity. All genera belonging to the orders Insectivora, Rodentia, and Lagomorpha are designated "small." All remaining genera are regarded as "large," namely Artiodactyla, Carnivora, Creodonta, Hyracoidea, Perissodactyla, Primates and Proboscidea, and Tubulidentata (74). In the absence of body size estimates for a large proportion of species in NOW, we justify our use of this coarse classification as follows. Bimodality in the body size distribution of mammals (26, 75, 76) has existed and intensified since the Eocene (45, 77). Moreover, body sizes are constrained by phylogeny with mammalian taxa having characteristic sizes over both space and time (78, 79). In addition, we used a nearest living relative approach to estimate mean body masses for the genera in the NOW database where possible (see next section for details).

SLOH. We used MammalBase, a database of living mammal attributes, based on refs. 80–83, compiled by one of us (K.L.), to extract SLOH information on extant species. SLOH attributes for genera used in this article are attached in Table S4 in Dataset S1 but further information is available on request. We coded all known extant species for SLOH behavior according to whether they (*i*) are hibernators, (*ii*) go into torpor, (*iii*) can be dormant, (*iv*) can go into aestivation mode, (*v*) make or use burrows or are fossorial, (*vi*) make or use tunnels or chambers, (*vii*) live in tree holes, or (*viii*) are cave dwellers. If one of these eight nonmutually exclusive traits is present for a given species, we coded it as 1 for its SLOH value and 0 if none was observed (Table S4 in Dataset S1). However, for data analyses, we treated the absence of mention of these traits as zeros instead of "NA" as coded, because of the common practice of not noting absent traits. If a NOW genus has any living species for which any one of the SLOH attributes is tabulated as present, we assign a SLOH value of "1" to that NOW genus and therefore to all of the NOW species that are members of the genus (nearest living relative approach). Additionally, one of us (L.F.) manually checked the entries described above and presented in Table S4 in Dataset S1, to ensure the quality of the data. L.F. provided alternative coding, and we show results based on amended assignments in Table S4 in

Dataset S1. These amendments, however, do not change our qualitative results or general conclusions.

Body Mass Estimates. With reference to the previous sections, for NOW genera with living species for which body masses are known from MammalBase, we similarly calculated the mean body mass from these living species and used these as mean body mass estimates for these genera. To supplement body mass data from MammalBase, we used body mass data compiled by two National Center for Ecological Analysis and Synthesis (NCEAS) working groups (ref. 84 and J. Damuth, personal communication).

ACKNOWLEDGMENTS. We thank Torbjørn Ergon, Jon Olav Vik, and Jim Nichols for technical advice and John Damuth for sharing unpublished data. Two anonymous reviewers offered valuable suggestions that improved the paper. This research is in part funded by an Academy of Finland grant (to M.F.) and core funding to the Center for Ecological & Evolutionary Synthesis.

1. Jablonski D (2007) Scale and hierarchy in macroevolution. *Palaeontology* 50:87–109.
2. Peters RH (1983) *The Ecological Implications of Body Size* (Cambridge Univ Press, Cambridge, UK).
3. Li W-H , Gouy M, Sharp PM, O'hUigin C, Yang Y-W (1990) Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. *Proc Natl Acad Sci USA* 87:6703–6707.
4. Martin AP, Palumbi SR (1993) Body size, metabolic-rate, generation time, and the molecular clock. *Proc Natl Acad Sci USA* 90:4087–4091.
5. Bromham L, Rambaut A, Harvey PH (1996) Determinants of rate variation in mammalian DNA sequence evolution. *J Mol Evol* 43:610–621.
6. Li WH, Ellsworth DL, Krushkal J, Chang BHJ, Hewett-Emmett D (1996) Rates of nucleotide substitution in primates and rodents and the generation time effect hypothesis. *Mol Phylogenet Evol* 5:182–187.
7. Rowe DL, Honeycutt RL (2002) Phylogenetic relationships, ecological correlates, and molecular evolution within the Cavioidea (Mammalia, Rodentia). *Mol Biol Evol* 19:263–277.
8. Gillooly JF, Allen AP, West GB, Brown JH (2005) The rate of DNA evolution: Effects of body size and temperature on the molecular clock. *Proc Natl Acad Sci USA* 102:140–145.
9. Omland KE (1997) Correlated rates of molecular and morphological evolution. *Evolution* 51:1381–1393.
10. Davies TJ, Savolainen V, Chase MW, Moat J, Barraclough TG (2004) Environmental energy and evolutionary rates in flowering plants. *Proc Biol Sci* 271:2195–2200.
11. Bromham L, Woolfit M, Lee MSY, Rambaut A (2002) Testing the relationship between morphological and molecular rates of change along phylogenies. *Evolution* 56:1921–1930.
12. Bromham L (2003) Molecular clocks and explosive radiations. *J Mol Evol* 57(Suppl 1):S13–S20.
13. McKinney ML (1997) Extinction vulnerability and selectivity: Combining ecological and paleontological views. *Annu Rev Ecol Syst* 28:495–516.
14. Harcourt AH, Schwartz MW(2001) Primate evolution: A biology of Holocene extinction and survival on the southeast Asian Sunda Shelf islands. *Am J Phys Anthropol* 114:4–17.
15. Cardillo M (2003) Biological determinants of extinction risk: Why are smaller species less vulnerable? *Anim Conserv* 6:63–69.
16. Cardillo M, *et al*. (2005) Multiple causes of high extinction risk in large mammal species. *Science* 309:1239–1241.
17. McKenzie NL, *et al*. (2007) Analysis of factors implicated in the recent decline of Australia's mammal fauna. *J Biogeogr* 34:597–611.
18. Koch PL, Barnosky AD (2006) Late quaternary extinctions: State of the debate. *Annu Rev Ecol Evol Syst* 37:215–250.
19. Lyons SK, Smith FA, Brown JH (2004) Of mice, mastodons and men: Human-mediated extinctions on four continents. *Evol Ecol Res* 6:339–358.
20. Purvis A, Gittleman JL, Cowlishaw G, Mace GM (2000) Predicting extinction risk in declining species. *Proc R Soc London Ser B Biol Sci* 267:1947–1952.
21. Johnson CN (2002) Determinants of loss of mammal species during the Late Quaternary ''megafauna'' extinctions: Life history and ecology, but not body size. *Proc R Soc London Ser B Biol Sci* 269:2221–2227.
22. Brashares JS (2003) Ecological, behavioral, and life-history correlates of mammal extinctions in West Africa. *Conserv Biol* 17:733–743.
23. Muñoz-Durán J (2002) Correlates of speciation and extinction rates in the Carnivora. *Evol Ecol Res* 4:963–991.
24. Viranta S (2003) Geographic and temporal ranges of Middle and Late Miocene carnivores. *J Mammal* 84:1267–1278.
25. Flynn LJ, *et al*. (1995) Neogene Siwalik mammalian lineages: Species longevities, rates of change, and modes of speciation. *Palaeogeogr Palaeoclimatol Palaeoecol* 115:249–264.
26. Van Valen L (1975) Group selection, sex and fossils. *Evolution* 29:87–94.
27. Van Valkenburgh B, Wang XM, Damuth J (2004) Cope's rule, hypercarnivory, and extinction in North American canids. *Science* 306:101–104.
28. McLain DK (1993) Cope's rules, sexual selection, and the loss of ecological plasticity. *Oikos* 68:490–500.
29. Brown JH, Stevens GC, Kaufman DM (1996) The geographic range: Size, shape, boundaries, and internal structure. *Annu Rev Ecol Syst* 27:597–623.
30. Biedermann R (2003) Body size and area-incidence relationships: Is there a general pattern? *Global Ecol Biogeogr* 12:381–387.
31. Blackburn TM, Gaston KJ (1999) The relationship between animal abundance and body size: A review of the mechanisms. *Adv Ecol Res* 28:181–210.
32. Brook BW, Bowman D (2005) One equation fits overkill: Why allometry underpins both prehistoric and modern body size-biased extinctions. *Popul Ecol* 47:137.
33. Brown JH, Sibly RM (2006) Life-history evolution under a production constraint. *Proc Natl Acad Sci USA* 103:17595–17599.
34. Sibly RM, Brown JH (2007) Effects of body size and lifestyle on evolution of mammal life histories. *Proc Natl Acad Sci USA* 104:17707–17712.
35. Symonds MRE (1999) Life histories of the Insectivora: The role of phylogeny, metabolism and sex differences. *J Zool* 249:315–337.
36. Stanley SM (1979) *Macroevolution: Pattern and Process* (John Hopkins Univ Press, Baltimore).
37. Popadin K, Polishchuk LV, Mamirova L, Knorre D, Gunbin K (2007) Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. *Proc Natl Acad Sci USA* 104:13390–13395.
38. Fortelius M (2008) *Neogene of the Old World Database of Fossil Mammals (NOW)* (University of Helsinki, Helsinki).
39. Muñoz-Durán J, Van Valkenburgh B (2006) The Rancholabrean record of carnivora: Taphonomic effect of body size, habitat breadth, and the preservation potential of caves. *Palaios* 21:424–430.
40. Badgley C (1986) Taphonomy of mammalian fossil remains from Siwalik rocks of Pakistan. *Paleobiology* 12:119–142.
41. Barry J, *et al*. (2002) Faunal and environmental change in the late Miocene Siwaliks of northern Pakistan. *Paleobiology* 28(Memoir 3, Suppl to Number 2):1–71.
42. Stenseth NC, Maynard Smith J (1984) Coevolution in ecosystems: Red Queen evolution or stasis. *Evolution* 38:870–880.
43. Stenseth NC (1984) The tropics: Cradle or museum. *Oikos* 43:417–420.
44. Allen AP, Gillooly JF, Savage VM, Brown JH (2006) Kinetic effects of temperature on rates of genetic divergence and speciation. *Proc Natl Acad Sci USA* 103:9130–9135.
45. Alroy J (1998) Cope's rule and the dynamics of body mass evolution in North American mammals. *Science* 280:731–734.
46. Agustí J, *et al*. (2001) A calibrated mammal scale for the Neogene of Western Europe: State of the art. *Earth Sci Rev* 52:247–260.
47. Geiser F (1998) Evolution of daily torpor and hibernation in birds and mammals: Importance of body size. *Clin Exp Pharmacol Physiol* 25:736–740.
48. Lehmer EM, Savage LT, Antolin MF, Biggins DE (2006) Extreme plasticity in thermoregulatory behaviors of free-ranging black-tailed prairie dogs. *Physiol Biochem Zool* 79:454–467.
49. Ramirez JM, Folkow LP, Blix AS (2007) Hypoxia tolerance in mammals and birds: From the wilderness to the clinic. *Annu Rev Physiol* 69:113–143.
50. Lovegrove BG, Lawes MJ, Roxburgh L (1999) Confirmation of pleisiomorphic daily torpor in mammals: The round-eared elephant shrew *Macroscelides proboscideus* (Macroscelidea). *J Comp Physiol B Biochem Syst Environ Physiol* 169:453–460.
51. Kawamichi M (1996) Ecological factors affecting annual variation in commencement of hibernation in wild chipmunks (*Tamias sibiricus*). *J Mammal* 77:731–744.
52. Nicol S, Andersen NA (2002) The timing of hibernation in Tasmanian echidnas: Why do they do it when they do? *Comp Biochem Physiol B Biochem Mol Biol* 131:603–611.
53. Voorhies MR (1975) Vertebrate burrows. *The Study of Trace Fossils* (Springer, New York), pp 325–350.
54. Humphries MM, Kramer DL, Thomas DW (2003) The role of energy availability in mammalian hibernation: An experimental test in free-ranging eastern chipmunks. *Physiol Biochem Zool* 76:180–186.
55. Smith AP, Quin DG (1996) Patterns and causes of extinction and decline in Australian conilurine rodents. *Biol Conserv* 77:243–267.
56. Vermeij GJ (1987) *Evolution and Escalation: An Ecological History of Life* (Princeton Univ Press, Princeton).
57. Bown TM, Holroyd PA, Rose KD (1994) Mammal extinctions, body-size, and paleotemperature. *Proc Natl Acad Sci USA* 91:10403–10406.
58. Schmidt-Nielsen K (1984) *Scaling: Why is Animal Size So Important?* (Cambridge Univ Press, Cambridge, UK).
59. Holliday R (2005) Ageing and the extinction of large animals. *Biogerontology* 6:151–156.
60. Damuth J (1981) Population-density and body size in mammals. *Nature* 290:699–700.
61. Damuth J (2007) A macroevolutionary explanation for energy equivalence in the scaling of body size and population density. *Am Nat* 169:621–631.
62. Steininger FF, *et al*., eds (1996) *Circum-Mediterranean Neogene (Miocene-Pliocene) Marine-Continental Chronologic Correlations of European Mammal Units* (Columbia Univ Press, New York), pp 7–46.
63. Fortelius M, *et al*. (1996) Provinciality, diversity, turnover, and paleoecology in land mammal faunas of the Later Miocene of Western Eurasia. *The Evolution of Western Eurasian Neogene Mammal Faunas*, eds Bernor RL, Fahlbusch V, Mittmann H-W (Columbia Univ Press, New York), pp 414–448.
64. Mein P (1975) Proposition de biozonation du Néogène méditerra néen a partir des mammifères. *Trabajos Sobre Neogeno Cuaternario* 4:112–113.
65. Mein P (1989) Updating of MN zones. *European Neogene Mammal Chronology*, eds Lindsay EH, Fahlbusch V, Mein P (Plenum, New York), pp 73–90.
66. Fahlbusch V (1991) The meaning of MN-zonation: Considerations for a subdivision of the European continental Tertiary using mammals. *Newsletters Stratigr* 24:159–173.
67. Foote M (2000) Origination and extinction components of taxonomic diversity: general problems. *Paleobiology* 26:74–102.

68. Liow LH, Stenseth NC (2007) The rise and fall of species: Implications for macroevolutionary and macroecological studies. *Proc Biol Sci* 274:2745–2752.
69. Liow LH (2007) Lineages with long durations are old and morphologically average: An analysis using multiple datasets. *Evolution* 61:885–901.
70. Nichols JD, Morris RW, Brownie C, Pollock KH (1986) Sources of variation in extinction rates, turnover, and diversity of marine invertebrate families during the Paleozoic. *Paleobiology* 12:421–432.
71. Connolly SR, Miller AI (2001) Joint estimation of sampling and turnover rates from fossil databases: Capture-mark-recapture methods revisited. *Paleobiology* 27:751–767.
72. Kroger B (2005)Adaptive evolution in Paleozoic coiled cephalopods. *Paleobiology* 31:253–268.
73. Burnham KP, Anderson DK (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York), 2nd Ed.
74. Heikinheimo H, Fortelius M, Eronen J, Mannila H (2007) Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. *J Biogeogr* 34:1053–1064.
75. Lovegrove BG, Haines L (2004) The evolution of placental mammal body sizes: Evolutionary history, form, and function. *Oecologia* 138:13–27.
76. Morgan ME, *et al*. (1995) Comparative paleoecology of Paleogene and Neogene mammalian faunas: Body-size structure. *Palaeogeogr Palaeoclimatol Palaeoecol* 115:287–317.
77. Alroy J (2003) Taxonomic inflation and body mass distributions in North American fossil mammals. *J Mammal* 84:431–443.
78. Smith FA, *et al*. (2004) Similarity of mammalian body size across the taxonomic hierarchy and across space and time. *Am Nat* 163:672–691.
79. Siemann E, Brown JH (1999) Gaps in mammalian body size distributions reexamined. *Ecology* 80:2788–2792.
80. Wilson DE, Reeder DM, eds (2005) *Mammal Species of the World* (Johns Hopkins Univ Press, Baltimore).
81. Myers P, *et al*. (2007) *The Animal Diversity Web* (Univ of Michigan, Ann Arbor), http://animaldiversity.org, accessed May 27, 2007.
82. Nowak RM (1991) *Walker's Mammals of the World* (Johns Hopkins Univ Press, Baltimore), 5th Ed.
83. Wilson DE, Ruff S (1999) *The Smithsonian Book of North American Mammals* (Smithsonian Institution Press, Washington, DC).
84. Smith F (2007) Maecroecological Database of Mammalian Body Mass (Knowledge Network for Biocomplexity Data Registry) (National Center for Ecological Analysis and Synthesis, Santa Barbara, CA), http://knb.ecoinformatics.org/knb/metacat/nceas.196.3/nceas, accessed August 31, 2007.

Table S1: Average durations of small and large extinct mammals.

Average durations are calculated for genera and species that are found no later than 0.5 Ma in our datasets and are not known to be alive today within our geographic region of focus. Localities are the number of localities represented in the data subsets. N = taxon sample size, Dur (mean) and Dur (med) are mean and median durations respectively (M.y.) . KS = p values from 2-tailed Kolmogorov-Smirnov tests determining if durations of small and large mammals arise from the same underlying distribution. MW = p values from 2-tailed Mann-Whitney U tests determining if the duration distributions of small and large mammals have equal medians. P-values < 0.05 are in bold.

Genera		**Large**			**Small**				
Data subsets	**Localities**	N	Dur (mean)	Dur (med)	N	Dur (mean)	Dur (med)	KS	MW
All	1102	308	3,94	2,90	246	5,34	4,04	**0,007**	**0,011**
5_occ	522	298	3,81	2,90	225	5,54	4,20	**0,002**	**0,001**
5_taxa	1076	156	5,20	5,15	122	7,06	6,25	**0,002**	**0,002**
10_occ_10_taxa	227	110	5,15	4,25	87	6,94	6,55	**0,005**	**0,005**

Species									
All	1106	887	2,17	1,50	868	2,40	1,50	0,872	0,670
5_occ	515	845	2,14	1,50	789	2,42	1,50	0,677	0,462
5_taxa	1015	247	3,26	2,70	247	3,71	2,90	0,587	0,406
10_occ_10_taxa	149	124	3,26	3,05	122	3,60	2,40	0,606	0,986

Table S2: Per capita rates of origination and extinction for large and small mammal species and genera

Mean per capita rates of origination, p and extinction, q in per millions years (darker cells are medians) for data subsets. Bins are in M.y. and MW are p-values for paired, 2 tailed Mann-Whitney U tests and with those which are significant at p < 0.05 marked in bold. N = number of time bins for which comparisons were possible. Because boundary crossers are used, the rates cannot be calculated for time bins at the "edges" of the time series.

Genera

			p					**q**				
Bins	**Data subsets**	**N**	**Large**		**Small**		**MW**	**Large**		**Small**		**MW**
1	*All*	18	0,40	0,18	0,25	0,14	**0,01**	0,30	0,17	0,15	0,12	0,10
	5_occ	17	0,23	0,18	0,13	0,14	**0,00**	0,35	0,18	0,15	0,11	0,06
	5_taxa	18	0,35	0,12	0,27	0,14	0,09	0,23	0,12	0,13	0,09	0,10
	10_occ_10_taxa	17	0,17	0,12	0,16	0,13	0,08	0,34	0,12	0,13	0,07	**0,05**
1,5	*All*	12	0,23	0,22	0,21	0,14	0,14	0,25	0,16	0,15	0,10	0,07
	5_occ	12	0,23	0,24	0,21	0,14	0,20	0,25	0,16	0,14	0,09	0,08
	5_taxa	12	0,19	0,16	0,18	0,12	0,45	0,21	0,13	0,12	0,08	0,08
	10_occ_10_taxa	11	0,18	0,14	0,20	0,10	0,19	0,19	0,11	0,12	0,06	0,07
2	*All*	9	0,41	0,21	0,26	0,16	**0,01**	0,28	0,24	0,15	0,14	**0,01**
	5_occ	8	0,24	0,20	0,14	0,15	**0,03**	0,28	0,25	0,15	0,13	**0,02**
	5_taxa	9	0,36	0,15	0,23	0,14	0,08	0,25	0,20	0,13	0,10	**0,01**
	10_occ_10_taxa	8	0,20	0,16	0,14	0,14	0,20	0,22	0,17	0,14	0,09	**0,05**

Species

Bins	Data subsets	N	p					q				
			Large		Small		MW	Large		Small		MW
1	*All*	18	0,61	0,43	0,49	0,33	**0,05**	0,55	0,36	0,40	0,38	0,10
	5_occ	17	0,48	0,44	0,38	0,32	0,23	0,63	0,34	0,41	0,35	0,35
	5_taxa	17	0,49	0,33	0,40	0,28	0,23	0,49	0,37	0,29	0,22	0,35
	10_occ_10_taxa	14	0,34	0,31	0,34	0,21	0,23	0,40	0,28	0,45	0,25	0,35
1,5	*All*	11	0,44	0,26	0,46	0,33	0,20	0,54	0,22	0,40	0,25	**0,05**
	5_occ	11	0,46	0,43	0,48	0,37	0,20	0,55	0,43	0,40	0,29	**0,05**
	5_taxa	11	0,38	0,30	0,38	0,29	0,20	0,47	0,37	0,32	0,26	**0,05**
	10_occ_10_taxa	9	0,37	0,31	0,36	0,31	0,20	0,42	0,31	0,47	0,27	**0,05**
2	*All*	9	0,62	0,37	0,50	0,26	0,20	0,57	0,33	0,40	0,19	**0,05**
	5_occ	8	0,52	0,47	0,42	0,36	0,20	0,63	0,47	0,43	0,33	**0,05**
	5_taxa	9	0,51	0,40	0,42	0,31	0,20	0,48	0,35	0,31	0,24	**0,05**
	10_occ_10_taxa	5	0,24	0,29	0,19	0,20	0,20	0,37	0,28	0,26	0,29	**0,05**

Table S3. Best models from model selection

Top 3 models from each binning scheme (Bins in M.y.) and various data subsets. QAIC values are corrected with the stated average c-hat from Test 2 and 3 in each case. Weights are model weights which sum to one for all 9 models (see SI Methods for details) although only the top 3 are shown.

Bins	**Data subsets**											
1	*All*			*5_occ*			*5_taxa*			*10_occ_10_taxa*		
	Model	**QAIC**	**Weights**	**Model**	**QAIC**	**Weights**	**Model**	**QAIC**	**Weights**	**Model**	**QAIC**	**Weights**
	Φ{gr+t}pr{t}	1432,45	0,611	Φ{gr+t}pr{gr+t}	1479,751	0,776	Φ{gr+t}pr{gr+t}	1285,151	0,809	Φ{gr+t}pr{t}	947,4754	0,680
	Φ{gr+t}pr{gr+t}	1433,36	0,387	Φ{gr+t}pr{t}	1482,243	0,223	Φ{gr+t}pr{t}	1288,075	0,188	Φ{gr+t}pr{gr+t}	949,0183	0,314
	Φ{t}pr{t}	1444,43	0,002	Φ{t}pr{t}	1495,128	0,000	Φ{t}pr{t}	1297,128	0,002	Φ{t}pr{t}	957,6332	0,004
	c-hat	3,12		c-hat	2,88		c-hat	2,41		c-hat	2,40	
1,5	Φ{gr+t}pr{t}	1308,577	0,507	Φ{gr+t}pr{gr+t}	1489,934	0,525	Φ{gr+t}pr{gr+t}	955,8409	0,649	Φ{gr+t}pr{t}	747,3297	0,495
	Φ{gr+t}pr{gr+t}	1308,673	0,484	Φ{gr+t}pr{t}	1491,066	0,298	Φ{gr+t}pr{t}	957,259	0,319	Φ{gr+t}pr{gr+t}	748,3553	0,297
	Φ{t}pr{t}	1317,218	0,007	Φ{gr*t}pr{t}	1492,584	0,140	Φ{t}pr{t}	962,2637	0,026	Φ{t}pr{t}	749,1989	0,195
	c-hat	2,73		c-hat	2,30		c-hat	2,46		c-hat	2,42	
2	Φ{gr*t}pr{t}	1591,699	0,378	Φ{gr*t}pr{gr+t}	1656,516	0,491	Φ{gr+t}pr{gr+t}	605,7544	0,574	Φ{t}pr{t}	394,4703	0,815
	Φ{gr+t}pr{gr+t}	1592,183	0,297	Φ{gr*t}pr{t}	1656,825	0,421	Φ{gr+t}pr{t}	606,6635	0,364	Φ{gr+t}pr{t}	398,6791	0,099
	Φ{gr*t}pr{gr+t}	1593,373	0,164	Φ{gr+t}pr{gr+t}	1660,467	0,068	Φ{t}pr{t}	610,8268	0,045	Φ{gr+t}pr{gr+t}	399,1409	0,079
	c-hat	1,73		c-hat	1,58		c-hat	3,14		c-hat	3,45	

Table S4. List of NOW genera for which SLOH behavior was estimable.

Genera for which SLOH behavior was estimable, their mean body masses, and type of SLOH behavior present/absent. The last column indicates changes to SLOH coding based on our own knowledge of these mammals, resulting in Fig. S3 as an alternative to text Fig 3. Using MammalBase (see text) to extract information on the life habitats of mammals, we obtained information on 125 NOW genera, as shown in Table S4 (this table). One of us (L.F.) then manually inspected the list and provided alternative life habitat coding. Specifically, Atlantoxerus, Blarinella, Cricetulus, Eolagurus, Myosorex, Sicista, and Urotrichus were identified to have burrowing habitats. These are all classified as small mammals and have low body masses. Considering these seven genera as SLOH in addition to those originally identified as SLOH using MammalBase does not change our qualitative conclusions. We also provide Fig. S3 an alternative plot of Fig. 3 based on these additional seven genera coded as SLOH.

	Mass (g)	Hibernate	Torpor	Dormancy	Aestivation	Fossorial	Burrow	Tunnel	Chambers	Tree Holes	Caves	Alternative
Acinonyx	5,35E+04	0	0	0	0	0	0	0	0	0	0	
Alcelaphus	1,67E+05	0	0	0	0	0	0	0	0	0	0	
Alces	5,13E+05	0	0	0	0	0	0	0	0	0	0	
Allactaga	1,81E+02	1	1	0	1	0	1	0	0	0	0	
Anourosorex	2,00E+01	0	0	0	0	1	0	0	0	0	0	
Aonyx	1,68E+04	0	0	0	0	0	1	0	0	0	0	
Apodemus	3,13E+01	0	0	0	0	0	1	1	0	0	0	
Arvicola	1,60E+02	0	0	0	0	0	1	0	0	0	0	
Atelerix	4,68E+02	1	1	0	1	0	0	0	0	0	0	
Atlantoxerus	6,23E+02	0	0	0	0	0	0	0	0	0	0	Burrower
Axis	6,51E+04	0	0	0	0	0	0	0	0	0	0	
Bison	6,75E+05	0	0	0	0	0	0	0	0	0	0	
Blarinella	1,25E+01	0	0	0	0	0	0	0	0	0	0	Burrower
Bos	8,30E+05	0	0	0	0	0	0	0	0	0	0	
Bubalus	4,25E+05	0	0	0	0	0	0	0	0	0	0	
Calomyscus	2,25E+01	0	0	0	0	0	1	0	0	0	0	
Camelus	4,95E+05	0	0	0	0	0	0	0	0	0	0	
Canis	1,94E+04	0	0	0	0	0	1	0	0	0	0	
Capra	8,23E+04	0	0	0	0	0	0	0	0	0	0	
Capreolus	3,25E+04	0	0	0	0	0	0	0	0	0	0	
Castor	1,85E+04	0	0	0	0	0	0	0	1	0	0	
Ceratotherium	2,50E+06	0	0	0	0	0	0	0	0	0	0	
Cervus	1,12E+05	0	0	0	0	0	0	0	0	0	0	
Clethrionomys	2,75E+01	0	0	0	0	0	0	0	0	0	0	
Condylura	6,25E+01	0	0	0	0	1	1	1	0	0	0	
Cricetulus	8,57E+01	0	0	0	0	0	0	0	0	0	0	Burrower
Cricetus	5,10E+02	1	0	0	0	0	1	1	0	0	0	
Crocidura	1,26E+01	0	0	0	0	0	0	1	0	0	0	
Crocuta	6,30E+04	0	0	0	0	0	1	0	0	0	0	
Cuon	1,55E+04	0	0	0	0	0	0	0	0	0	0	
Dama	7,00E+04	0	0	0	0	0	0	0	0	0	0	
Desmana	3,83E+02	0	0	0	0	0	0	0	0	0	0	
Dicerorhinus	1,40E+06	0	0	0	0	0	0	0	0	0	0	
Diceros	1,10E+06	0	0	0	0	0	0	0	0	0	0	
Dicrostonyx	6,76E+01	0	0	0	0	0	1	0	0	0	0	
Dinaromys	5,60E+01	0	0	0	0	0	0	0	0	0	0	
Dryomys	2,60E+01	1	1	0	0	0	0	0	0	1	0	
Elephas	4,06E+06	0	0	0	0	0	0	0	0	0	0	
Eliomys	1,03E+02	1	0	0	0	0	1	0	0	0	0	
Ellobius	7,43E+01	0	0	0	0	1	0	0	0	0	0	
Eolagurus	2,60E+01	0	0	0	0	0	0	0	0	0	0	Burrower
Eozapus	NA	0	0	0	0	0	0	0	0	0	0	

Episoriculus	5,90E+00	0	0	0	0	0	0	0	0	0	0	
Equus	2,96E+05	0	0	0	0	0	0	0	0	0	0	
Erinaceus	7,50E+02	1	0	0	0	0	1	0	0	0	0	
Felis	4,78E+03	0	0	0	0	0	1	0	0	0	0	
Galemys	5,75E+01	0	0	0	0	0	1	0	0	0	1	
Gazella	4,56E+04	0	0	0	0	0	0	0	0	0	0	
Gerbillus	3,66E+01	0	0	0	0	0	1	1	0	0	0	
Giraffa	1,24E+06	0	0	0	0	0	0	0	0	0	0	
Glirulus	NA	0	0	0	0	0	0	0	0	0	0	
Glis	1,25E+02	1	0	0	0	0	1	1	0	1	0	
Hemiechinus	2,60E+02	1	1	0	0	0	1	0	0	0	0	
Hemitragus	7,50E+04	0	0	0	0	0	0	0	0	0	0	
Herpestes	2,17E+03	0	0	0	0	0	1	0	0	1	0	
Hexaprotodon	2,35E+05	0	0	0	0	0	0	0	0	0	0	
Hippopotamus	2,75E+06	0	0	0	0	0	0	0	0	0	0	
Homo	NA	0	0	0	0	0	0	0	0	0	0	
Hyaena	4,14E+04	0	0	0	0	0	0	0	0	0	0	
Hylopetes	1,66E+02	0	0	0	0	0	0	0	0	0	0	
Hystrix	1,12E+04	0	0	0	0	0	1	1	1	0	0	
Kobus	1,54E+05	0	0	0	0	0	0	0	0	0	0	
Lagurus	3,00E+01	0	0	0	0	0	1	0	0	0	0	
Lemmus	7,60E+01	0	0	0	0	0	1	1	0	0	0	
Lepus	3,98E+03	0	0	0	0	0	1	1	0	0	0	
Lutra	8,50E+03	0	0	0	0	0	0	1	0	0	0	
Lynx	1,17E+04	0	0	0	0	0	1	0	0	0	1	
Macaca	8,48E+03	0	0	0	0	0	0	0	0	0	0	
Marmota	5,25E+03	1	0	1	0	0	1	1	1	0	0	
Martes	1,74E+03	0	0	0	0	0	1	0	0	1	1	
Meles	1,30E+04	0	0	0	0	0	1	1	1	0	0	
Mellivora	1,00E+04	0	0	0	0	0	0	0	0	0	0	
Meriones	7,72E+01	1	0	0	0	0	1	1	1	0	0	
Mesocricetus	1,28E+02	0	1	0	0	0	1	0	0	0	0	
Micromys	6,00E+00	0	0	0	0	0	1	0	0	0	0	
Microtus	4,82E+01	0	0	0	0	1	1	1	0	0	0	
Muntiacus	2,10E+04	0	0	0	0	0	0	0	0	0	0	
Mus	1,47E+01	0	0	0	0	0	1	1	1	0	0	
Muscardinus	2,75E+01	1	0	0	0	0	1	0	0	1	0	
Mustela	5,86E+02	0	0	0	0	0	1	1	0	0	1	
Myomimus	NA	0	0	0	0	0	0	0	0	0	0	
Myosorex	1,33E+01	0	0	0	0	0	0	0	0	0	0	Burrower
Nannospalax	1,60E+02	0	0	0	0	1	0	0	0	0	0	
Neomys	1,50E+01	0	0	0	0	0	1	1	0	0	0	
Neurotrichus	1,00E+01	0	0	0	0	0	0	1	1	0	0	
Nyctereutes	5,00E+03	1	0	0	0	0	1	0	0	0	0	
Ochotona	2,62E+02	0	0	0	0	0	1	1	0	0	0	
Orycteropus	7,00E+04	0	0	0	0	0	1	0	0	0	0	
Oryctolagus	1,80E+03	0	0	0	0	0	1	0	0	0	0	
Oryx	1,55E+05	0	0	0	0	0	0	0	0	0	0	
Ovibos	3,05E+05	0	0	0	0	0	0	0	0	0	0	
Ovis	1,13E+05	0	0	0	0	0	0	0	0	0	0	
Panthera	1,39E+05	0	0	0	0	0	0	0	0	0	0	
Parascalops	6,25E+01	0	0	0	0	1	0	1	0	0	0	

Pelomys	1,08E+02	0	0	0	0	0	0	0	0	0	0	
Phoca	1,15E+05	0	0	0	0	0	0	0	0	0	0	
Prolagus	NA	0	0	0	0	0	0	0	0	0	0	
Puma	5,16E+04	0	0	0	0	0	0	0	0	0	0	
Rangifer	1,89E+05	0	0	0	0	0	0	0	0	0	0	
Rattus	1,76E+02	0	0	0	0	0	1	0	1	0	0	
Ratufa	2,25E+03	0	0	0	0	0	0	0	0	1	0	
Rhagamys	NA	0	0	0	0	0	0	0	0	0	0	
Rupicapra	3,70E+04	0	0	0	0	0	0	0	0	0	0	
Scapanulus	7,50E+01	0	0	0	0	1	0	0	0	0	0	
Scaptonyx	3,00E+01	0	0	0	0	1	0	0	0	0	0	
Sciurotamias	4,54E+02	0	0	0	0	0	0	0	0	0	0	
Sciurus	6,00E+02	0	0	0	0	0	0	0	0	1	0	
Semnopithecus	1,45E+04	0	0	0	0	0	0	0	0	0	0	
Sicista	1,00E+01	0	0	0	0	0	0	0	0	0	0	Burrower
Sorex	1,00E+01	0	0	0	0	0	1	1	0	0	0	
Soriculus	9,95E+00	0	0	0	0	0	0	0	0	0	0	
Spalax	3,93E+02	0	0	0	0	1	0	0	0	0	0	
Spermophilus	5,43E+02	1	1	1	1	0	1	1	1	0	0	
Suncus	1,45E+01	0	1	0	0	0	0	0	0	0	0	
Sus	1,19E+05	0	0	0	0	0	0	0	0	0	0	
Synaptomys	3,55E+01	0	0	0	0	0	1	0	0	0	0	
Talpa	9,22E+01	0	0	0	0	1	1	1	1	0	0	
Tamias	7,48E+01	1	1	0	0	0	1	1	1	1	1	
Tapirus	2,50E+05	0	0	0	0	0	0	0	0	0	0	
Thallomys	1,03E+02	0	0	0	0	0	0	0	0	0	0	
Urotrichus	1,70E+01	0	0	0	0	0	0	0	0	0	0	Burrower
Ursus	2,43E+05	1	1	0	0	0	0	0	0	1	1	
Viverra	8,00E+03	0	0	0	0	0	1	0	0	0	0	
Vormela	5,43E+02	0	0	0	0	0	0	0	0	0	0	
Vulpes	3,55E+03	0	0	0	0	0	1	1	1	0	1	

Paper II

Eronen, J.T., Puolamäki, K., Liu, L., Lintulaakso, K.,
Damuth, J., Janis, C., Fortelius, M.
Precipitation and large herbivorous mammals I: estimates from present-day communities.
Evolutionary Ecology Research, 12(2), 2010.

Evolutionary Ecology Research, 2010, **12**: 217–233

Precipitation and large herbivorous mammals I: estimates from present-day communities

J.T. Eronen[1,2], K. Puolamäki[3], L. Liu[1,4], K. Lintulaakso[1], J. Damuth[5], C. Janis[6] and M. Fortelius[1,7]

[1]*Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland,*
[2]*Department of Computer Science, University of Helsinki, Helsinki, Finland,*
[3]*Department of Media Technology, Helsinki University of Technology, Helsinki, Finland,*
[4]*Laboratory of Evolutionary Systematics of Vertebrates, Chinese Academy of Science, Beijing, China,*
[5]*Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, California, USA,* [6]*Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island, USA and* [7]*Institute of Biotechnology, University of Helsinki, Helsinki, Finland*

ABSTRACT

Question: How can mammalian community characteristics be used to estimate regional precipitation?

Data: Global distribution data of large mammals and their ecomorphology; global climate data.

Research methods: Non-linear regression-tree analysis and linear regression.

Conclusions: The methods unravelled the complex relationships between the environment and the characteristics of mammalian communities. The regression trees described here provide a reasonably accurate estimate of precipitation values for today's world. The strongest correlations are for annual precipitation versus diet ($R^2 = 0.665$), precipitation versus tooth crown height ($R^2 = 0.658$), and precipitation versus diet and tooth crown height combined ($R^2 = 0.742$)

Keywords: climate, community structure, decision trees, herbivorous mammals, hypsodonty, precipitation.

INTRODUCTION

There is a growing need to provide numerical estimates of environmental variables for the present and the past. This is especially true for climate and ecosystem models, which need such estimates as input for boundary conditions. As climatic modelling has moved deeper into the historical past and our data coverage of older time periods has improved, the need

---

Correspondence: J.T. Eronen, Department of Geosciences and Geography, University of Helsinki, PO Box 64, FI-00014 Helsinki, Finland. e-mail: jussi.t.eronen@helsinki.fi

to estimate environmental conditions for pre-Quaternary times has become more urgent. Furthermore, there is a growing need to estimate biotic and abiotic variables for past times to validate the growing number of results from paleoclimate models.

Although there are many different methodologies for reconstructing past environmental conditions based on plant data [e.g. Leaf Margin Analysis (Wolfe, 1979); Climate Leaf Analysis Multivariate Program (Wolfe, 1993); Coexistence Approach (Mosbrugger and Utescher, 1997); Climate Amplitude Method (Fauquette *et al.*, 1998)], data from large fossil mammals, which are both spatially and temporally well-sampled, have not been extensively used for quantitative estimates of past environmental conditions. Large mammals are an important component of the biosphere, and are increasingly threatened by human activities. There is a rich literature on the relationships between large mammals (especially herbivores) and a multitude of biotic and abiotic variables (see below).

It is known that plant productivity and quality are influenced by precipitation and nutrient availability (Reich *et al.*, 1997; Walker and Langridge, 1997). Not surprisingly, herbivore community structure and distributions are in turn affected by plant quality and productivity (e.g. Du Toit *et al.*, 1989; Belovsky, 1997; Owen-Smith, 2002). Previous studies have linked rainfall and animal biomass (Rosenzweig, 1968; Coe *et al.*, 1976), primary productivity and herbivore community biomass (McNaughton *et al.*, 1989), rainfall and species richness of herbivores (Rosenzweig, 1995; Danell *et al.*, 1996), and environmental controls of large herbivore biomass (Olff *et al.*, 2002). These relationships are usually made through the food-source of mammals, and therefore herbivores are the trophic group of choice in such studies.

A number of researchers have shown a relationship between aspects of large mammal herbivore communities and local environmental conditions (e.g. Bell, 1971; Jarman, 1974; Gordon and Illius, 1989; Caughley and Gunn, 1993; Clutton-Brock *et al.*, 1997; Mysterud *et al.*, 2001; Owen-Smith, 2002). More specifically, we have been able to show a relationship between the average value of tooth crown height (hypsodonty) in a herbivore community and the precipitation level of the community's local habitat (Damuth and Fortelius, 2001; J. Damuth, unpublished data). Damuth and Fortelius (2001) introduced and Fortelius *et al.* (2002) applied the use of mean hypsodonty to estimate past precipitation values from fossil molar teeth of large herbivorous mammals. This method has also been applied by Fortelius *et al.* (2003, 2006), Eronen and Rook (2004), and Eronen (2006). Hypsodonty-based estimates have also been used in several studies, including Jernvall and Fortelius (2004) and Eronen *et al.* (2009). Damuth *et al.* (2002) introduced and Janis *et al.* (2004) used a technique called 'per species mean hypsodonty', where mean hypsodonty was related to species diversity (the number of species). The method we introduce here also includes species diversity, although the implementation is different. Here we present a more detailed study on how these aspects of mammalian communities can be used to estimate precipitation.

Tooth crown height (specifically of the cheek teeth used for mastication) is a measure of dental durability. Mammals have only a single adult dentition (i.e. only one set of replacement teeth). Thus, if a herbivore consumes abrasive vegetation, the teeth will be liable to be worn down before the end of the natural lifespan, with repercussions on reproductive output, and hence fitness. To extend the useful life of teeth, the dental materials themselves cannot be made much more durable, so the height of the tooth crown must increase (making the tooth hypsodont). Hypsodonty has evolved multiple times among herbivorous mammals (Janis and Fortelius, 1988). Brachydont, or low crowned, teeth are seen in herbivores that eat relatively non-abrasive food such as soft browse (leaves of dicotyledonous plants) in a relatively grit-free environment. A greater degree of hypsodonty indicates a diet that is more abrasive. This usually means a diet containing greater amounts of grass, but other plants in more arid areas may also contain abrasive dust and grit with similar effects (see Janis and Fortelius, 1988).

Following Janis and Fortelius (1988) and Fortelius *et al.* (2002, 2003, 2006), we argue that hypsodonty is fundamentally an adaptive response to increasing demands for wear tolerance and functional durability of the dentition. Thus, hypsodonty is related to habitats that are open and contain arid-adapted vegetation, whose available plant material is more fibrous and abrasive (Van Valen, 1960; Fortelius, 1985; Janis and Fortelius, 1988; Solounias *et al.*, 1994; Fortelius and Solounias, 2000; C. Janis, unpublished data). In short, we can expect details reflecting regional ecology to be recorded in the herbivore dental morphology (Fortelius and Hokkanen, 2001; Fortelius *et al.*, 2002; Jernvall and Fortelius, 2002). In fact, we contend that communities containing primarily hypsodont herbivores imply a local condition of the consumable vegetation that might be termed 'generalized water stress'. Herein, we will quantify this relationship.

METHODS AND MATERIALS

Precipitation variables were obtained from Hijmans *et al.* (2005; available online at http://www.worldclim.org). The data consist of global climate layers in the 10-min resolution ($18.6 \times 18.6 = 344$ km^2 at the equator) version of the data set. Climate values were converted to a gridded format with a resolution of 0.5° latitude and longitude for grid cells (~55 km at the equator, 720 × 360 cells) to match the mammal data. For climate analysis, we used the following bioclimatic (BIOCLIM) variables that are contained in the WorldClim data set: (1) total annual precipitation; (2) precipitation of the wettest month; (3) precipitation of the driest month; (4) precipitation seasonality, coefficient of variation; (5) precipitation of the wettest quarter; (6) precipitation of the driest quarter; (7) precipitation of the warmest quarter; (8) precipitation of the coldest quarter. Precipitation data are in millimetres.

We selected the orders Artiodactyla, Perissodactyla, and Primates for our investigation. These orders encompass the majority of large herbivorous land mammal species. We excluded the two elephant species (order Proboscidea) from our analysis, but this had a negligible effect on the results (see Discussion below). We also excluded carnivorous omnivores and carnivores because their relationship to climate is more complex than that of herbivores.

The geographic ranges of the mammalian species were derived from the World Wildlife Fund's species distributions (WWF WildFinder; http://www.worldwildlife.org/science/). The ranges are based on occurrences lists for the ecoregion divisions of Olson *et al.* (2001). Ecoregions are contiguous regions across which environmental conditions are similar, as are the fauna and flora. The ranges from this data set are estimated by identifying the ecoregions where the species is present and then making the assumption that the species distribution is ecoregion-wide (i.e. present throughout that entire ecoregion). The ranges used here are thus likely overestimated, covering a larger geographic region than is actually inhabited by the species; however, this is not an issue for this study because the range of climatic conditions associated with the species is not affected by these extensions (since the additional areas belong to the same ecoregions, which have by definition the same climate), and because the ranges estimated here are only slightly different from ranges reported in the literature. [For further details, see Olson *et al.* (2001).]

We converted the WildFinder data to a grid format with a resolution of 0.5° latitude and longitude for grid cells (~55 km at the equator, 720 × 360 cells) and recorded all the species present within each cell. When more than one ecoregion was included in the area of a cell, we chose the ecoregion with the highest proportion of area within the cell. We excluded the Indo-Pacific area because we lack data for most species there. We also excluded Australasia because the herbivores there are marsupials, for which we lack good data. After this conversion, we exported all the grid cells associated with each species to a matrix format where one occurrence of a species in a grid cell is represented by one row.

For ecomorphology, we used the following variables for each species: body mass, body length, diet, and tooth crown height. We gathered the data from published literature sources (Janis, 1988, 1995; Janis and Fortelius, 1988; Nowak, 1991; Wilson and Ruff, 1999; Fortelius and Solounias, 2000; Smith *et al.*, 2003; Wilson and Reeder, 2005; Myers *et al.*, 2007; NatureServe, 2007), and complemented these with unpublished data (from M. Fortelius, C. Janis, J.T. Eronen and L. Liu). For the body mass variable, we used the mean recorded body mass (kg), and for body length, the total mean snout-to-vent length (cm). For diet, we used the following categories: herbivore, frugivore, and omnivore. We further split the herbivore group into grazer, mixed feeder, and browser (based on Hoffmann and Stewart, 1972; Hoffmann, 1989). For tooth crown height, we used three different categories: low crowned (brachydont), medium high crowned (mesodont), and high crowned (hypsodont), as described in Fortelius *et al.* (2002).

Our aim here is to predict a precipitation-related quantity in a grid cell with a set of covariates specific to that cell. We need regression analysis. After preliminary analyses, we decided to concentrate on three precipitation-related climate variables: annual precipitation, precipitation of the wettest quarter, and precipitation of the driest quarter. These are easiest to interpret and use, and they give the strongest correlations with ecomorphological variables. As covariates or predictor variables, we used the average body size, average body mass, a set of variables related to tooth crown height, and a set of variables related to diet. We also report the results for the combination of tooth crown variables and diet variables. The covariates used in this paper are reported in detail in Table 1.

We used both linear regressions and regression trees (Breiman *et al.*, 1984) to predict the annual precipitation (mm). Both are multivariate methods – that is, they can handle several covariates at the same time. We used the implementation of linear regression in GNU R (R Development Core Team, 2009), and the regression tree library *rpart* in GNU R, for all computations (Therneau *et al.*, 2009). The *rpart* library uses cross-validation and other procedures to prune the trees in order to avoid over-fitting the data [see Venables and Ripley (2002) and Breiman *et al.* (1984) for discussion and references about how to construct a regression tree and about the cross-validation procedures]. Note that for linear regressions, one can use stepwise model selection to prune out unnecessary parameters and avoid over-fitting the data (Venables and Ripley, 2002). For example, cladistic analysis uses linear regressions with stepwise model selection. However, unlike linear models, regression trees can express complex non-linear relationships between the covariates.

We used only those grid cells with at least two recorded values of tooth crown height and diet relevant for the analysis. There are 53,054 grid cells and 584,341 taxon occurrences in our data set. The mean number of species per locality is thus 11.014 (= 584,341/53,054). We tested all of the precipitation-related BIOCLIM variables against our mammal ecomorphological variables with all combinations (not reported here).

Linear regression and the regression tree both minimize the squared prediction error between the actual and the predicted values of annual precipitation. We can measure the discrepancy between the actual and predicted values, because we know both the actual value (from the WorldClim data set) and the value predicted by the regression. We use

Table 1. Variables, specific to a grid cell, used as covariates in a regression to estimate the average annual precipitation, precipitation of the wettest quarter, and the precipitation of the driest quarter (all in mm) in the grid cell

Name of set of covariates	Covariates specific to a grid cell
Body size	SVLENGTH (mean animal estimated snout-to-vent length in centimetres, without tusk or tail, in a grid cell)
Body mass	MASS (mean animal body mass in kilograms in a grid cell)
Tooth crown height	NHYP (number of species with hypsodonty data), NHYP1 (number of low crowned species), NHYP2 (number of medium high crowned species), NHYP3 (number of high crowned taxa), pHYP1 (fraction of low crowned species), pHYP2 (fraction of medium high crowned species), pHYP3 (fraction of high crowned taxa)
Diet	NDIET (number of species with diet data), DIETB (number of browsers), DIETF (number of fruit eaters), DIETG (number of grazers), DIETM (number of mixed feeders), DIETO (number of omnivores), pDIETB (fraction of browsers), pDIETF (fraction of fruit eaters), pDIETG (fraction of grazers), pDIETM (fraction of mixed feeders), pDIETO (fraction of omnivores)
Diet + tooth crown height	The ‘tooth crown height’ and ‘diet’ variables listed above, respectively

the R^2 measure to assess the error. The R^2 statistic is commonly interpreted to be the proportion of variance explained by the regression.

To facilitate the use of the regression tree method to resolve environmental proxies, we describe, in the Results section, the step-by-step process of using the tree. Our example is the tree that uses tooth crown height to estimate annual precipitation (Fig. 1). We also provide other regression trees that can be used to estimate annual, wettest, and driest quarter precipitation based on different sets of characteristics (see online Appendix 1: evolutionary-ecology.com/data/2538A1.pdf), together with a description of the variables necessary to use the trees. The regression tree can be visualized as a binary tree (hence the name; for an example, see Fig. 1). To use the tree, traverse it from its root node at the top to a leaf node at the bottom. At each node, choose the left or right branches on the tree based on the covariates: if the condition associated with the node is true, take the left branch, otherwise follow the right branch (see Results for further description). The regression tree has an advantage over a linear regression if the response to the covariates is non-linear, as is often the case in biology. Another advantage of the regression tree is that it can identify context-dependent associations among multiple correlated covariate variables; regression trees do not, for example, assume that the covariates are independent, as is the case with linear regression. Moreover, the same covariate variables can occur several times at different levels of the tree. Finally, a regression tree can be represented in an easily understandable graphical format, allowing ready interpretation of the results. See online Appendix 2 (evolutionary-ecology.com/data/2538A2.pdf) for further discussion on the comparison of linear regression and regression trees, and Legendre and Legendre (1998) for further discussion of linear regression in ecology.

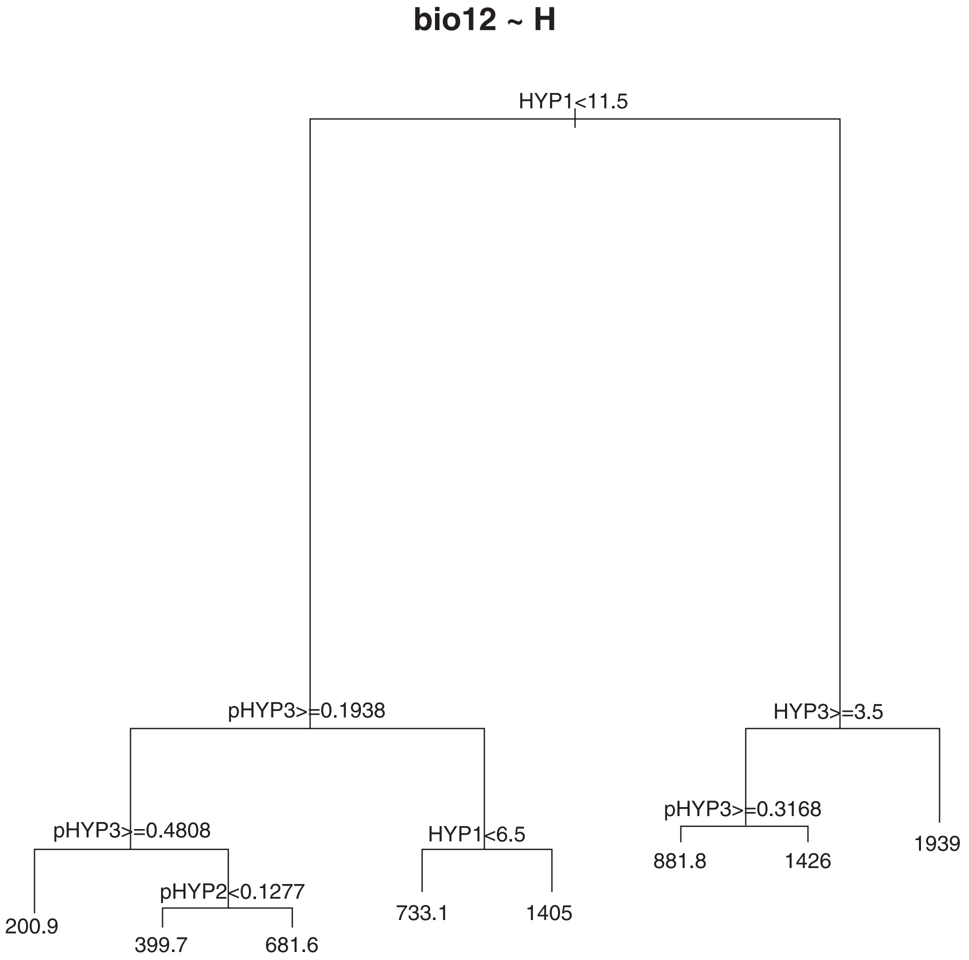


Fig. 1. Decision tree for annual precipitation using hypsodonty alone as regressor (see online Appendix 1 for other decision trees generated: evolutionary-ecology.com/data/2538A1.pdf).

We have ignored spatial autocorrelation and trends in our analysis. [See Diniz-Filho *et al.* (2003) and Hawkins *et al.* (2007) for discussion about the effects of ignoring spatial autocorrelation in regression analysis.] In other words, we have ignored the fact that estimated values (e.g. mean annual precipitation) in nearby grid cells may be correlated. In so doing, we may have overestimated the statistical significance of our results, or over-fitted the regression tree model. Consequently, the resulting regression tree – while more accurate for the present-day mammal and climatic data used to fashion the regression tree structure – may be too detailed. Taking the autocorrelation carefully into account could, in principle, result in smaller regression trees in which some of the branches near the leaf nodes would have been pruned out. However, our regression trees have only 8–12 leaf nodes, meaning that each leaf node describes a large number of grid cells, on average about 5000 cells; hence, we are confident that potential over-fitting is not a serious issue here. Even if there were some over-fitting, the regression tree would still fit present-day data well, as shown by the R^2 values. [For previous use of regression trees in ecology, see, for example, Roff and

Roff (2003), Sankaran *et al.* (2005), Jones *et al.* (2006), and Davidson *et al.* (2009).] Most relevant to our study is Mendoza's (2007) use of decision trees to generate dietary groupings of herbivorous mammals according to ecomorphological species traits.

To visualize the results, we used thematic mapping and grid interpolation in MapInfo 8.5 with the following settings: IDW-interpolation, 50 km grid size, 100 km search radius, 100 grid border. The interpolated maps were then imported as grids to Vertical Mapper, where they were assigned to classes and then contoured to connect the areas with similar values. Inflection values vary between different variables. (See the individual figure legends.)

RESULTS

To estimate the global relationship of precipitation with mammalian characteristics, we used both linear regression and regression-tree methods. The best single correlations for annual precipitation using linear regression are for diet [$R^2 = 0.658$ (mean error = 384 mm)] and tooth crown height [$R^2 = 0.581$ (mean error = 388 mm)]. As single regressors, both body mass and body size have a weak correlation with precipitation ($R^2 = 0.109$ and 0.389, respectively) (Table 2). The best combination for linear correlation is diet and tooth crown height [$R^2 = 0.705$ (mean error = 337 mm)]. When diet is used to estimate precipitation values, the strongest correlations are for annual precipitation ($R^2 = 0.658$) and wettest quarter precipitation ($R^2 = 0.600$) (Table 2). The driest quarter has a weak correlation ($R^2 = 0.350$). Tooth crown height gives a similar range of correlations as diet (annual precipitation = 0.581, wettest quarter = 0.499, driest quarter = 0.334) (Table 2). Our regression tree uses both the proportions and absolute numbers of species in a representative taxon to estimate precipitation. The method uses decimals in absolute numbers (such as 6.5), which is just a technicality of the methodology. Of course, in reality we cannot have

Table 2. Results of the linear regression and regression-tree analysis

	Annual precipitation		Driest quarter precipitation		Wettest quarter precipitation	
Variable	Linear	Tree	Linear	Tree	Linear	Tree
Body size	0.389 (518.33)	0.566 (437.10)	0.110 (87.09)	0.442 (68.93)	0.381 (232.75)	0.496 (209.89)
Body mass	0.109 (625.89)	0.444 (494.57)	0.084 (88.37)	0.416 (70.52)	0.071 (285.06)	0.374 (234.07)
Diet	0.658 (388.05)	**0.665** **(384.11)**	0.350 (74.42)	**0.546** **(62.23)**	0.600 (187.05)	**0.659** **(172.57)**
Tooth crown height	0.581 (429.31)	**0.658** **(387.98)**	0.334 (75.31)	**0.486** **(66.20)**	0.499 (209.38)	**0.663** **(171.69)**
Diet + tooth crown height	0.705 (360.02)	0.742 (337.04)	0.446 (68.71)	0.548 (62.03)	0.649 (175.09)	0.700 (162.02)

Note: The values comparing diet alone and hypsodonty alone are in **bold**. The table shows the R^2 values and the respective standard deviations (in parentheses). The regression trees are shown in Fig. 1 and online Appendix 1 (evolutionary-ecology.com/data/2538A1.pdf).

half of a grazer in a community; for example, >6.5 should be understood as 'more than six'. In our example, we try to estimate the mean annual precipitation of each grid cell based on the characters of the herbivorous mammal community in that grid cell.

The regression tree method returns larger correlation values, suggesting that the mammal characteristics used have a non-linear relationship with precipitation. The correlation values are larger when we use the full set of absolute and relative numbers than when we use only relative numbers. The strongest correlations are for annual precipitation (diet = 0.665, tooth crown height = 0.658, diet and tooth crown height = 0.742) and wettest quarter precipitation (diet = 0.659, tooth crown height = 0.663, diet and tooth crown height = 0.700) (Table 2). The fact that hypsodonty performs almost as well as diet for the annual and wettest quarter precipitation estimation suggests that both variables may capture much of the same environmental signal.

Now, we describe the step-by-step process of using the tree. Our example is the tree that uses tooth crown height to estimate annual precipitation (Fig. 1). The tree starts from the root at the top and splits into two at each node according to the covariates in question. The vertical length of the line between the nodes illustrates the fraction of variance that is explained by the covariate in question. The tree should be read down as far as one of the leaf nodes, thus providing an estimate of the mean annual precipitation based on the particulars of the community in question.

The regression tree for annual precipitation using only tooth crown height information as the estimator yields an accuracy of 0.658 and a mean error of 388 mm (Table 2). The first and most important determinant is the absolute number of brachydont species (Fig. 1). If the number of brachydont species (HYP1) is less than 11.5, we go to the left. If it is more than 11.5, we go to the right. On the second level, the split uses the number and proportion of hypsodont species (pHYP3 and HYP3). Looking at the left branch of the tree, if the proportion of hypsodont species (pHYP3) is larger than 19.38%, the determining factor is again the proportion of hypsodont species (pHYP3). If it is larger than 48.08%, then the precipitation estimate is very low (200.9 mm). If it is small (less than 48.08%), then the proportion of medium high crowned species (pHYP2) determines whether the precipitation is low (399.7 mm) or medium (681.6 mm). Looking at the other branch on the right, we see that if the proportion of hypsodont species (pHYP3) is less than 19.38%, then it is the number of brachydont species (NHYP1) that determines whether the precipitation is medium (733.1 mm) or high (1405 mm). On the right side of the tree, when the number of brachydont species is large, the second level split is determined by the number of hypsodont species (HYP3). If it is large (more than 3), then the proportion of hypsodont species determines whether the precipitation is medium (881.8 mm) or high (1426 mm). If the number of hypsodont species (HYP3) is less than 3, then the precipitation estimate is very high (1939 mm). The height of the branches in the visualization correspond to the variance explained; for example, in the tree in Fig. 1, the choice at the root node (whether HYP1 is above or below 11.5) already explains most of the variation in the data. Relative to this first choice, the other choices, or branches of the tree, account only for fine-tuning of the result.

Consider the following. Cell 10.25 N, 11.75 E, located in West Sudanian Savanna (WWF ecoregion code AT0772, for a description of ecological conditions and fauna, see: http://www.worldwildlife.org/wildworld/profiles/terrestrial/at/at0722_full.html), has the following faunal community structure: number of species (NHYP) = 23; number of brachydont species (HYP1) = 11; number of mesodont species (HYP2) = 3; number of hypsodont species (HYP3) = 9; proportion of brachydont species (pHYP1) = 0.478;

proportion of mesodont species (pHYP2) = 0.13; proportion of hypsodont species (pHYP3) = 0.39. (Online Appendix 3 has the species list; evolutionary-ecology.com/data/2538A3.pdf.) Using the hypsodonty predictive tree (Fig. 1), we will now estimate the annual precipitation for this grid cell. We start from the root (the upmost node). Because the fauna has fewer than 12 brachydont species (HYP1 < 11.5), we continue to the left side in the regression tree. The next node is pHYP3 > 0.1938. Because the proportion of hypsodont species (pHYP3) in this cell is 0.39, we continue to the left. The next node is pHYP3 > 0.4808. As pHYP3 in this cell is 0.39, we go right. The next node is pHYP2 < 0.1277. In this cell, the proportion of mesodont species (pHYP2) is 0.13, so we continue to the right. We arrive at the value of 681.6 mm annual precipitation. The actual annual precipitation (from WorldClim data set) for this grid cell is 743 mm.

To test the geographical resolution of our precipitation estimates, we estimated the mean annual precipitation values for the present-day world using: hypsodonty alone (Fig. 2A); diet alone (Fig. 2B; for regression tree: evolutionary-ecology.com/data/2538A1.pdf); hypsodonty and diet in combination (Fig. 2C and evolutionary-ecology.com/data/2538A1.pdf). For comparison, we also show the actual precipitation map calibrated to the same ranges (Fig. 2D). Plotting the variables on a map makes it easier to understand the environmental context of the splitting sequence in the regression tree. To this end, we also show the number of species with low crowned teeth (Fig. 3A), the proportion of frugivores (Fig. 3B), and the proportion of mixed feeders (Fig. 3C) within the grid cells. These plots show how the variables resolve the spatial pattern, and what regions they resolve best.

DISCUSSION

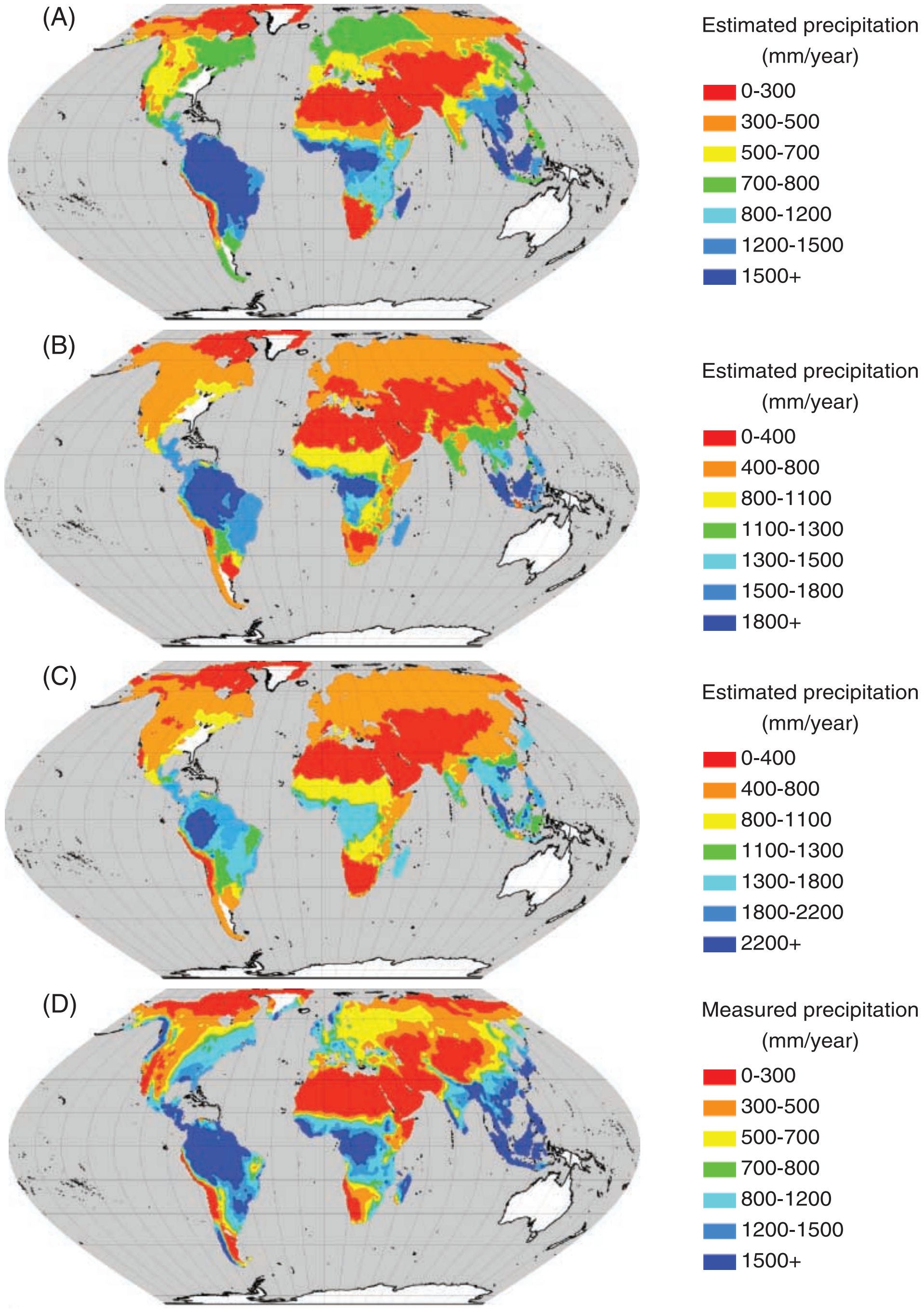
The regression tree appears to be a good way to estimate rainfall because it allows for non-linear responses and it takes the correlations between the input variables into account. Our results show that a non-linear algorithm estimates the precipitation better than does linear correlation. So the community structure of mammals may well respond non-linearly to changes in environmental variation.

Mammal communities reflect the vegetational habitat, which in turn reflects patterns of rainfall. The causal chain between rainfall and animal abundances goes through grass and foliage production, which is approximately linearly related to rainfall (Huxman *et al.*, 2004). The slope of the rainfall–production relationship, which can be expressed as WUE (water use efficiency), is a function of soil nutrient availability (Huxman *et al.*, 2004). In more moist regions, animal productivity is limited not by the gross availability of fodder but by its protein (nitrogen) content (Ellery *et al.*, 1996). Le Houerou (1984) and Le Houerou *et al.* (1988) related rain use efficiency (RUE) to productivity in arid environments. Huxman *et al.* (2004) demonstrated that RUE decreases as rainfall increases. They also related the rainfall to annual net primary production (ANNP) in world biomes. The slope of the ANNP–rainfall relationship decreases in mesic environments and increases in more arid environments (Huxman *et al.*, 2004). This means that production in the arid environments is more rainfall limited, whereas in more humid environments the production is temperature or nutrient limited. In high northern latitudes, primary production is controlled by temperature and light more than by precipitation (Churkina and Running, 1998; Nemani *et al.*, 2003), and there, the relationship between precipitation and mammalian community traits becomes less clear. The same applies to the extremely humid areas where rainfall does not limit primary production. In Figure 2 of

Huxman *et al.* (2004), one can see that in regions with precipitation of over 2000 mm per year, production increases very little when precipitation increases. In our regression tree (Fig. 1), we have only one (or a few; see evolutionary-ecology.com/data/2538A1.pdf) precipitation values in excess of 1500 mm. So our estimation method saturates and mammalian communities cannot be confidently used to estimate rainfall above ~2000 mm. Note, however, that this is a very high level of rainfall in present-day habitats. Most present-day habitats with rainfall exceeding 1500 mm are tropical forests and those over 2000 mm are tropical rainforests. Our regression trees can, in principle, differentiate between tropical forest and tropical rainforest. Beyond saying a place is a tropical rainforest, however, we cannot estimate its precipitation.

The estimated precipitation values presented here have quite a large standard error (see Table 2). This is in part caused by known exceptions to the general trend that result from re-invasion of humid habitats by hypsodont forms. In today's mammal fauna, the main examples of re-invasion are grazing bovid species that range into tropical rainforest (e.g. *Bos gaurus*, *Bos javanicus*, *Syncerus caffer*), and the elephants (*Elephas* and *Loxodonta*), which are highly mobile habitat generalists. We excluded the elephants from the analysis because their large home ranges and broad spectrum of habitats and climate conditions mean that they would contribute mostly noise to our analysis. We retained the bovids, however, because their grazing diet is in harmony with their hypsodonty despite their atypical environmental setting. (The rainforest grasses on which they feed represent an interesting topic for future research that we cannot pursue here.) The exclusion of the two species of elephant from our regressions had only a small effect on the results. They would have contributed to just one further split in the extremely high precipitation estimate, a split that we are not able to resolve well in any case because of the saturation effect discussed above. However, despite these complications, the estimated annual precipitation values based on regression trees using hypsodonty, diet, and hypsodonty plus diet (Fig. 2A–C) depict the actual annual precipitation pattern (Fig. 2D) remarkably well. Africa in particular, excluding Madagascar, is extremely well estimated, and the Eurasian continent is relatively well estimated. The New World is also relatively well estimated, but there are more differences between estimated and actual values than in Africa or Eurasia. Below we discuss some possible reasons for the errors of estimation.

Fig. 2. (A) Predicted annual precipitation based on decision tree analysis using hypsodonty alone as regressor. The precipitation values have been classified to seven classes, reflecting the distribution of prediction outcome from Fig. 1. (B) Predicted annual precipitation based on decision tree for annual precipitation using diet alone as regressor. The precipitation values have been classified to seven classes, reflecting the distribution of prediction outcome from Appendix 1, Tree 1. Note that the classification is slightly different from that in (A). The main difference to (A) is a slight increase in the resolution of the high precipitation range (over 1500 mm). The difference to the pattern of (A) is that the areas of high and extremely high precipitation in the tropics are spatially better differentiated than they are in (A). (C) Predicted annual precipitation based on decision tree for annual precipitation using diet and hypsodonty in combination as regressors. The precipitation values have been classified to seven classes, reflecting the distribution of prediction outcome from Appendix 1, Tree 2. Note that the classification is slightly different from that in (A) and (B). The high precipitation range is better resolved than in (A) and (B) (over 1500 mm range), and the high precipitation areas in the tropics are spatially slightly better differentiated than they are in (A) and (B). (D) Measured annual precipitation from WorldClim data set calibrated to the same classes as (A).



The likely reason for the poor estimates in Madagascar is that it has almost no terrestrial herbivores. That appears to be an historical accident. There never were any large terrestrial herbivores on Madagascar, and most of the large terrestrial and arboreal species have gone

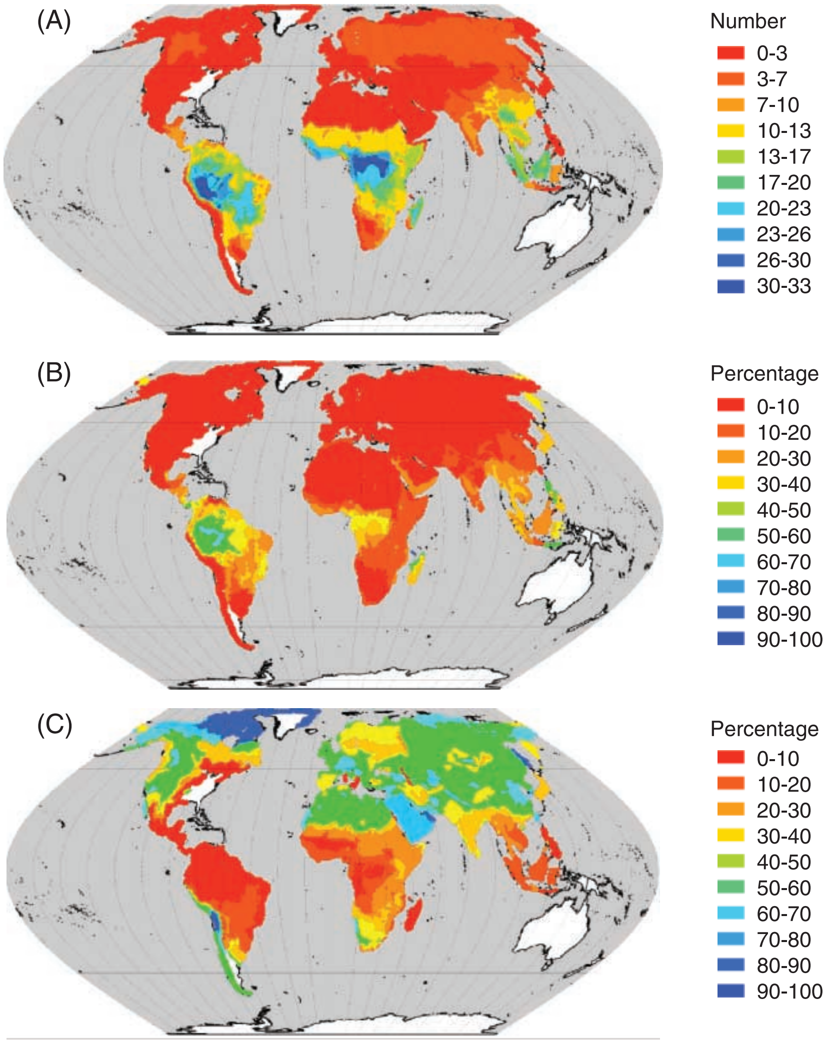


Fig. 3. (A) The number of species with low crowned teeth within grid cells. (B) The proportion of frugivores within grid cells. (C) The proportion of mixed feeders within grid cells.

extinct during the last 10,000 years due to excessive human hunting and human-induced environmental change (Burney *et al.*, 2004). All Madagascan primates have low crowned teeth, as is the case for all primates, and their diet alone is not enough to capture regional environmental variation in Madagascar. The proportion of frugivores (Fig. 3B) does capture some environmental variation, but this variable by itself is not adequate.

In Eurasia, the main differences between the maps (Fig. 2A–C vs. 2D) are in the alpine regions of Europe and the highlands of central India. Neither the hypsodonty nor the diet values capture the precipitation variation in the Alps. This is probably because our mammal data are too coarse to reveal the orographically driven precipitation increase in the Alps or on the western coast of Norway. Comparing Fig. 2A and Fig. 2B reveals that the diet-based and hypsodonty-based prediction both capture some of the variation on the Indian subcontinent, although the regional variation is somewhat different. Their combined effect is seen in Fig. 2C. The orographically driven Indian monsoon in the South Himalayan region (Fig. 2D) is not captured by our estimation method. On the other hand, both hypsodonty- and diet-based estimators, separately and combined, captured the Asian monsoon precipitation regime to a fair extent (Fig. 2A–C). Perhaps our method is too coarse to detect orographically driven changes in precipitation. A likely reason is that mammals are mobile and track the seasonal variation in mountains, which causes mammal species to be recorded in multiple adjacent ecoregions.

The New World exhibits more differences between estimated and actual precipitation values than does the Old World. Part of the reason may be the large-mammal extinctions at the end of the Pleistocene (Koch and Barnosky, 2006). The wetter southeast of North America is not represented in our maps at all because of lack of ecomorphological data in these regions (not enough species with diet and tooth crown data). On the other hand, the drier, western areas of North America are represented in all of the estimation maps (Fig. 2A–C), as well as its more humid northeastern corner. In South America, tropical rainforest, with its extremely high precipitation, covers large areas. This is especially troubling for hypsodonty estimation (Fig. 2A). Yet the combined hypsodonty and diet estimation captures much of the precipitation variation in South America despite the existence of some problematical areas.

We expected that the substantial end-Pleistocene global loss of large mammal diversity (Koch and Barnosky, 2006) that has not been regenerated (Barnosky, 2008) would bias our results in some way. As shown above, there are some small-scale discrepancies, but much of the regional and all of the global scale patterns are well represented in our predictions. This is further confirmed by predictions for Africa. Africa was the continent least affected by end-Pleistocene megafaunal extinctions (Koch and Barnosky, 2006), and indeed Africa does show the best correspondence between predicted and actual rainfall patterns. This is significant for the potential use of this method to estimate climatic conditions of the past, as the African continent has the best available approximation of the mammal community structures that could be thought to represent pre-anthropogenic conditions.

CONCLUSIONS

Estimation methods based on regression trees offer powerful tools to unravel the complex relationships between the environment and the characteristics of mammalian communities. We developed a method to estimate precipitation from a set of ecomorphological characteristics of large herbivorous mammal communities. The regression trees described above all give reasonably accurate estimates of precipitation values for today’s world, and thus we have some confidence that our methodology can be applied to paleocommunities to estimate past environmental conditions. The applicability of different regression trees to paleontological material is severely limited by the little information that is available for living species, such as direct observations of diet. Therefore, the trees that are preferable for reconstruction of past conditions are those that successfully use information derived directly from preserved morphology that correlates with diet, such as hypsodonty.

ACKNOWLEDGEMENTS

This work was supported by the Academy of Finland (M.F.), the Kone Foundation (J.T.E., L.L.), and the Emil Aaltonen Foundation (K.L.). We thank Arne Micheels, Torsten Utescher, Volker Mosbrugger, Louis Francois, and Angela Bruch for discussions on different proxy methods and climate modelling. We thank Jukka Jernvall for discussions on mammal ecomorphology and community structure and Susanna Sova for advice on Malagasy faunas. Finally, we thank Michael Rosenzweig for his thorough review and editing of the paper, which improved it considerably.

REFERENCES

Barnosky, A.D. 2008. Megafauna biomass tradeoff as a driver of Quaternary and future extinctions. *Proc. Natl. Acad. Sci. USA*, **105**(suppl. 1): 11543–11548.

Bell, 1971. A grazing system in Serengeti. *Sci. Am.*, **225**: 86–93.

Belovsky, G.E. 1997. Optimal foraging and community structure: the allometry of herbivore food selection and competition. *Evol. Ecol.*, **11**: 641–672.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. 1984. *Classification and Regression Trees* [software]. Belmont, CA: Wadsworth.

Burney, D.A., Burney, L.P., Godfrey, L.R., Jungers, W.L., Goodman, S.M., Wright, H.T. *et al.* 2004. A chronology for late prehistoric Madagascar. *J. Human Evol.*, **47**: 25–63.

Caughley, G. and Gunn, A. 1993. Dynamics of large herbivores in deserts: kangaroos and caribou. *Oikos*, **67**: 47–55.

Churkina, G. and Running, S.W. 1998. Contrasting climatic controls on the estimated productivity of global terrestrial biomes. *Ecosystems*, **1**: 206–215.

Clutton-Brock, T.H., Illius, A.W., Wilson, K., Grenfell, B.T., MacColl, A.D.C. and Albon, S.D. 1997. Stability and instability in ungulate populations: an empirical analysis. *Am. Nat.*, **49**: 195–219.

Coe, M.J., Cumming, D.H. and Phillipson, J. 1976. Biomass and production of large African herbivores in relation to rainfall and primary production. *Oecologia*, **22**: 341–354.

Damuth, J. and Fortelius, M. 2001. Reconstructing mean annual precipitation, based on mammalian dental morphology and local species richness. In *EEDEN Plenary Workshop on Late Miocene to Early Pliocene Environments and Ecosystems* (J. Agustí and O. Oms, eds.), pp. 23–24. Sabadell, Spain: EEDEN Programme, European Science Foundation.

Damuth, J., Fortelius, M., Andrews, P., Badgley, C., Hadly, E.A., Hixon, S. *et al.* 2002. Reconstructing mean annual precipitation based on mammalian dental morphology and local species richness. *J. Vert. Paleontol.*, **22**(suppl.): 48A.

Danell, K., Lundbarg, P. and Niemelä, P. 1996 Species richness in mammalian herbivores: patterns in the boreal zone. *Ecography*, **19**: 404–409.

Davidson, A.D., Hamilton, M.J., Boyer, A.G., Brown, J.H. and Ceballos, G. 2009. Multiple ecological pathways to extinction in mammals. *Proc. Natl. Acad. Sci. USA*, **106**: 10702–10705.

Diniz-Filho, J.A.F., Bini, L.M. and Hawkins, B.A. 2003. Spatial autocorrelation and red herrings in geographical ecology. *Global Ecol. Biogeogr.*, **12**: 53–64.

Du Toit, J.T. and Owen-Smith, N. 1989. Body size, population metabolism, and habitat specialization among large African herbivores. *Am. Nat.*, **133**: 736–740.

Ellery, W.N., Scholes, R.J. and Scholes, M.C. 1996. The distribution of sweetveld and sourveld in South Africa's grassland biome in relation to environmental factors. *Afr. J. Range Forage Sci.*, **12**: 38–45.

Eronen, J.T. 2006. Eurasian Neogene large herbivorous mammals and climate. *Acta Zool. Fenn.*, **216**: 1–72.

Eronen, J.T. and Rook, L. 2004. The Mio-Pliocene European primate fossil record: dynamics and habitat tracking. *J. Human Evol.*, **47**: 323–341.

Eronen, J.T., Mirzaie Ataabadi, M., Micheels, A., Karme, A., Bernor, R.L. and Fortelius, M. 2009. Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. *Proc. Natl. Acad. Sci. USA*, **106**: 11867–11871.

Fauquette, S., Guiot, J. and Suc, J.-P. 1998. A method for climatic reconstruction of the Mediterranean Pliocene using pollen data. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **144**: 183–201.

Fortelius, M. 1985. Ungulate cheek teeth: developmental, functional, and evolutionary interrelations. *Acta Zool. Fenn.*, **180**: 1–76.

Fortelius, M. and Hokkanen, A. 2001. The trophic context of hominoid occurrence in the later Miocene of western Eurasia – a primate-free view. In *Phylogeny of the Neogene Hominoid Primates of Eurasia* (L. De Bonis, G. Koufos and A. Andrews, eds.), pp. 19–47. Cambridge: Cambridge University Press.

Fortelius, M. and Solounias, N. 2000. Functional characterization of ungulate molars using the abrasion–attrition wear gradient: a new method for reconstructing paleodiets. *Am. Mus. Nov.*, **3301**: 1–36.

Fortelius, M., Eronen, J.T., Jernvall, J., Liu, L., Pushkina, D., Rinne, J. *et al.* 2002. Fossil mammals resolve regional patterns of Eurasian climate change during 20 million years. *Evol. Ecol. Res.*, **4**: 1005–1016.

Fortelius, M., Eronen, J.T., Liu, L.P., Pushkina, D., Tesakov, A., Vislobokova, I. *et al.* 2003. Continental-scale hypsodonty patterns, climatic paleobiogeography and dispersal of Eurasian Neogene large mammal herbivores. In *Distribution and Migration of Tertiary Mammals in Eurasia: A Volume in Honour of Hans De Bruijn* (J.W.F. Reumer and W. Wessels, eds.). *DEINSEA*, **10**: 1–11.

Fortelius, M., Eronen, J.T., Liu, L., Pushkina, D., Tesakov, A., Vislobokova, I. *et al.* 2006. Late Miocene and Pliocene large land mammals and climatic changes in Eurasia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **238**: 219–227.

Gordon, I.J. and Illius, A.W. 1989. Resource partitioning by ungulates on the Isle of Rhum. *Oecologia*, **79**: 383–389.

Hawkins, B.A., Diniz-Filho, J.A.F., Bini, L.M., De Marco, P. and Blackburn, T.M. 2007. Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, **30**: 375–384.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, **25**: 1965 –1978.

Hoffmann, R.R. 1989. Evolutionary steps of ecophysiological adaptation and diversification of ruminants: a comparative view of their digestive system. *Oecologia*, **78**: 443–457.

Hoffmann, R.R. and Stewart, D.R.M. 1972: Grazer or browser: classification based on the stomach structure and feeding habits of East African ruminants. *Mammalia*, **36**: 226–240.

Huxman, T.E., Smith, M.D., Fay, P.A., Knapp, A.K., Shaw, M.R., Lolk, M.E. *et al.* 2004. Convergence across biomes to a common rain-use efficiency. *Nature*, **429**: 651–654.

Janis, C.M. 1988. An estimation of tooth volume and hypsodonty indices in ungulate mammals, and the correlation of these factors with dietary preferences. In *Teeth Revisited: Proceedings of the VIIth International Symposium on Dental Morphology* (D.E. Russell, J.-P. Santoro and D. Sigogneau-Russell, eds.), pp. 367–387. Paris: Mémoirs de Musée d'Histoire Naturelle, series C.

Janis, C.M. 1995. Correlations between craniodental anatomy and feeding in ungulates: reciprocal illumination between living and fossil taxa. In *Functional Morphology in Vertebrate Paleontology* (J.J. Thomason, ed.), pp. 76–98. Cambridge: Cambridge University Press.

Janis, C. and Fortelius, M. 1988. On the means whereby mammals achieve increased functional durability of their dentitions, with special reference to limiting factors. *Biol. Rev. (Camb.)*, **63**: 197–230.

Janis, C.M., Damuth, J. and Theodor, J.M. 2004. The species richness of Miocene browsers, and implications for habitat type and primary productivity in the North American grassland biome. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **207**: 371–398.

Jarman, P.J. (1974). The social organisation of antelope in relation to their ecology. *Behaviour*, **48**: 215–267.

Jernvall, J. and Fortelius, M. 2002. Common mammals drive the evolutionary increase of hypsodonty in the Neogene. *Nature*, **417**: 538–540.

Jernvall, J. and Fortelius, M. 2004. Maintenance of trophic structure in fossil mammal communities: site occupancy and taxon resilience. *Am. Nat.*, **164**: 614–624.

Jones, M.J., Fielding, A. and Sullivan, M. 2006. Analysing extinction risk in parrots using decision trees. *Biodivers. Conserv.*, **15**: 1993–2007.

Koch, P.L. and Barnosky, A.D. 2006. Late Quaternary extinctions: state of the debate. *Annu. Rev. Ecol. Evol. Syst.*, **37**: 215–250.

Legendre, P. and Legendre, R. 1998. *Numerical Ecology*, 2nd edn. Amsterdam: Elsevier.

Le Houerou, H.N. 1984. Rain use efficiency: a unifying concept in arid-land ecology. *J. Arid Environ.*, **7**: 213–247.

Le Houerou, H.N., Bingham, R.L. and Skerbek, W. 1988. Relationship between the variability of primary production and the variability of annual precipitation in world arid lands. *J. Arid Environ.*, **15**: 1–18.

McNaughton, S.J., Oesterheld, M., Frank, D.A. and Williams, K.J. 1989. Ecosystem- level patterns of primary productivity and herbivory in terrestrial habitats. *Nature*, **341**: 142–144.

Mendoza, M. 2007. Decision trees: a machine learning methodology for characterizing morphological patterns resulting from ecological adaptations. In *Automated Taxon Identification in Systematics: Theory, Approaches and Applications* (N. MacLeod, ed.), pp. 261–276. London: Taylor & Francis.

Mosbrugger, V. and Utescher, T. 1997. The coexistence approach – a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, **134**: 61–86.

Myers, P., Espinosa, R., Parr, C.S., Jones, T., Hammond, G.S. and Dewey, T.A. 2007. *Animal diversity* (available at: http://animaldiversity.ummz.umich.edu/site/index.html).

Mysterud, A., Stenseth, N.C., Yoccoz, N.G., Langvatn, R. and Steinheim, G. 2001. Nonlinear effects of large-scale climatic variability on wild and domestic herbivores. *Nature*, **410**: 1096–1099.

NatureServe. 2007. *NatureServe Explorer: An Online Encyclopedia of Life*, version 6.2. Arlington, VA: NatureServe (http://www.natureserve.org/explorer).

Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J. *et al.* 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**: 1560–1563.

Nowak, R.M. 1991. *Walkers' Mammals of the World*. Baltimore, MD: Johns Hopkins University Press.

Olff, H., Ritchie, M. and Prins, H.H.T. 2002. Global environmental controls of diversity in large herbivores. *Nature*, **415**: 901–904.

Olson, D.M., Dinerstein, D., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C. *et al.* 2001. Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**: 933–938.

Owen-Smith, N. 2002. *Adaptive Herbivore Ecology: From Resources to Populations in Variable Environments*. Cambridge: Cambridge University Press.

R Development Core Team. 2009. *R: A Language and Environment for Statistical Computing*. Vienna, Ausatria: R Foundation for Statistical Computing (http://www.R-project.org).

Reich, P.B., Walters, M.B. and Ellsworth, D.S. 1997. From tropics to tundra: global convergence in plant functioning. *Proc. Natl. Acad. Sci. USA*, **94**: 13730–13734.

Roff, D.A. and Roff, R.J. 2003. Of rats and Maoris: a novel method for the analysis of patterns of extinction in the New Zealand avifauna before European contact. *Evol. Ecol. Res.*, **5**: 759–779.

Rosenzweig, M. 1968. Net primary productivity of terrestrial communities: prediction from climatological data. *Am. Nat.*, **102**: 67–74.

Rosenzweig, M.L. 1995. *Species Diversity in Space and Tim*e. Cambridge: Cambridge University Press.

Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S. *et al.* 2005. Determinants of woody cover in African savannas. *Nature*, **438**: 846–849.

Smith, F.A., Lyons, S.K., Ernest, S.K.M., Jones, K.E., Kaufman, D.M., Dayan, T. *et al.* 2003. Body mass of late Quaternary mammals. *Ecology*, **84**: 3403.

Solounias, N., Fortelius, M. and Freeman, P. 1994. Molar wear rates in ruminants: a new approach. *Ann. Zool. Fenn.*, **31**: 219–227.

Therneau, T.M., Atkinson, B. and Ripley, B. 2009. *rpart library*, version 3.1-45. [computer program] (http://cran.r-project.org/web/packages/rpart/index.html).

Van Valen, L. 1960. A functional index of hypsodonty. *Evolution*, **14**: 531–532.

Venables, W.M. and Ripley, B.D. 2002. *Modern Applied Statistics with S*, 4th edn. New York: Springer.

Walker, B.H. and Langridge, J.L. 1997. Predicting savanna vegetation structure on the basis of plant available moisture (PAM) and plant available nutrients (PAN): a case study from Australia. *J. Biogeogr.*, **24**: 813–825.

Wilson, D.E. and Reeder, D.M. 2005. *Mammal Species of the World*. Baltimore, MD: Johns Hopkins University Press.

Wilson, D.E. and Ruff, S. 1999. *The Smithsonian Book of North American Mammals*. Washington, DC: Smithsonian Institution.

Wolfe, J.A. 1979. Temperature parameters of humid and mesic forests of eastern Asia and relation to forests of other regions of the northern hemisphere and Australasia. *US Geol. Surv. Prof. Pap.*, **1106**: 1–37.

Wolfe, J.A. 1993. A method for obtaining climate parameters from leaf assemblages, *US Geol. Surv. Bull.*, **204**: 1–71.

Paper III

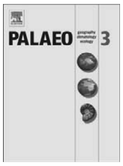
Lintulaakso, K., Kovarovic, K.
Diet and locomotion, but not body size, differentiate mammal communities in worldwide tropical ecosystems.
Palaeogeography, Palaeoclimatology, Palaeoecology, 454, 2016.

Palaeogeography, Palaeoclimatology, Palaeoecology 454 (2016) 20–29

Contents lists available at ScienceDirect

Palaeogeography, Palaeoclimatology, Palaeoecology

journal homepage: www.elsevier.com/locate/palaeo



Diet and locomotion, but not body size, differentiate mammal communities in worldwide tropical ecosystems

Kari Lintulaakso [a,*], Kris Kovarovic [b]

[a] Department of Geosciences and Geography, University of Helsinki, Gustaf Hällströmin katu 2a, 00014 Helsinki, Finland
[b] Department of Anthropology, Durham University, South Road, Durham DH1 3LE, UK

ARTICLE INFO

Article history:
Received 5 August 2015
Received in revised form 31 March 2016
Accepted 5 April 2016
Available online 13 April 2016

Keywords:
CAP
Mammal community structure
NPMANOVA
PERMANOVA
SIMPER
Taxon-free analysis

ABSTRACT

Aim: To test whether tropical habitat groups across the world can be differentiated by using taxon-free mammalian community structures and to discuss the implications of this analysis for palaeoecological community studies.
Materials and methods: We used mammalian community data for 169 localities, which were assigned a priori to hierarchical Olson (1983) vegetation categories. Species over 500 g were classified into dietary, locomotion, and body mass groups and the resulting group structures were analysed using community structure analyses (NPMANOVA, CAP, SIMPER).
Results: The test results show that the mammalian community structures are significantly different between all of Olson's categories. These differences are highest at Olson's major and minor ecosystem levels, and require the least number of variable categories. At the vegetation level, the number of variable categories required to distinguish between them becomes higher. Of the dietary groups, the number of frugivore–granivores, frugivore–omnivores, grazers and mixed feeders contribute most to these differences, while the number of arboreal, arboreal–terrestrial and subterranean–terrestrial species are the key locomotor groups. Body mass was not a good discriminator.
Main conclusions: As general ecosystem categories are broken down into more precisely defined habitats, it requires more detailed knowledge of the species adaptations to distinguish between them. Many of Olson's vegetation groups represent a continuum of cover that are, at least at the worldwide comparison, too detailed to differentiate when broad generalities are sought. We suggest using three worldwide tropical major ecosystems in mammalian community structure analyses: "Humid, closed forests", "Seasonal or interrupted forests and grasslands", and "Seasonal, open drylands". Our results also demonstrate that community structures defined by both dietary and locomotor adaptations are powerful discriminators of tropical ecosystems and habitats across the continents we examined, but body mass should be interpreted with caution when the research question pertains to multiple continents.

1. Introduction

The use of mammalian communities as proxies for past habitats and environmental conditions has a long established history (e.g. Andrews et al., 1979). Such synecological studies may employ a taxonomic approach in which communities with similar taxonomic compositions are considered to have inhabited analogous environments (Simpson, 1943; Fleming, 1973). However, a "taxon-free" approach is favoured by palaeoecologists concerned about the difficulties in identifying species in the fossil record and the assumption that fossil taxa behaved the same as their modern forms (Harrison, 1962; Andrews et al., 1979; Andrews and Hixson, 2014). Such studies focus on how the resident species exploit available spatial and trophic niches irrespective of taxonomic identity through an evaluation of each taxon's adaptations to specific ecological variables, including dietary preferences and locomotor patterns.

Freed from the need to identify specific taxa, communities from different continents and time periods can be compared, giving us the tools to analyse past environmental changes, palaeoclimates and habitat distribution on a global scale. However, many mammalian community studies are restricted to a single geographical realm (e.g. Andrews et al., 1979; Reed, 1997, 1998; Mendoza et al., 2004), with few concentrating specifically on mammalian communities between continents (e.g. Kelt et al., 1996; Andrews and Humphrey, 1999; Rodríguez et al., 2006a). An implicit assumption in these studies is that mammalian communities from the same types of environment will have similar community structures between continents.

Comparisons of communities from different world regions can reveal global generalities, assuming that the relationship between size and abundance of species within specific geographical locations and habitats tells us something about the partitioning of resources in natural communities (Peters and Raelson, 1984; Fa and Purvis, 1997). For

---

[*] Corresponding author.
E-mail address: kari.lintulaakso@gmail.com (K. Lintulaakso).

http://dx.doi.org/10.1016/j.palaeo.2016.04.012
0031-0182/

example, the distribution of species in various classes of body sizes, feeding adaptations, and food habits suggest that the structure of Old and New World tropical mammalian communities is very similar (Fleming, 1973). However, other research demonstrates that despite general resemblances communities from regions with similar climates can radically differ in certain features (Dubost, 1984; Fleagle and Reed, 1996; Fa and Purvis, 1997; Rodríguez et al., 2006a), possibly relating to historic effects. Ecosystems evolve over time (Olson, 1966; Andrews et al., 1979; Janis, 1993); tectonic, regional, historical and climatic processes may all play a role. Mountain uplift (e.g. the rainshadow cast by the uplift of the East African rift; blocking of the monsoonal winds to Central Asia by uplift of Himalayas), changes in oceanic circulation (opening of Drake's passage between Antarctica and Australia, resulting in a cold circumpolar current around Antarctica), closing of the Tethys sea in Late Miocene to form the Mediterranean, isolation (e.g. Australia with its unique fauna), interchange between continents (e.g. the Great American Interchange of the late Pliocene) and major ecosystem changes (e.g. the current arid configuration of Sahara at the end of the Holocene climatic optimum) each contribute to differences between regional species pools and vegetation (Janis, 1993; De Vivo and Carmignotto, 2004; Louys et al., 2011; Fortelius, 2013; Owen-Smith, 2013). This presents a challenge in distinguishing between historical and ecological factors in community structure (Endler, 1982) and likewise suggests that mammal community structures around the world will present some significant differences.

In light of this, we analyse modern mammalian community structures across the tropical belt of four continents: the Americas, Africa, Asia and Australia, which is frequently neglected. We address the following question: can tropical habitats across the world be differentiated on the basis of their mammalian community structures?

2. Methods

2.1. Localities

This study includes 169 extant tropical localities between 23° 30′ N and 23° 30′ S (Fig. 1, Appendix S1 in Supporting information) from previously published studies including Andrews et al. (1979) (n = 23), Andrews and Humphrey (1999) (n = 16), Reed (1997) (n = 18) and Reed (2008) (n = 8) and the database compiled by the National Center for Ecological Analysis and Synthesis (NCEAS) Workshop on Mammalian Communities (see Badgley et al., 2001; van Dam et al., 2001; Damuth et al., 2002) (n = 100). This is supplemented by four unpublished localities from P. Andrews (n = 3) and K. Reed (n = 1). We endeavoured to use only communities comprised of more than 20 species to ensure completeness of the locality species lists, although in some cases (i.e. tundra and montaneous habitats) the number is naturally not likely to be this high (Appendix S1).

Coordinates for each locality were taken from Andrews et al. (1979), Andrews and Humphrey (1999), the NCEAS database and WDPA, the World Database on Protected Areas (IUCN and UNEP, 2009). Site-specific references were also consulted (Appendix S1).

2.2. Vegetation categories

We used the vegetation classification system of the “Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation” GIS dataset, which characterises the vegetative cover of the Earth's land surface in 0.5 by 0.5 degree grid cells (Olson et al., 1983, 1985). The classes can be applied across continents because the system explicitly uses natural vegetation categories without defaulting to generalised physiognomic categories or specific vegetation types (i.e. coconut groves) like those considered in other systems (White and UNESCO, 1983; Eiten, 1992; Lawesson, 1994).

Olson et al. (1983) organises vegetation categories into a hierarchical structure with six levels of increasing complexity. Our analyses are conducted at three different levels corresponding to Olson et al.'s (1983) third, fourth and sixth. Their third level, which describes the landscape, is a broad grouping which we call ‘major ecosystems’. The fourth we call ‘minor ecosystems’, which takes into account latitudinal and altitudinal effects. Their sixth level takes into account temperature, moisture and seasonality. For this we use the term ‘vegetation category’ (Table 1). We excluded seven categories associated with human impact and agriculture (Leemans et al., 1996) and 11 marginal categories that did not describe our localities. Detailed descriptions of selected categories are found in Olson et al. (1983).

Localities may cover tens or thousands of kilometres of space depending on whether they are national parks, small conservancies, etc. This is not unlike other community analyses (e.g. Reed, 1998; Mendoza et al., 2004). Although one can expect there to be some

Fig. 1. 169 localities used in this study. For the list of names and coordinates see Appendix S1.

Table 1
The vegetation categories used in this study with the percentages they represent in the Olson et al. (1983) tropical grid cells.

Major ecosystem		Minor ecosystem		Vegetation		Percentage
FW	Major forest and woodland	MBF	Mostly mid-latitude broad-leaved and mixed	TBC	Deciduous forests, warm woods with conifers	0.9
		TSF	Main tropical/subtropical forest	RGD	Tropical dry forest and woodland	9.2
				TBS	Tropical/subtropical broad-leaved humid forest	11.7
				TRF	Equatorial rain forest	8.3
IW	Interrupted woods	DWM	Other dry woods mosaics	STW	Succulent and thorn woods and shrub	6.9
		TSM	Tropical savanna or montane	SGW	Tropical savanna and woodlands	13.0
				TMC	Tropical montane complexes	2.2
GS	Grass and shrub complexes	MGS	Main grassland or shrubland	MGS	Warm or hot shrub and grassland	15.4
TD	Tundra and desert	SDS	Nonpolar desert or semidesert	SDS	Desert and semidesert (no winter snow)	9.3
		TUN	Tundra, arctic desert, and ice	TUN	Tundra	0.4
WL	Major wetlands	WHW	Warm or hot wetlands	SWP	Warm or hot wetlands, swamp/marsh	2.3

variation in topography and vegetation on the ground, especially where larger localities are concerned, our samples were selected for their relative homogeneity. Each locality was assigned to a vegetation category by comparing the detailed descriptions in our primary sources to those of Olson's to further ensure that we were correct in our interpretation and that each of the localities did not include multiple broad vegetation complexes. The comparison was first made against the vegetation groups within the same grid cells where the localities are situated, then against the neighbouring ones, followed by those applicable to the tropical region of the locality's country and finally for the whole country. The number of localities included in each vegetation category is summarised in Table 2.

2.3. Locality species lists

Species lists for each locality are derived from the published literature (Appendix S1, S2). We only consider species weighing more than 500 g (n = 607), as micromammals may be under-sampled in modern studies and are poorly represented in fossil sites. We restrict our analysis to non-volant mammals and exclude Chiroptera, Sirenia and Cetacea. This follows a well-established practice of excluding these groups in palaeocommunity analyses due to the practical difficulties with including them as fossil species; they do not frequently fossilise or they are fossilised under very different taphonomic conditions compared to other land-dwelling or larger mammals. Given this, and their rarity in the fossil record, most workers thus exclude them (e.g. Andrews et al., 1979; Reed, 1998; Kovarovic et al., 2002; Mendoza et al., 2004; Louys et al., 2011). All of the species were matched against the taxonomy of Wilson and Reeder (2005).

2.4. Ecological variables

We assigned species to categories in four ecological variables: body mass, locomotor pattern, and two hierarchical dietary categories. These data are from NCEAS (Badgley et al., 2001; van Dam et al., 2001; Damuth et al., 2002), PanTHERIA (Jones et al., 2009) and MammalBase (Liow et al., 2008, 2009; Lintulaakso, 2013), a compilation of species attributes and diets based on hundreds of published sources.

2.4.1. Body mass

Taxa were assigned to a body mass category based on median body weights from several published sources (Appendix S2). Mass categories are modified from those in Andrews et al. (1979): 0.5–8 kg (A), 8–45 kg (B), 45–90 kg (C), 90–180 kg (D), 180–360 kg (E), and 360+ kg (F).

2.4.2. Locomotion

We divided the species into six locomotor groups which describe the substrate(s) in which the species moves (Reed, 1998; Miljutin, 2009): arboreal (A), arboreal–terrestrial (AT), subterranean (S), subterranean–terrestrial (ST), terrestrial (T) and terrestrial–aquatic (TA). If a species was found in several different references, the most common locomotion group among these was used (Appendix S2).

2.4.3. Diet

The dietary variable is divided into two levels of food resource preference (Table 3). Diet Level 1 represents the broadest categories of feeding strategy: animalivorous, frugivorous and herbivorous (Miljutin, 2009). Diet Level 2 follows that of Eisenberg (1981) with some modifications: mixed-feeders are added to the herbivore group, and aerial insectivores and foliage-gleaning insectivores are combined into one category: insectivores. As with locomotion, where a species was described in multiple references, the mostly frequently cited dietary preference was used (Appendix S2).

2.5. Analysis

Community structure data may be summarised in two ways: as raw counts of species in each category at each locality or converted to a

Table 2
Number of localities used in the analyses. AF = Africa, AM = America, AS = Asia, and AU = Australia. Ecosystem/vegetation abbreviations are presented in Table 1.

Major ecosystem						Minor ecosystem						Vegetation					
	AF	AM	AS	AU	Total		AF	AM	AS	AU	Total		AF	AM	AS	AU	Total
FW	34	39	23	2	98	MBF		3	1		4	TBC		3	1		4
						TSF	34	36	22	2	94	RGD	16	3	2	2	23
												TBS	11	16	8		35
												TRF	7	17	12		36
IW	23	18	2	1	44	DWM	4	8			12	STW	4	8			12
						TSM	19	10	2	1	32	SGW	12	3		1	16
												TMC	7	7	2		16
GS	11	4		1	16	MGS	11	4		1	16	MGS	11	4		1	16
TD	2	2	1		5	SDS	2		1		3	SDS	2		1		3
						TUN		2			2	TUN		2			2
WL	4	2			6	WHW	4	2			6	SWP	4	2			6
Total	74	65	26	4	169		74	65	26	4	169		74	65	26	4	169

Table 3
Dietary categories used in this study (Eisenberg, 1981; Miljutin, 2009).

Diet Level 1	Diet Level 2
A, animalivore	C, carnivore P, piscivore M, myrmecophage I, insectivore
F, frugivore	FG, frugivore–granivore FH, frugivore–herbivore FO, frugivore–omnivore IO, insectivore–omnivore U, gummivore
H, herbivore	G, grazer MF, mixed-feeder B, browser

percentage of the total number of species at each locality. Percentage data is common (Harrison, 1962; Fleming, 1973; Andrews et al., 1979; Reed, 1997, 1998; Kovarovic et al., 2002; Louys, 2007; Louys et al., 2009, 2011), raw count data much less so (Mendoza et al., 2004, 2005; Rodríguez et al., 2006a). However, the use of percentages is potentially problematic because they generate a closed system (Aitchison, 1986; Hernández Fernández and Vrba, 2006), giving rise to the so called 'closure problem' (Butler, 1981) in which the variables lack statistical independence (Jackson, 1997). It is most noticeable with few variables and decreases in magnitude as the number of variables is increased (Grayson, 1984; Jackson, 1997). As some of our analyses had only three groups, we use the number of species within each group transformed to $\log2(x) + 1$ for $x > 0$, where x is the number of species (Anderson et al., 2006; decostand in R package vegan (Oksanen et al., 2007)). After transformation, a Bray–Curtis distance matrix was calculated (Bray and Curtis, 1957; Legendre and Legendre, 1998). Analyses were conducted in R 3.0.0, using R-packages *stats* and *vegan* (Oksanen et al., 2007; R Development Core Team, 2005).

We tested spatial autocorrelation using the Mantel test (Mantel, 1967). Mantel r of association between Bray–Curtis dissimilarity and distance among sites ranged from a moderate of 0.43 (mass) to a low of 0.16 (locomotion) with randomization p-value of 0.001 (999 permutations, results not shown). Mantel correlograms showed that spatial correlation is highest in 414, 1243, 2072, and 2901 km class indices, i.e. localities which are closer together (within a continent) did tend to have more similar species composition. However, both CAP and NPMANOVA (see below) are non-parametric tests and, as such, are not affected by autocorrelation. If the observations are correlated with one another (e.g., temporally or spatially), then randomly shuffling them will eliminate this inherent structure, if it is present (Anderson, 2005). Therefore, a sensitivity analysis was performed using the original Bray–Curtis distance and the distance between NMDS space. The results were visualized with a Shepard plot (not shown here). For two axes ($k = 2$), the resulting stress values were 0.09, 0.11, 0.12, and 0.15 for body mass, locomotion, Diet Level 1, and Diet Level 2 respectively. Values <0.1 correspond to a good ordination and values <0.2 still provide useful 2-dimensional ordination of the data (Clarke and Warwick, 2001).

We performed a constrained ordination (CAP, canonical analysis of principal coordinates; Anderson and Willis, 2003) to determine if there are differences among mammalian communities. CAP finds the canonical axis that best separates groups in multivariate space, and it also tests the hypothesis of no significant differences in multivariate location among groups. We plotted the product-moment correlation coefficient for each ecological variable value with each of the canonical axes. High absolute correlation is interpreted as an important factor for the nature of group differences (Anderson and Willis, 2003). A non-parametric a priori test for differences among groups was also conducted using the permutational multivariate analysis of variance (PERMANOVA or NPMANOVA; Anderson, 2001), followed by SIMPER (similarity percentage) which determines the relative contribution of each ecological variable to the dissimilarity between groups (Clarke, 1993). If a specific variable consistently contributes to both the within-group similarity and between group dissimilarity, it is considered a good discriminator.

3. Results

3.1. General analyses

Both the NPMANOVA and CAP test results show that the mammalian community structures are significantly different between all categories (NPMANOVA, $p < 0.001$, CAP, $P = 0.005$ Tables 4 & 5).

3.2. Ecological variable analyses

3.2.1. Diet

Most of the communities are significantly different between major ecosystems at both dietary levels (NPMANOVA). In the CAP analysis, the number of frugivorous (mainly frugivore–granivores and frugivore–omnivores) and herbivorous (grazer and mixed feeder) species contribute most to the constrained ordination between major ecosystems (Fig. 2a). SIMPER results indicate that the number of frugivorous species is important at Diet Level 1, and at Diet Level 2 frugivore–granivores contribute most to the dissimilarities between major ecosystem communities. As Olson's categories become more specific (i.e. major ecosystem to minor ecosystem to vegetation category), the number of dietary categories required to distinguish between them becomes higher (Table 6).

3.2.2. Locomotion

Between most of the category pairs, the mammalian community structures are significantly different (NPMANOVA). In the CAP analyses, the number of arboreal, arboreal–terrestrial and subterranean–terrestrial species contribute most to the differences (Fig. 2b). SIMPER results indicate that the number of arboreal or arboreal–terrestrial species is important for the dissimilarities between communities; the subterranean–terrestrial, terrestrial–aquatic, and terrestrial species contributed much less so (Table 7).

3.2.3. Mass categories

In NPMANOVA body mass categories were not as dissimilar between mammalian groups as the other variables. CAP shows that the number of small species contributes most to the group differences (Fig. 2c). In SIMPER, the number of large (360+ kg) and small (0.5–8 kg; 8–

Table 4
Results for NPMANOVA analyses (9999 permutations, Bray–Curtis distance) of different Olson categories in tropical global habitats. All mammalian ecological variables show statistically significant ($p \ll 0.001$) differences within Olson's major and minor ecosystems and vegetation categories.

Variable	Major ecosystem				Minor ecosystem				Vegetation			
	df	F	R^2	p	df	F	R^2	p	df	F	R^2	p
Diet Level 1	4, 164	8.08	0.16	<0.001	7, 161	8.03	0.26	<0.001	10, 158	8.02	0.34	<0.001
Diet Level 2	4, 164	7.48	0.15	<0.001	7, 161	6.41	0.22	<0.001	10, 158	7.50	0.32	<0.001
Locomotion	4, 164	12.96	0.24	<0.001	7, 161	10.24	0.31	<0.001	10, 158	9.68	0.38	<0.001
Mass	4, 164	3.38	0.08	<0.001	7, 161	3.41	0.13	<0.001	10, 158	3.32	0.17	<0.001

Table 5
Results of CAP in tropical global habitats. g = number of Olson groups, P = dimensions, M = number of PCO axes in CAP, Axis1% and Axis2% = percentage of variance explained on axis, constrained eigenvalues of two first axes, df = degrees of freedom (between, within groups), F = approximate F of groups, Sig = the significance of the percentage of correct predictions by CAP.

Olson category	Variable	g	P	M	Axis1%	Axis2%	Eigenvalue 1	Eigenvalue 2	df	F	Sig
Major ecosystem	Diet Level 1	5	3	4	81.18	12.05	0, 360	0, 053	4, 164	6, 06	0, 005
	Diet Level 2	5	12	4	77.80	14.05	1, 302	0, 235	4, 164	5, 76	0, 005
	Locomotion	5	6	4	85.53	8.23	1, 177	0, 113	4, 164	9, 40	0, 005
	Mass	5	6	4	62.95	26.63	0, 288	0, 122	4, 164	2, 67	0, 005
Minor ecosystem	Diet Level 1	8	3	6	67.89	14.71	0, 496	0, 107	7, 161	6, 20	0, 005
	Diet Level 2	8	12	6	60.23	14.49	1, 440	0, 346	7, 161	4, 91	0, 005
	Locomotion	8	6	6	70.33	11.51	11, 267	0, 207	7, 161	7, 43	0, 005
	Mass	8	6	6	63.88	18.72	0, 522	0, 153	7, 161	2, 82	0, 005
Vegetation	Diet Level 1	11	3	6	56.75	24.89	0, 524	0, 230	10, 158	5, 79	0, 005
	Diet Level 2	11	12	6	52.85	22.93	1, 862	0, 808	10, 158	5, 53	0, 005
	Locomotion	11	6	6	64.7	15.72	1, 433	0, 348	10, 158	6, 77	0, 005
	Mass	11	6	6	56.19	25.01	0, 647	0, 288	10, 158	2, 88	0, 005

45 kg) species has the highest contribution to the dissimilarities between category pairs (Table 7).

3.3. Mammalian community structures in Olson's categories

3.3.1. Major forest and woodland (FW) and sub-categories within it

FW differs from the other major ecosystems; the only exception is "major wetlands" (WL). It has the largest number of frugivorous and arboreal species compared with more open ecosystems "grass and shrub complexes" (GS) and "tundra and desert" (TD). SIMPER indicates that the lower number of herbivorous species in FW and the lower number of arboreal species in "interrupted woods" (IW) distinguish these two ecosystems.

At the minor ecosystem level, the "main tropical/subtropical forest" (TSF) had similar results as FW. The "mid-latitude broad-leaved and mixed forest" (MBF) differed from the IW ecosystems by having lower

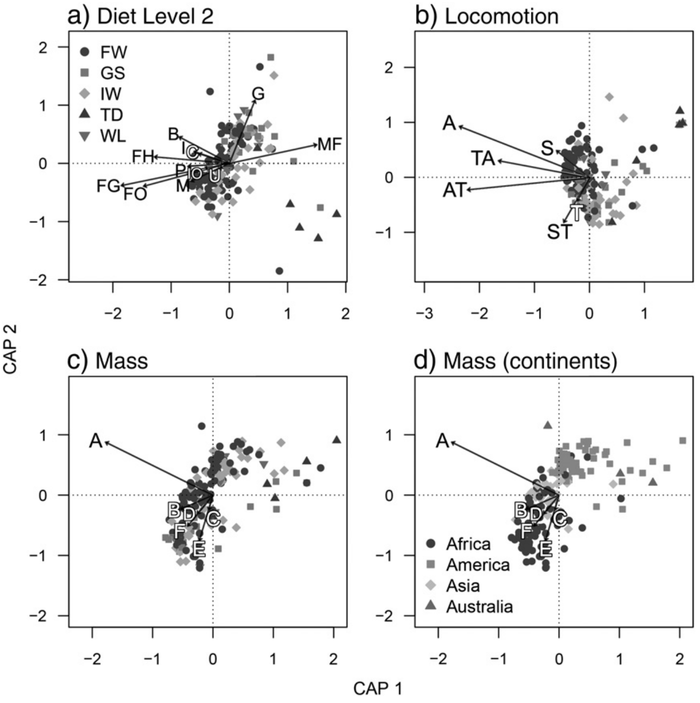


Fig. 2. Canonical analysis of principal coordinates (CAP, Anderson and Willis, 2003) for 169 tropical mammal communities. The communities are analysed by diet (Level 2) (a), locomotion (b), and body mass (c, d). In (a)–(c), communities are grouped by Olson et al. (1983) major ecosystems. In (d), the labels of (c) are replaced by continent names. In the constrained ordination of CAP results, forested, more closed environment can be identified by frugivorous–granivorous and frugivorous–omnivorous species (a), or by arboreal or arboreal–terrestrial species (b). More open environment can be identified by grazers and mixed feeder (a), or by subterranean–terrestrial or terrestrial species (b). Body mass groups do not clearly differentiate mammal communities in the major ecosystems of the Tropics (c); instead the continents have different body mass distribution. This is illustrated in (d), in which the constrained ordination of body masses of (c) is labelled with continent names. Africa has generally larger species than America and Australia. Asian communities have more similar distribution of body masses with Africa than with the other continents. Major ecosystems (Olson et al., 1983) are: major forest and woodland (FW), interrupted woods (IW), grass and shrub complexes (GS), tundra and desert (TD), and major wetlands (WL). Dietary categories are (Eisenberg, 1981): animalivores: carnivore (C), piscivore (P), myrmecophage (M), insectivore (I); frugivores: frugivore–granivore (FG), frugivore–herbivore (FH), frugivore–omnivore (FO), insectivore–omnivore (IO), gummivore (U); herbivores: grazer (G), mixed feeder (MF), and browser (B). Locomotion categories (Miljutin, 2009) are: arboreal (A), arboreal–terrestrial (AT), subterranean (S), subterranean–terrestrial (ST), terrestrial (T) and terrestrial–aquatic (TA). Mass categories are modified from those in Andrews et al. (1979): 0.5–8 kg (A), 8–45 kg (B), 45–90 kg (C), 90–180 kg (D), 180–360 kg (E), and 360+ kg (F).

Table 6
Diet Level 1 and 2 results of pairwise NPMANOVA and SIMPER for 169 tropical mammal communities grouped by vegetation categories (Olson et al., 1983). The asterisks represent statistically significant NPMANOVA results (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). The highlighted dietary categories in the upper left, and their corresponding highlighted asterisks show the first variable that contributes most to these differences (SIMPER analyses). Values above and beside these variables are the mean numbers of species in the analysed groups. For example, the Diet Level 1 communities between minor ecosystems TSF and MGS are significantly different (***$p < 0.001$) in the NPMANOVA analysis. SIMPER indicates that this difference is mostly due to the number of frugivorous (F) species. Reading either down or across from the F column, the mean number of frugivores is 15.2 in TSF and 5.9 in MGS. Major ecosystems: FW = major forest and woodland, IW = interrupted woods, GS = grass and shrub complexes, TD = tundra and desert, WL = major wetlands; minor ecosystems: MBF = mostly mid-latitude broad-leaved and mixed, TSF = main tropical/subtropical forest, DWM = other dry woods mosaics, TSM = tropical savanna or montane, MGS = main grassland or shrubland, SDS = nonpolar desert or semidesert, TUN = tundra, arctic desert, and ice, WHW = warm or hot wetlands; vegetation: TBC = deciduous forests, warm woods with conifers, RGD = tropical dry forest and woodland, TBS = tropical/subtropical broad-leaved humid forest, TRF = equatorial rain forest, STW = succulent and thorn woods and shrub, SGW = tropical savanna and woodlands, TMC = tropical montane complexes, MGS = warm or hot shrub and grassland, SDS = desert and semidesert (no winter snow), TUN = tundra, SWP = warm or hot wetlands, swamp/marsh; Diet Level 1: F = frugivore, H = herbivore; Diet Level 2: FG = frugivore–granivore, FH = frugivore–herbivore, FO = frugivore–omnivore, B = browser, G = grazer, C = carnivore.

Major Ecosystem
Diet 1 and mean # of species

			IW	GS	TD	WL
	F		10.1	5.9	1.8	8.2
		H	9.1	13.2	5.2	9.7
			Pairwise NPMANOVA and SIMPER			
FW	14.8	8.2	*	***	***	-
IW	10.1	9.1		*	**	-
GS	5.9	13.2			-	-
TD	1.8	5.2				*

Minor Ecosystem
Diet 1 and mean # of species

			TSF	DWM	TSM	MGS	SDS	TUN	WHW
	F		15.2	7.2	11.2	5.9	3.0	0.0	8.2
		H	8.4	8.8	9.2	13.2	7.0	2.5	9.7
			Pairwise NPMANOVA and SIMPER						
MBF	6.2	4.0	**	-	*	-	-	-	-
TSF	15.2	8.4		**	-	***	*	**	-
DWM	7.2	8.8			*	**	*	*	-
TSM	11.2	9.2				-	-	**	-
MGS	5.9	13.2					-	*	-
SDS	3.0	7.0						-	-
TUN	0.0	2.5							*

Vegetation
Diet 1 and mean # of species

			RGD	TBS	TRF	STW	SGW	TMC	MGS	SDS	TUN	SWP
	F		9.0	15.3	19.1	7.2	9.1	13.3	5.9	3.0	0.0	8.2
		H	14.2	6.9	6.1	8.8	12.6	5.8	13.2	7.0	2.5	9.7
			Pairwise NPMANOVA and SIMPER									
TBC	6.2	4.0	*	**	***	-	-	-	-	-	-	-
RGD	9.0	14.2		***	***	-	-	***	-	-	*	-
TBS	15.3	6.9			*	**	*	-	***	**	**	-
TRF	19.1	6.1				***	***	**	***	***	**	**
STW	7.2	8.8					-	**	**	*	*	-
SGW	9.1	12.6						**	-	-	*	-
TMC	13.3	5.8							**	*	**	-
MGS	5.9	13.2								-	*	-
SDS	3.0	7.0									-	-
TUN	0.0	2.5										*

Major Ecosystem
Diet 2 and mean # of species

			IW	GS	TD	WL
	FG		3.0	1.1	0.0	2.8
		G	4.4	6.5	1.4	5.7
			Pairwise NPMANOVA and SIMPER			
FW	5.8	3.1	*	***	***	-
IW	3.1	4.4		*	***	-
GS	1.1	6.5			*	-
TD	0.0	1.4				**

Minor Ecosystem
Diet 2 and mean # of species

					TSF	DWM	TSM	MGS	SDS	TUN	WHW
	FG				5.9	1.8	3.6	1.1	0.0	0.0	2.8
		FH			5.8	2.9	5.3	4.0	2.3	0.0	3.7
			B		4.3	2.7	3.8	4.5	2.0	0.0	3.2
				G	3.2	5.0	4.2	6.5	2.0	0.5	5.7
					Pairwise NPMANOVA and SIMPER						
MBF	2.0	2.5	2.8	1.0	**	-	*	*	-	-	-
TSF	5.9	5.8	4.3	3.2		***	-	***	**	***	-
DWM	1.8	2.9	2.7	5.0			*	*	-	*	-
TSM	3.6	5.3	3.8	4.2				**	-	**	-
MGS	1.1	4.0	4.5	6.5					-	*	-
SDS	0.0	2.3	2.0	2.0						-	*
TUN	0.0	0.0	0.0	0.5							*

Vegetation
Diet 2 and mean # of species

							RGD	TBS	TRF	STW	SGW	TMC	MGS	SDS	TUN	SWP
	FG						1.6	6.0	8.7	1.8	2.1	5.1	1.1	0.0	0.0	2.8
		FH					4.9	5.6	6.7	2.9	5.2	5.4	4.0	2.3	0.0	3.7
			FO				2.4	3.7	3.7	2.4	1.8	2.9	0.9	0.7	0.0	1.7
				B			5.3	3.8	4.2	2.7	4.1	3.4	4.5	2.0	0.0	3.2
					G		8.0	2.2	1.1	5.0	6.8	1.5	6.5	2.0	0.5	5.7
						C	13.0	10.2	9.6	10.7	11.9	8.9	11.4	9.7	3.0	11.0
							Pairwise NPMANOVA and SIMPER									
TBC	2.0	2.5	1.8	2.8	1.0	9.2	*	**	***	-	-	*	*	-	-	-
RGD	1.6	4.9	2.4	5.3	8.0	13.0		***	***	*	-	***	*	*	**	-
TBS	6.0	5.6	3.7	3.8	2.2	10.2			**	**	**	-	***	**	**	*
TRF	8.7	6.7	3.7	4.2	1.1	9.6				***	***	-	***	***	**	**
STW	1.8	2.9	2.4	2.7	5.0	10.7					-	***	*	-	*	-
SGW	2.1	5.2	1.8	4.1	6.8	11.9						**	-	-	**	-
TMC	5.1	5.4	2.9	3.4	1.5	8.9							***	*	**	-
MGS	1.1	4.0	0.9	4.5	6.5	11.4								-	*	-
SDS	0.0	2.3	0.7	2.0	2.0	9.7									-	*
TUN	0.0	0.0	0.0	0.0	0.5	3.0										*

number of subterranean–terrestrial species. MBF had also fewer terrestrial species than "main grassland or shrubland" (MGS), and more arboreal–terrestrial species than "nonpolar desert or semidesert" (SDS).

At the vegetation level, TSF has three sub-groups. These groups are dissimilar between each other and most of the other vegetation groups. However, there were no significant differences between the "tropical dry forest and woodland" (RGD) vs. "tropical savanna and woodlands" (SGW) and "tropical/subtropical broad-leaved humid forest" (TBS) vs. "tropical montane complexes" (TMC) pairs (Tables 6, 7).

3.3.2. Interrupted woods (IW) and sub-categories within it

IW differed from the other major ecosystems, WL being the only exception. However, at the minor ecosystem level, the "other dry woods mosaics" (DWM) could be differentiated from the wetlands by having fewer arboreal species.

At the vegetation level, there were no significant differences between SGW and the "warm or hot shrub and grassland" (MGS) nor between the SGW and RGD as mentioned earlier (Tables 6, 7).

3.3.3. Grass and shrub complexes (GS)

This major ecosystem had only one minor ecosystem and vegetation group, MGS. GS differs from the other major ecosystems at some of the four community variables. At the minor ecosystem level, MGS did not differ from "nonpolar desert or semidesert" (SDS) and, at the vegetation level, MGS was not different from "tropical savanna and woodlands" (SGW) or "desert and semidesert (no winter snow)" (SDS) (Tables 6, 7).

3.3.4. Tundra and desert (TD) and sub-categories within it

TD differs from the other major ecosystems. It had the least number of species in all of the community variables. TD is divided into two minor ecosystems, SDS and "tundra, arctic desert and ice" (TUN[1]), but they do not differ. As mentioned above, MGS and SDS were similar with each other, but MGS and TUN were not (Tables 6, 7).

---

[1] The only localities from TUN were high Alpine tundra localities from South America.

Table 7

Locomotion and mass category results of pairwise NPMANOVA and SIMPER for 169 tropical mammal communities grouped by vegetation categories (Olson et al., 1983). The asterisks represent statistically significant NPMANOVA results (*p < 0.05; **p < 0.01; ***p < 0.001). The highlighted dietary categories in the upper left and their corresponding highlighted asterisks show the first variable that contributes most to these differences (SIMPER analyses). Values above and beside these variables are the mean numbers of species in the analysed groups. For example, the locomotion communities between minor ecosystems TSF and MGS are significantly different (***p < 0.001) in the NPMANOVA analysis. SIMPER indicates that this difference is mostly due to the number of arboreal (A) species. Reading either down or across from the A column, the mean number of arboreal species is 8.1 in TSF and 1.7 in MGS. Major ecosystems: FW = major forest and woodland, IW = interrupted woods, GS = grass and shrub complexes, TD = tundra and desert, WL = major wetlands; minor ecosystems: MBF = mostly mid-latitude broad-leaved and mixed, TSF = main tropical/subtropical forest, DWM = other dry woods mosaics, TSM = tropical savanna or montane, MGS = main grassland or shrubland, SDS = nonpolar desert or semidesert, TUN = tundra, arctic desert, and ice, WHW = warm or hot wetlands; vegetation: TBC = deciduous forests, warm woods with conifers, RGD = tropical dry forest and woodland, TBS = tropical/subtropical broad-leaved humid forest, TRF = equatorial rain forest, STW = succulent and thorn woods and shrub, SGW = tropical savanna and woodlands, TMC = tropical montane complexes, MGS = warm or hot shrub and grassland, SDS = desert and semidesert (no winter snow), TUN = tundra, SWP = warm or hot wetlands, swamp/marsh; locomotion groups: A = arboreal, AT = arboreal–terrestrial, S = subterranean, ST = subterranean–terrestrial, T = terrestrial, TA = terrestrial–aquatic; mass categories: A = 0.5–8 kg, B = 8–45 kg, E = 180–360 kg, F = 360 + kg.

Major Ecosystem

Locomotion and mean # of species

				IW	GS	TD	WL
A				3.8	1.7	0.2	3.5
	AT			4.8	4.1	1.0	3.5
		TA		1.8	1.3	0.2	3.0
				Pairwise NPMANOVA and SIMPER			
FW	7.9	6.1	2.8	***	***	***	-
IW	3.8	4.8	1.8		*	***	-
GS	1.7	4.1	1.3			*	*
TD	0.2	1.0	0.2				**

Minor Ecosystem

Locomotion and mean # of species

					TSF	DWM	TSM	MGS	SDS	TUN	WHW
A					8.1	1.2	4.7	1.7	0.3	0.0	3.5
	AT				6.2	5.0	4.8	4.1	1.7	0.0	3.5
		ST			3.2	5.4	3.5	3.2	2.7	1.0	3.8
			T		17.3	18.5	18.1	22.6	15.7	6.0	18.5
				TA	2.9	1.7	1.9	1.3	0.3	0.0	3.0
					Pairwise NPMANOVA and SIMPER						
MBF	2.8	4.5	1.2	10.5	1.5	**	*	*	**	*	-	-
TSF	8.1	6.2	3.2	17.3	2.9		***	***	***	***	***	*
DWM	1.2	5.0	5.4	18.5	1.7			*	-	**	**	*
TSM	4.7	4.8	3.5	18.1	1.9				*	**	**	-
MGS	1.7	4.1	3.2	22.6	1.3					-	*	*
SDS	0.3	1.7	2.7	15.7	0.3						-	*
TUN	0.0	0.0	1.0	6.0	0.0							*

Vegetation

Locomotion and mean # of species

					RGD	TBS	TRF	STW	SGW	TMC	MGS	SDS	TUN	SWP
A					3.4	8.1	11.2	1.2	2.9	6.5	1.7	0.3	0.0	3.5
	AT				4.4	6.3	7.2	5.0	4.1	5.4	4.1	1.7	0.0	3.5
		ST			3.7	2.9	3.1	5.4	4.4	2.6	3.2	2.7	1.0	3.8
			T		24.2	15.7	14.6	18.5	22.6	13.6	22.6	15.7	6.0	18.5
				TA	2.6	2.4	3.4	1.7	1.6	2.2	1.3	0.3	0.0	3.0
					Pairwise NPMANOVA and SIMPER									
TBC	2.8	4.5	1.2	10.5	1.5	*	**	**	*	**	-	**	*	-	-
RGD	3.4	4.4	3.7	24.2	2.6		***	***	***	-	**	**	**	**	-
TBS	8.1	6.3	2.9	15.7	2.4			**	***	***	-	***	***	**	**
TRF	11.2	7.2	3.1	14.6	3.4				***	***	***	***	***	**	***
STW	1.2	5.0	5.4	18.5	1.7					-	***	-	**	*	*
SGW	2.9	4.1	4.4	22.6	1.6						***	-	*	**	-
TMC	6.5	5.4	2.6	13.6	2.2							**	**	**	-
MGS	1.7	4.1	3.2	22.6	1.3								-	*	*
SDS	0.3	1.7	2.7	15.7	0.3									-	-
TUN	0.0	0.0	1.0	6.0	0.0										*

Major Ecosystem

Body mass and mean # of species

				IW	GS	TD	WL
A				15.6	11.6	5.4	14.2
	B			8.9	10.2	5.6	8.8
		F		1.9	2.8	0.2	2.0
				Pairwise NPMANOVA and SIMPER			
FW	21.0	8.3	1.7	-	**	***	-
IW	15.6	8.9	1.9		-	***	-
GS	11.6	10.2	2.8			*	-
TD	5.4	5.6	0.2				-

Minor Ecosystem

Body mass and mean # of species

				TSF	DWM	TSM	MGS	SDS	TUN	WHW
A				21.2	15.6	15.6	11.6	7.0	3.0	14.2
	B			8.6	8.3	9.2	10.2	8.3	1.5	8.8
		F		1.8	1.6	2.0	2.8	0.3	0.0	2.0
				Pairwise NPMANOVA and SIMPER						
MBF	14.2	2.8	0.8	**	-	*	-	-	-	-
TSF	21.2	8.6	1.8		-	-	**	**	***	-
DWM	15.6	8.3	1.6			-	-	-	*	-
TSM	15.6	9.2	2.0				-	-	**	-
MGS	11.6	10.2	2.8					-	*	-
SDS	7.0	8.3	0.3						-	-
TUN	3.0	1.5	0.0							*

Vegetation

Body mass and mean # of species

				RGD	TBS	TRF	STW	SGW	TMC	MGS	SDS	TUN	SWP
A				15.5	20.8	25.4	15.6	14.2	16.9	11.6	7.0	3.0	14.2
	B			10.0	8.1	8.1	8.3	10.6	7.8	10.2	8.3	1.5	8.8
		F		3.0	1.6	1.2	1.6	2.7	1.4	2.8	0.3	0.0	2.0
				Pairwise NPMANOVA and SIMPER									
TBC	14.2	2.8	0.8	*	**	**	-	-	-	-	-	-	-
RGD	15.5	10.0	3.0		**	***	-	-	**	-	*	*	-
TBS	20.8	8.1	1.6			*	-	-	-	*	**	**	-
TRF	25.4	8.1	1.2				*	**	*	***	***	**	-
STW	15.6	8.3	1.6					-	-	-	-	**	-
SGW	14.2	10.6	2.7						-	-	-	*	-
TMC	16.9	7.8	1.4							-	-	*	-
MGS	11.6	10.2	2.8								-	*	-
SDS	7.0	8.3	0.3									-	-
TUN	3.0	1.5	0.0										*

3.3.5. Major wetlands (WL)

This major ecosystem had only one vegetation category, "warm or hot wetlands, swamp/marsh" (SWP). WL is the only major ecosystem that does not differ from FW and IW. At the vegetation level, it can only be differentiated from forested (TBS, TRF) or open habitats (SDS, TUN) (Tables 6, 7).

4. Discussion

We tested whether tropical habitats across the world can be differentiated on the basis of their mammalian community structures. Generally, our results show that the diet and locomotion categories differentiate the communities best, while body mass performed poorly. As Olson's categories become more tightly defined, the number of ecological variable groups required for differentiation becomes higher. Differences between treeless categories, such as desert and grasslands, are not significant.

4.1. Ecological variables

Differences noted between dietary categories are largely accounted for by the number of frugivores. At Diet Level 2, this signal comes from frugivore–granivores (FG) and frugivore–herbivores (FH) (but see ordination in CAP where frugivore–omnivores are also important, Fig. 2a). Many of these species are primates but there are other frugivorous groups, such as ungulates, also present in high numbers in tropical forests which are under-represented in open habitats such as savannas and grasslands (Bodmer, 1990).

Others have found that the separation of vegetation groups relates more to the grazer/browser continuum, rather than the abundance of frugivores (Sponheimer et al., 1999; Janis et al., 2000). In our study, the number of herbivores (mainly grazers and browsers in SIMPER, grazers and mixed-feeders in the CAP ordination) is important only at the minor ecosystem and vegetation level where "interrupted woods" (IW) and treeless areas like grasslands and tundra, are differentiated from each other. Animalivorous groups do not impact on the differences between major or minor ecosystems but, at the vegetation level, the low number of carnivorous species in MGS distinguishes this habitat from RGD. Similar results were obtained by Reed (1997, 1998) where frugivores, fruit/insect eaters and grazers discriminated between habitats with no clear contribution from animalivorous taxa. Herbivores and frugivores are mainly primary or secondary consumers while animalivores are higher in the trophic hierarchy. This may be the reason why habitats are reflected more by their primary and secondary consumers and the 'signal' from the primary production weakens towards the top of the trophic hierarchy.

When we compare results between Diet Level 1 and 2, the differentiation of the major and minor ecosystems and the vegetation categories is based on similar variables which relate to the abundance of frugivores and herbivores. The differences are more nuanced when we consider Diet Level 2, with between one and six variables required to distinguish between various habitat pairs. The broader classification (i.e. Diet Level 1), requiring a consideration of only frugivory and herbivory, may make it a more robust approach for worldwide ecosystem and vegetation comparisons, although clearly less detail is captured.

In terms of the locomotion groups, the number of arboreally adapted species differentiates closed (FW, IW) from open habitats (GS, TD). This is a conclusion borne out by many others (Reed, 1997; Louys et al., 2011) who also demonstrate that an abundance of arboreal species indicates the presence of well-developed tree cover. Primates and tree-dwelling rodents comprise the majority of these taxa. The higher number of terrestrial–aquatic species in wetlands (WL) only separated it from TD; otherwise WL did not differentiate from FW and IW. We conjecture that the species lists from wetlands include taxa from nearby areas that periodically visit for drinking and maybe counted as inhabitants of both areas because they are observed in each. Interestingly, the number of terrestrial species contributed very little to the differences between community structures in SIMPER. Terrestriality may be too general of a category for describing species that move about on the land surface. By adding more detailed locomotor classes (i.e. cursorial), the separation between open and closed habitats could be clearer. In the CAP analysis, the number of subterranean–terrestrial and terrestrial species did distinguish between humid and forested localities (TBS, TRF) and seasonal or less forested areas (RGD, STW, SGW, MGS). Seasonality and the inconsistent availability of food throughout the year can be dealt with in two ways: migration, which requires terrestrial movement, or by caching food resources, which subterranean–terrestrial species generally do.

Body mass categories did not clearly differentiate community structures, particularly in vegetation categories. This may be a consequence of the scale of the investigation. Work by Brown and Maurer (1989) has shown that the frequency distribution of body sizes appears to be general; it is modal and skewed towards smaller species when large areas (i.e. continents or biomes) are sampled, whereas there are approximately equal numbers of species in each size category in smaller, local areas. They suggest that a strong negative interaction prevents local coexistence of similar-sized species, so one would therefore not expect that community structures could be distinguished at this scale. However, further analysis of the data showed that the size ranges of mammals differed on each continent. Africa has a higher proportion of larger mammals, whilst America and Australia have smaller species with Asia somewhere in between (Fig. 2d). This pattern is also observable in the palaeontological record of the Pleistocene and is likely to have roots deep in time (Janis, 1993; De Vivo and Carmignotto, 2004; Louys et al., 2011; Fortelius, 2013; Owen-Smith, 2013).

We found that only the number of small (0.5–8 kg, 8–45 kg) and large (360 + kg) species has the highest contribution (SIMPER) for some of the community pairs. The importance of small species is likely that, as a group, there is more variety in both dietary preferences and locomotion. Large species are almost always terrestrial herbivores, separating grasslands and savannas from forests, whereas small species can be arboreal, terrestrial, subterranean, as well as insectivores, frugivores, and herbivores. Although a recent study (Rodríguez et al., 2006b) has shown that body size can be useful, this is only clear at higher latitudes suggesting that the warmer the annual temperature, the less likely body size is helpful in distinguishing habitats.

4.2. Distinguishing Olson's categories

Generally, major ecosystems can be differentiated by the dietary and locomotor adaptations of their constituent species, but body mass is not a robust discriminator. For example, mass does not differentiate the FW–IW–GS continuum. This is likely to relate to historical differences between the continents that have resulted in unique body size ranges that are not correlated with habitat (Louys et al., 2011).

Minor ecosystems, particularly forested areas, are the easiest to distinguish from other categories. However, some pairs share similar mammalian community structures. These include MGS–SDS and SDS–TUN. Most of these are open environments that gradually transform into each other across the landscape and are likely to share species that are comfortably adapted to ecotones and a wider array of ecological conditions.

At the detailed level of vegetation category, the general pattern of differences between forested and open areas are still noticeable: the number of frugivorous and arboreal species account for most of the differences. However, the same types of pairs that did not differentiate at the minor ecosystem level can now be distinguished. These include SGW–RGD; SGW–MGS; TMC–TBS; MGS–SDS and SDS–TUN. The division of the minor ecosystem TSM into SGW and TMC reveals the true nature of this group: SGW represents more open and seasonal areas and TMC is more forested. This may be the reason why SGW does not clearly differ from RGD and MGS. Both SGW and RGD have some number of dry months during the year, are widespread and often mixed vegetation complexes (Olson et al., 1983). SGW is a mosaic that contains dry, grassy areas as well as places that are similar to tropical seasonal forests. On the other hand, RGD may also alternate with grassy savannas, so that points of separation become arbitrary (Olson et al., 1983). In the constrained ordination (CAP), RGD plotted near or within the SGW, STW, and MGS, which indicates that seasonality plays a major role for the mammalian composition of this vegetation group. TMC consists of rainy forests that may be locally denser than nearby lowlands and are more similar to woodlands, such as RGD (Olson et al., 1983). Other studies have subsumed montane forests into a general forest category (e.g. Reed, 1998). Although Olson classifies TMC as a part of interrupted woods (due to lower mean carbon density), our results also suggest that the community structure of TMC is similar to forest complexes and that it is not appropriate to consider this habitat a part of IW.

4.3. Palaeoecological implications

Our study confirms that palaeoecological reconstructions of tropical fossil localities based on dietary and locomotor variables can be obtained. Questions raised regarding the appropriateness of modern comparative datasets derived from localities across multiple continents are possibly unwarranted, unless body size has been given significant consideration. This is not to say that ecological differences between the continents do not exist in the structure of communities. We showed some body mass differences with results similar to others (e.g. Smith et al., 2004; Louys et al., 2011), particularly in demonstrating that Africa has a higher number of large-bodied taxa; our tentative suggestion is that the differences are perhaps more meaningful than other workers who emphasise that body size distributions across continents are not significant (e.g. Smith et al., 2004). For example, during the drier late Pleistocene last glacial maximum, the body mass groups of mammalian faunas in South America and Africa were quite similar. However, during the wetter Holocene climatic optimum there was a decrease in open vegetation areas that lead to the loss of larger-sized mammal lineages in South America, whilst they survived in Africa. As a result, today Africa is richer in the number of species above 5 kg (De Vivo and Carmignotto, 2004). In addition, Louys et al. (2011) hypothesise that faunal exchange over the past several million years between Africa and Asia contributes to community convergence, whereas the much longer separation of the New World from the Old World relates to differences between their communities. Furthermore, they posit that historical factors function on a much longer temporal scale in shaping community structure, which suggests that some caution is warranted in selecting comparative communities from across the world for palaeoecological reconstructions of sites located on a continent that has been long separated from the others, or for analyses involving deep time. This can be interpreted from our results, too. Although we did not analyse continental differences in the other ecovariables, our results show only relatively small numbers of similar mammalian community structures within worldwide tropical vegetation classes. This indicates that historical and ecological factors are shaping mammalian community structure slightly differently in relatively similar primary production environments (Endler, 1982).

In addition to caveats regarding the usefulness of body mass as a habitat discriminator, our results provide an additional principal observation for palaeoecological community studies. Some vegetation groups represent a continuum of cover (TMC–TBS–TRF, RGD–SGW–MGS–SDS) that are, at least at the worldwide comparison, too detailed to differentiate when broad generalities are sought. This may be related to the methodology, in which habitats on this continuum are forced into discrete categories for the sake of analysis. The ecological reality is that differences between these groups are probably much more subtle than these discrete categories can be analysed using our ordination-based approaches and other common tests. One way around this problem may be to focus on reconstructing not habitat categories specifically, but the variable amount of canopy cover available in them, as Louys et al. (2015) have done. However, one may seek more ecological detail than this, whereby the use of well-defined habitat categories is both convenient and analytically manageable. Where this option is pursued, we suggest using three worldwide tropical major ecosystems: “Humid, closed forests” that contains TMC, TBS, and TRF; “Seasonal or interrupted forests and grasslands” that contains RGD, SGW, and STW; and “Seasonal, open drylands” that contains MGS, SDS, and TUN (TBC and SWP were difficult to differentiate from the other groups so they are excluded from our three groups; the former exists mostly in mid-latitudes and is a marginal habitat in the Tropics, whilst the latter is a generalised category that can be found naturally within many other habitats). These broad groups are similar in structure to those described by Mendoza et al. (2004): “evergreen forests”, “wooded savannahs that share grass and browse” and “arid communities with sparse plant cover”. They identified these in strictly African communities, but they appear to hold at the global level, as well, further highlighting their utility in ecological reconstructions. Interestingly, these authors do not consider locomotion in their study, but assign each large mammal taxon to a group which combines feeding strategy and body size.

The data as we have analysed them herein demonstrate their wide applicability in distinguishing between modern habitats and, by extension, fossil localities where relatively complete species lists can be obtained. We appreciate that defining palaeospecies is not a straightforward task, but since we find that more precise habitat reconstructions are every bit as possible as broad ecosystem discrimination where increasingly detailed species information is analysed, palaeontological research should continue to approach questions of individual species adaptive complexes, as well as palaeocommunity structures and habitat affiliations.

Acknowledgements

We wish to thank Mikael Fortelius for suggesting this study and Peter Andrews whose earlier studies inspired it. We thank Peter Andrews, John Damuth, and Kaye Reed for assistance with mammalian datasets. We are grateful to Lee Hsiang Liow, Mikael Fortelius, Kaye Reed, Russ Hill, Jussi Eronen, Julien Louys and an anonymous reviewer for useful comments. We thank Eduard Szöcs (Institute for Environmental Sciences, University of Koblenz-Landau) who provided the SIMPER R code and Joonas Lintulaakso who helped us with finalising the CAP figure. We are grateful for Sara Bolton's contributions in testing dataset compatibility during her MSc at Durham University. Funding for KL came from the Emil Aaltonen Foundation.

Map. KML file containing Google map of the most important areas described in this article.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.palaeo.2016.04.012. These data include the Google map of the most important areas described in this article.

References

Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Springer, Netherlands.
Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46.
Anderson, M.J., 2005. PERMANOVA. Permutational Multivariate Analysis of Variance. Department of Statistics, University of Auckland, Auckland.
Anderson, M.J., Willis, T.J., 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84, 511–525.
Anderson, M.J., Ellingsen, K.E., McArdle, B.H., 2006. Multivariate dispersion as a measure of beta diversity. Ecol. Lett. 9, 683–693.
Andrews, P., Hixson, S., 2014. Taxon-free methods of palaeoecology. Ann. Zool. Fenn. 51, 269–284.
Andrews, P., Humphrey, L., 1999. African Miocene environments and the transition to early hominines. In: Bromage, T.G., Schrenk, F. (Eds.), African Biogeography, Climate Change, and Human Evolution. Oxford University Press, New York, pp. 282–300.
Andrews, P., Lord, J.M., Evans, E.M.N., 1979. Patterns of ecological diversity in fossil and modern mammalian faunas. Biol. J. Linn. Soc. 11, 177–205.
Badgley, C., Damuth, J., Theodor, J., Werdelin, L., 2001. Global variation in ecological structure of mammalian faunas and paleoenvironmental inferences. J. Vertebr. Paleontol. 21 (suppl.), 30A.
Bodmer, R.E., 1990. Ungulate frugivores and the browser–grazer continuum. Oikos 57, 319–325.
Bray, J.R., Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27, 325–349.
Brown, J.H., Maurer, B.A., 1989. Macroecology: the division of food and space among species on continents. Science 243, 1145–1150.
Butler, J., 1981. Effect of various transformations on the analysis of percentage data. J. Int. Assoc. Math. Geol. 13, 53–68.
Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. Austral Ecol. 18, 117–143.
Clarke, K., Warwick, R., 2001. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. second ed. Primer-E Ltd., Plymouth, UK.
Damuth, J., Fortelius, M., Andrews, P., Badgley, C., Hadly, E.A., Hixon, S., Janis, C., Madden, R.H., Reed, K., Smith, F.A., Theodor, J., Van Dam, J.A., Van Valkenburgh, B., Werdelin, L., 2002. Reconstructing mean annual precipitation based on mammalian dental morphology and local species richness. J. Vertebr. Paleontol. 22 (suppl.), 48A.
De Vivo, M., Carmignotto, A.P., 2004. Holocene vegetation change and the mammal faunas of South America and Africa. J. Biogeogr. 31, 943–957.
R Development Core Team, 2005. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.
Dubost, G., 1984. Comparison of the diets of frugivorous forest ruminants of Gabon. J. Mammal. 65, 298–316.
Eisenberg, J.F., 1981. The Mammalian Radiations: An Analysis of Trends in Evolution, Adaptation, and Behavior. University of Chicago Press.
Eiten, G., 1992. How names are used for vegetation. J. Veg. Sci. 3, 419–424.
Endler, J.A., 1982. Problems in distinguishing historical from ecological factors in biogeography. Am. Zool. 22, 441–452.
Fa, J.E., Purvis, A., 1997. Body size, diet and population density in Afrotropical forest mammals: a comparison with neotropical species. J. Anim. Ecol. 66, 98–112.
Fleagle, J.G., Reed, K.E., 1996. Comparing primate communities: a multivariate approach. J. Hum. Evol. 30, 489–510.
Fleming, T.H., 1973. Numbers of mammal species in North and Central American forest communities. Ecology 54, 555–563.
Fortelius, M., 2013. The grassiness of all flesh. J. Biogeogr. 40, 1213–1214.
Grayson, D.K., 1984. Quantitative Zooarchaeology: Topics in the Analysis of Archaeological Faunas. Academic Press. Harcourt Brace Jovanovich.
Harrison, J.L., 1962. The distribution of feeding habits among animals in a tropical rain forest. J. Anim. Ecol. 31, 53–63.
Hernández Fernández, M., Vrba, E.S., 2006. Plio-Pleistocene climatic change in the Turkana Basin (East Africa): evidence from large mammal faunas. J. Hum. Evol. 50, 595–626.
IUCN, UNEP, 2009. The World Database on Protected Areas (WDPA). UNEP-WCMC, Cambridge, UK (Available from http://www.wdpa.org).
Jackson, D.A., 1997. Compositional data in community ecology: the paradigm or peril of proportions? Ecology 78, 929–940.
Janis, C.M., 1993. Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. Annu. Rev. Ecol. Syst. 24, 467–500.
Janis, C.M., Damuth, J., Theodor, J.M., 2000. Miocene ungulates and terrestrial primary productivity: where have all the browsers gone? Proc. Natl. Acad. Sci. U. S. A. 97, 7899–7904.
Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., Safi, K., Sechrest, W., Boakes, E.H., Carbone, C., 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 90, 2648.
Kelt, D.A., Brown, J.H., Heske, E.J., Marquet, P.A., Morton, S.R., Reid, J.R.W., Rogovin, K.A., Shenbrot, G., 1996. Community structure of desert small mammals: comparisons across four continents. Ecology 77, 746–761.
Kovarovic, K., Andrews, P., Aiello, L., 2002. The palaeoecology of the Upper Ndolanya Beds at Laetoli, Tanzania. J. Hum. Evol. 43, 395–418.
Lawesson, J.E., 1994. Some comments on the classification of African vegetation. J. Veg. Sci. 5, 441–444.
Leemans, R., Cramer, W., Van Minnen, J.G., 1996. Prediction of global biome distribution using bioclimatic equilibrium models. In: Breymeyer, A.I., Hall, D.O., Melillo, J.M., Ågren, G.I. (Eds.), Global Change: Effects on Coniferous Forests and Grasslands. SCOPE. John Wiley and Sons, New York, pp. 414–450.

Legendre, P., Legendre, L., 1998. Numerical Ecology. 2nd English ed. Elsevier Science BV, Amsterdam, the Netherlands.
Lintulaakso, K., 2013. MammalBase — database of recent mammals. http://www.mammalbase.net.
Liow, L.H., Fortelius, M., Bingham, E., Lintulaakso, K., Mannila, H., Flynn, L., Stenseth, N.C., 2008. Higher origination and extinction rates in larger mammals. Proc. Natl. Acad. Sci. U. S. A. 105, 6097–6102.
Liow, L.H., Fortelius, M., Lintulaakso, K., Mannila, H., Stenseth, N.C., 2009. Lower extinction risk in sleep-or-hide mammals. Am. Nat. 173, 264–272.
Louys, J., 2007. Ecology and Extinction of Southeast Asia's Megafauna (PhD Thesis) University of New South Wales Sydney, Australia.
Louys, J., Travouillon, K.J., Bassarova, M., Tong, H., 2009. The use of protected natural areas in palaeoecological analyses: assumptions, limitations and application. J. Archaeol. Sci. 36, 2274–2288.
Louys, J., Meloro, C., Elton, S., Ditchfield, P., Bishop, L.C., 2011. Mammal community structure correlates with arboreal heterogeneity in faunally and geographically diverse habitats: implications for community convergence. Glob. Ecol. Biogeogr. 20, 717–729.
Louys, J., Meloro, C., Elton, S., Ditchfield, P., Bishop, L.C., 2015. Analytical framework for reconstructing heterogeneous environmental variables from mammal community structure. J. Hum. Evol. 78, 1–11.
Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer research 27.2 Part 1, 209–220.
Mendoza, M., Goodwin, B., Criado, C., 2004. Emergence of community structure in terrestrial mammal-dominated ecosystems. J. Theor. Biol. 230, 203–214.
Mendoza, M., Janis, C.M., Palmqvist, P., 2005. Ecological patterns in the trophic-size structure of large mammal communities: a 'taxon-free' characterization. Evol. Ecol. Res. 7, 505–530.
Miljutin, A., 2009. Substrate utilization and feeding strategies of mammals: description and classification. Est. J. Ecol. 58, 60–71.
Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M., 2007. The Vegan Package (Community ecology package).
Olson, E.C., 1966. Community evolution and the origin of mammals. Ecology 47, 291–302.
Olson, J.S., Watts, J.A., Allison, L.J., 1983. Carbon in Live Vegetation of Major World Ecosystems. Oak Ridge National Laboratory, Oak Ridge TN.
Olson, J.S., Watts, J.A., Allison, L.J., 1985. Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation (NDP-017).
Owen-Smith, N., 2013. Contrasts in the large herbivore faunas of the southern continents in the late Pleistocene and the ecological implications for human origins. J. Biogeogr. 40, 1215–1224.
Peters, R.H., Raelson, J.V., 1984. Relations between individual size and mammalian population density. Am. Nat. 124, 498–517.
Reed, K.E., 1997. Early hominid evolution and ecological change through the African Plio-Pleistocene. J. Hum. Evol. 32, 289–322.
Reed, K.E., 1998. Using large mammal communities to examine ecological and taxonomic structure and predict vegetation in extant and extinct assemblages. Paleobiology 24, 384–408.
Reed, K.E., 2008. Paleoecological patterns at the Hadar hominin site, Afar Regional State, Ethiopia. J. Hum. Evol. 54, 743–768.
Rodríguez, J., Hortal, J., Nieto, M., 2006a. An evaluation of the influence of environment and biogeography on community structure: the case of Holarctic mammals. J. Biogeogr. 33, 291–303.
Rodríguez, M.Á., López-Sañudo, I.L., Hawkins, B.A., 2006b. The geographic distribution of mammal body size in Europe. Glob. Ecol. Biogeogr. 15, 173–181.
Simpson, G.G., 1943. Mammals and the nature of continents. Am. J. Sci. 241, 1–31.
Smith, F.A., Brown, J.H., Haskell, J.P., Lyons, S.K., Alroy, J., Charnov, E.L., Dayan, T., Enquist, B.J., Morgan Ernest, S.K., Hadly, E.A., Jones, K.E., Kauman, D.M., Marquet, P.A., Maurer, B.A., Niklas, K.J., Porter, W/.P., Tiffney, B., Willig, M.R., 2004. Similarity of mammalian body size across the taxonomic hierarchy and across space and time. Am. Nat. 163, 672–691.
Sponheimer, M., Reed, K.E., Lee-Thorp, J.A., 1999. Combining isotopic and ecomorphological data to refine bovid paleodietary reconstruction: a case study from the Makapansgat limeworks hominin locality. J. Hum. Evol. 36, 705–718.
van Dam, J.A., Andrews, P., Badgley, C., Damuth, J., Fortelius, M., Hadly, E.A., Hixson, S., Janis, C., Madden, R.H., Reed, K., Smith, F.A., Theodor, J., Van Valkenburgh, B., Werdelin, L., 2001. Within-habitat mammal diversity and productivity and their recent patterns across latitude. J. Vertebr. Paleontol. 21 (suppl), 43A.
White, F., UNESCO, 1983. The Vegetation of Africa: A Descriptive Memoir to Accompany the UNESCO/AETFAT/UNSO Vegetation Map of Africa. UNESCO.
Wilson, D.E., Reeder, D.A.M., 2005. Mammal Species of the World: A Taxonomic and Geographic Reference. Johns Hopkins University Press.

Kari Lintulaakso is a PhD student at the University of Helsinki. He specialises in recent mammals and his main interest is linking current mammalian community structures with key environmental factors that can be used in palaeoclimatological and environmental studies.

Kris Kovarovic is a Senior Lecturer in Human Evolution at Durham University, UK. She specialises in mammalian community palaeoecology, ungulate ecomorphology and the use of fossil mammals in reconstructing the environment of hominin evolution.

Paper IV

Lintulaakso, K., Polly, P.D., Eronen, J.T.
Land mammals form eight distinct biogeographical clusters
in North America but only one in Europe.
Journal of Biogeography (submitted), 2018.

Land mammals form eight functionally and climatically distinct faunas in North America but only one in Europe

Kari Lintulaakso[1,5], P. David Polly [2], Jussi T. Eronen[1,3,4]

[1] *Department of Geosciences and Geography, University of Helsinki, Finland*
[2] *Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN 47405, USA*
[3] *BIOS Research Unit, Meritullintori 6, Helsinki, Finland*
[4] present address: Helsinki Institute of Sustainability Science, Faculty of Biological and Environmental Sciences, Ecosystems and Environment Research Programme, University of Helsinki, P.O. Box 65 (Viikinkaari 1), 00014 University of Helsinki, Finland
[5] present address: Finnish Museum of Natural History LUOMUS, P.O. Box 44 (Kumpula Manor, Jyrängöntie 2), 00014 University of Helsinki, Finland

Abstract

Aim

We use cluster analysis and spatial statistics to delimit mammalian faunal clusters that are both climatically and functionally distinct. These, we argue, are the entities that form regional species pools and are relevant to community assembly processes. Such faunal clusters can be differentiated in the fossil record on the basis of occurrences and traits, which also makes them relevant for inferring differences in environmental conditions in the past.

Location

North America within W 178°, W 14°, N 83°, N 7° and Europe within W 32°, E 35°, N 80°, N 35°

Major taxa studied

575 and 124 land mammal species from North America and Europe

Methods

Point-sampled species occurrences of mammals from North America and Europe were each clustered with k-means to identify faunas at spatial scales ranging from three clusters at the largest scale and 21 at the smallest. At each scale, the faunas were tested for significant differences in climate (mean annual precipitation, mean annual temperature) and functional traits (body mass, locomotion and diet).

Results

In North America, climatic differentiation exists at the scale where mammals are divided into 11 distinct faunas and, in Europe, at the scale where there are five faunas. Functional trait differentiation in body mass occurs at a larger spatial scale in North America (8 distinct faunas), but locomotor differentiation is present at all spatial scales, and dietary differentiation is not present at any scale. No significant differentiation in any functional trait at any scale is present in Europe.

Main conclusions

Faunal clusters can be constructed at any spatial scale, but clusters are climatically and functionally meaningful only at larger scales. Climatic differences (and linked environmental differences) and their associated functional trait specialisations are likely to be barriers to large-scale mixing. We argue, therefore, that faunal clusters that are climatically and functionally distinct are the entities that form regional species pools for community assembly processes. In North America, there are eight such mammal species pools, but only one in Europe. Since the functional traits in our study are observable in the fossil record, community-level functional trait analysis has the potential to be used to reconstruct past climatic gradients.

Keywords: climate variables, clustering, communities, Europe, functional traits, mammals, North America, species pools

1 Introduction

An important problem in biogeography is the relationship between species pools, communities, climate and environments, and functional traits (e.g., Fox and Brown, 1993; McGill et al., 2006). The interaction between functional traits and the environment has been well studied in the context of local community assembly, but their interactions in the formation of larger species pools have been much less studied (but see Zobel, 1999; Zobel *et al.*, 1998 for examples). A species pool is the larger group of species from which local communities are assembled (Weiher and Keddy, 2001). Those species by definition must be functionally compatible with the local and regional environments, even if they are not all mutually compatible in a local community. Therefore, a good working definition of a species pool is that it is distinctive in both its species composition and in those functional traits that are related to the shared regional environment. Because species pools serve as the sources of species for community assembly processes (Weiher and Keddy, 2001), not to mention the null context for the statistical evaluation of community assembly problems (e.g., Connor and Simberloff, 1979; Gotelli, 2000; Kraft and Ackerly, 2010), it is important to understand their geographic extent and their relationship to broader patterns of climate, environment, and functional traits.

Our primary purpose is to identify faunas that are both compositionally and functionally different in ways that correspond to environmental (specifically climatic) differences. We use geographic ranges, trait data, and climate data with clustering algorithms and Monte Carlo statistics to identify diagnosable faunas of North American mammals based on species composition and to determine at what spatial scale faunas are differentiated from one another in terms of their climate and their functional traits. Functional traits are the mechanisms by which species interact with habitat variation, so we would expect differences in functionally relevant features like body size, locomotion, and dietary preferences from one fauna to another unless the sorting of species into regional pools is purely by chance.

Our secondary purpose is to test Heikinheimo et al.'s (2007) findings on European mammal faunas that showed that clusters of mammal species correspond closely with geographic barriers like rivers and mountains, paralleling the boundaries of modern nation states (presumably for parallel reasons). Those authors used gridded presence/absence records of European land mammals for clustering the faunas. The two independent clustering methods they used produced spatially coherent faunal clusters that were interpreted to represent larger metacommunities whose boundaries were influenced by the interaction of natural barriers (rivers and mountains) and climatic gradients. Similar community cluster analyses have been performed for North America and Africa (Polly 2010; Linder et al. 2012). Majority of these studies have been done for a single continent only. The clusters in Heikinheimo et al. (2007) were shown to be strongly related to climate and were similar to an independently derived environmental zonation of Europe (Metzger et al., 2005). Clusters were geographically alike for multiple categorisations of mammalian data, such as trophic groups, small or large mammals or the extinction risk status of the species. The results were especially noteworthy, as the clustering methods used did not take geography (spatial adjacency of grid cells) into account in any way yet produced spatially coherent clusters. They interpreted the community clusters as biologically distinct, metacommunity entities (species pools). The clusters differed significantly ($p < 0.05$) in pairwise ANOVA comparisons of their primary climate variables (precipitation, temperature, and annual temperature range) and elevation. The authors concluded that the community clusters, which were based only on species co-occurrences, were associated with major physiographic and environmental features. Heikinheimo et al. (2012) later combined climate, plant and mammal datasets in community clustering to show that mammal and plant clusterings are closely linked in spatial extent. Coherent floristic groups (biomes) are usually associated with climate (temperature, temperature range, and rainfall), mediated by elevation differences (Holdridge, 1967; Whittaker, 1975), thus driving the association of mammal faunas with these same factors.

Finally, functionally differentiated species pools provide an important bridge to the study of paleoecology and paleoenvironment. Local community composition at fossil sites is often poorly understood because of preservational biases, but regional, which are arguably equivalent to species pools in both space and time (*c.f.*, Olson 1952), are more robustly defined because they are derived from the combined occurrences at many sites (Eronen et al., 2009; Figueirido et al., 2012). As we show below, functional differentiation between faunas manifests itself at a larger spatial scale than climatic differentiation (and thus environmental differentiation). The existence of functionally differentiated faunal clusters therefore implies environmentally (climatically) differentiated regions, thus linking biogeographic analysis of the fossil record with paleoenvironmental reconstruction.

2 Materials and methods

Geographic ranges for mammals in North America were taken from *Digital Distribution Maps of the Mammals of the Western Hemisphere, 3.0* (Patterson et al., 2003), which is an update of Hall (1981), and in Europe from the *Atlas of European Mammals* (Mitchell-Jones et al., 1999). Because the focus of this study is on terrestrial mammals, bats and aquatic species were excluded. Non-native species were also excluded, except for the raccoon dog (*Nyctereutes*), which has become established in Europe through recent range expansion from its native eastern Asia. Because of their commensalism with human habitation, rodents *Mus* and *Rattus* were also excluded. A total of 575 and 124 species were included for the two continents respectively.

To facilitate clustering and other spatial analyses, the ranges of each species were point sampled using a grid of equidistant points spaced 50 km apart (Polly, 2010). This strategy avoids the common problems of latitudinal biases in sampling density founded in data that are gridded by longitudinal degrees and the problems of spatial scaling associated with amalgamating data contained within grid cells (Polly, 2010; Polly and Sarwar, 2014; Lawing et al., 2016).

Species occurrences, climate variables (Willmott and Legates, 1988), biomes (Bailey & Hogg, 1986; Bailey, 1989), and elevation (Hastings and Dunbar, 1998) were sampled using this grid. Bailey's ecoregions, which are spatially localised areas with common temperature, precipitation, and vegetation, were used as our biome categories. The ecoregions are classified hierarchically into domains, divisions and provinces. North America has four domains and 28 divisions, and Europe has three domains and 15 divisions. To assess the association between faunal clusters and biomes, the biomes occurring within each core cluster were tabulated as percentages of the total number of grid points in that cluster.

For each species, functional traits of log body mass, locomotion, and diet were compiled. Data were from the National Center for Ecological Analysis and Synthesis (NCEAS) Workshop on Mammalian Communities (see Badgley et al., 2001; van Dam et al., 2001; Damuth et al., 2002), PanTHERIA (Jones et al., 2009). We also used MammalBase, a compilation of species attributes and diets based on hundreds of published sources (Lintulaakso, 2013). Body mass was quantified as the natural log of median body mass (in grams) for each species from the data sources listed in Appendix S1 in the Supporting Information. These were categorised into 1.0 log unit bins ranging between 0 and 14. Locomotion was classified into six substrate-based categories: arboreal (e.g., opossums and two-toed sloths); arboreal-terrestrial (e.g., raccoons and grey squirrels), subterranean (e.g., pocket gophers and moles), subterranean-terrestrial (*e.g.,* ground squirrels and deer mice), terrestrial (e.g., cotton-tailed rabbits and deer), and terrestrial-aquatic (e.g., beaver and otters) (Reed, 1998; Miljutin, 2009). If published sources disagreed, the most commonly reported substrate was used (Appendix S1). Dietary categories were assigned by K.L. into three broad categories, animalivorous (a combined category for carnivores and insectivores), frugivorous and herbivorous, which were subdivided into a total of 28 categories based on listed food resources (Appendix S1).

Matrices of species distribution for both continents were created in which each row represented a 50 km grid point and each column a species with 0 or 1 for absence or presence. K-means clustering (Steinhaus, 1956) was applied to partition the occurrences into k number of clusters. This method chooses *k* random centroids; then each observation is assigned to the closest centroid using a distance metric (Euclidean in this study). Next, a new centroid is chosen for each cluster based on the same distance metric, after which each observation is assigned again. The procedure is repeated until none of the observations changes clusters or an agreed limit for iterations is achieved (see Heikinheimo et al., 2007; Polly 2010). *K*-means clustering can find different solutions in successive runs of the same data, so we adopted a "*core clusters*" strategy in which points that were not consistently assigned to the same cluster in 10 randomized clustering iterations were excluded, thus leaving a core cluster of points that is robust to randomly inconsistent behavior of the clustering algorithm. The whole core clustering procedure was repeated for k-values between 3 and 21.

Summary statistics for each faunal cluster were calculated, including the total number of species (standing diversity), the number of endemic species (species not found in any other faunal cluster), and ubiquitous species (species that are found in every grid point of the cluster). Endemic and ubiquitous species define the coherency of the fauna such that it can be diagnosed in the real world.

We used climate and functional traits to determine at which value of *k* faunas are meaningfully differentiated. We defined "climatic units*"* as those faunal clusters with the highest value of *k* for which both climate variables (mean annual precipitation and mean annual temperature) were statistically different. The faunal clusters were plotted in a climate space defined by mean annual temperature and annual precipitation (the same space that Whittaker, 1975 used to categorise vegetative biomes) as double box plots (*sensu* Shinichiro, 2013). For clarity, only the plus and minus one standard deviation were plotted, leaving out the whiskers and outliers, clarifying to assess the significance of the differences between faunal clusters. Overlapping plots were considered as climatically similar. The significance of climatic differences was also tested using a randomisation procedure in which the observed variance of the faunas in the climate space was compared to a null distribution of variances generated by randomising the climate variables with respect to the faunal points 1,000 times.

We identified functionally distinct faunas using a similar approach of identifying the highest value of k for which the species trait composition (average body mass, locomotion, and diet groups) were statistically different. To determine the significance of body mass we used another random permutation test with 1,000 replicates. However, we treated the three variables as frequency distributions instead of means (this is required for the two categorical variables). Observed differences among the faunas were calculated as the variance of their distributions around the overall distribution for the entire continent using a chi-square distance (sum of the squared differences between values in each bin). The observed distance was compared to a null distribution of distances derived from randomising the trait variables with respect to the faunas.

Cenograms, which are rank ordered distributions of body mass in a group of species (Valverde 1964; Legendre 1986), were used to visualise gaps in body mass distributions among the faunal clusters. Cenograms from open environments have a gap in the medium-sized species (500–8000 g), whereas closed environments have a continuous distribution (Legendre 1986). A gap is defined to be at least two-fold difference of the body mass (in g).

3 Results

Number of climatically and functionally distinct faunas

We found eleven faunal units in North America and five in

Table 1. Descriptive statistics of faunal clusters. The predominant Bailey's ecoregion domain and division are indicated of each cluster with the percentage of the area of the cluster that it occupies. *No=cluster number in supplementary material at k=8, 11 (North America) and at k=5 (Europe).*

Name and abbreviation	No	Domain	% Domain	Division	% Division	Temperature (°C) ± SD	Precipitation (mm) ± SD	Elevation (m)
North American functionally distinct clusters (k=8)								
High Arctic Canada (HC)	5	Polar	97	Tundra	78	-11.4 ± 5.2	289 ± 273	270
Eastern Beringia (EB)	8	Polar	92	Subarctic Mountains	66	-5.3 ± 3.1	448 ± 294	837
Northern High Canada (NC)	7	Polar	100	Subarctic	85	-5.2 ± 1.0	560 ± 168	360
Southern Canada (SC)	1	Polar	77	Subarctic	77	0.4 ± 2.4	705 ± 242	370
Great Basin (GB)	3	Dry	98	Temperate Desert	64	6.8 ± 2.7	336 ± 113	1782
Eastern US (EU)	4	Humid Temperate	100	Hot Continental	55	10.9 ± 1.7	1036 ± 133	287
Northern Mexico (NM)	2	Dry	87	Tropical/Subtropical Desert	58	18.2 ± 3.1	441 ± 172	1479
Mesoamerica (MA)	6	Humid Tropical	99	Savanna	31	23.8 ± 3.4	1737 ± 786	620
North America climatically distinct clusters (k=11)								
High Arctic Canada (HC)	11	Polar	98	Tundra	74	-12 ± 6.0	299 ± 191	293
Eastern Beringia (EB)	2	Polar	92	Subarctic Mountains	65	-5.4 ± 3.1	445 ± 296	827
Northern High Canada (NC)	9	Polar	100	Subarctic	97	-3.7 ± 1.1	715 ± 187	348
Southern Canada (SC)	6	Polar	87	Subarctic	87	0.3 ± 1.1	616 ± 185	417
British Columbia (BC)	10	Humid Temperate	55	Warm Continental Mountains	33	1.4 ± 3.0	772 ± 451	1234
Northern Rocky Mountains (NR)	1	Dry	100	Temperate Steppe	90	5.6 ± 1.7	361 ± 47	1022
Great Basin (GB)	8	Dry	98	Temperate Desert	58	6.3 ± 3.0	355 ± 129	1790
Eastern US (EU)	7	Humid Temperate	100	Hot Continental	82	9 ± 2.0	918 ± 136	271
Southeastern US (SU)	5	Humid Temperate	99	Subtropical	69	17 ± 2.5	1294 ± 160	129
Northern Mexico (NM)	3	Dry	87	Tropical/Subtropical Desert	41	17.4 ± 3.2	478 ± 165	1262
Mesoamerica (MA)	4	Humid Tropical	100	Savanna	30	23.9 ± 3.3	1804 ± 787	576
European climatically distinct clusters (k=5)								
Northern Scandinavia – Finland (NS)	3	Polar	54	Subarctic	54	1.1 ± 2.4	700 ± 282	360
Central Europe and The Baltic countries (CE)	1	Humid Temperate	96	Hot Continental	29	8.1 ± 2.0	678 ± 172	284
Southern Scandinavia – UK (SS)	4	Humid Temperate	62	Marine	46	8.2 ± 2.6	837 ± 284	171
France (FR)	2	Humid Temperate	96	Marine	45	9.7 ± 2.5	839 ± 187	463
Mediterranean (ME)	5	Humid Temperate	71	Mediterranean	36	12.4 ± 3.7	726 ± 287	426

Europe that were statistically distinct in climate space (mean annual precipitation and mean annual temperature) (Table 1; Figs 1b, 2). We also identified eight functionally distinct faunas in North America based on trait differences in body mass and locomotion (Table 2; Fig 1a). Diet did not differ between faunas in North America at any spatial scale, nor did any of the functional traits differ among faunas in Europe at any spatial scale. In North America, there was a close correspondence between climatic and functionally distinct faunas (R=0.86 for the number of species that were ubiquitous to both climatic and functional clusters, Appendix S3; and R= 0.99 for mean annual temperature, Table 1).

In the following sections, we report on eight North American climatically and functionally distinct faunas, followed by a report on the three remaining climatic units. Finally, we will summarise the results of the five climatic units for Europe. See Figures 1, 2 and 3, Tables 1 and 2, and Appendix S3 for detailed statistics about the faunas.

Description of climatically and functionally distinct faunas in North America

Eight North American community clusters differed statistically both in their mammalian body mass (in *k*=5 and *k*=8, P=0.04) and locomotion (*k*=4 to 21, P < 0.02) distribution, indicating the smallest functionally distinct faunas at the continental scale (Fig. 1a, Appendix S3, Table 1). Dietary groups did not differentiate the clusters (*k*=3 to 21, P>0.20; Appendix S2). The majority of the faunas were composed of geographically contiguous points. Four functionally distinct faunas had the majority of their grid points within Bailey's *Polar Domain*, one within *Humid Temperate Domain*, two within *Dry Domain*, and one within *Humid Tropical Domain* (Table 1). We have designated a name for each biogeographical unit based on their geography. Starting from the coldest unit to the warmest one, we describe the main findings for each of them. A full summary of results is in Table 1 and 2.

High Arctic Canada (Cluster 5 at *k*=8 and Cluster 11 at *k*=11) is found dominantly in Bailey's Tundra and Tundra Mountains divisions (93% of the unit's area falls within these ecological divisions). Three non-contiguous areas are clustered with this fauna: the Alaska Peninsula (Marine Mountains division), Vancouver Island (Marine Mountains division), and the southern coast of Cuba (Savanna Mountains division). The latter is almost certainly an artefact caused by the absence of species linking Cuba to other faunal clusters, and the other two outliers may be as well. It is the coldest (mean annual temperature -11.4 ± 5.2 °C) and driest (289 ± 273 mm year-^{1}) of the faunal clusters. This cluster also has the largest mammals with a median body mass of 933 g. In the cenogram data (Appendix S5), there are gaps between 30,000-75,000 g, 285-750 g, and 8-18 g. Terrestrial (45%) and subterranean-terrestrial (35%) species are the primary locomotion groups. **High Arctic Canada** has no subterranean species, and the percentage of arboreal-terrestrial species (9%) is the lowest among units (Table 2).

Eastern Beringia (Cluster 8 at *k*=8 and Cluster 2 at *k*=11) is found in the Subarctic and Subarctic Mountains divisions (84% of the total area). This fauna occurs at the highest elevation (837 ± 542 m) of the Polar Domain. This fauna is the only northern one with a body mass gap in the large mammals, 195,000-460,000 g. There is also a gap between 285-750 g. Terrestrial (46%) and subterranean-terrestrial (33%) species are the primary locomotion groups, the terrestrial percentage being highest among the faunas. There are no subterranean species, and the portion of arboreal species (2%) is the lowest among the faunas.

Northern High Canada (Cluster 7 at *k*=8 and Cluster 9 at *k*=11) is found in the Subarctic division (85% of its total area). It has the lowest standing diversity (49 species), none of which are endemic to it. Body mass is also large in this fauna, with a median of 747 g and it has body mass gaps between 30,000-75,000 g and 285-750 g. Terrestrial (45%) and subterranean-terrestrial (31%) species are the major locomotion groups, the subterranean-terrestrial percentage being the lowest among the faunas. There are no subterranean species, and the portion of terrestrial-aquatic species (8%) is the highest among units.

Southern Canada (Cluster 1 at *k*=8 and Cluster 6 at *k*=11) straddles the Polar domain's Subarctic division (77% of its total area) and the Humid Temperate domain's Warm Continental division (16% of its area). Median body mass is 286 g. This fauna is the only one with no gaps in mammalian body masses. Terrestrial (38%) and subterranean-terrestrial (33%) species are the primary locomotion groups.

Great Basin (Cluster 3 at *k*=8 and Cluster 8 at *k*=11) is found in the Temperate Desert and Mountains division of the Dry domain (74% of its total area). It is the second driest fauna (336±113 mm year-^{1}) and has the highest elevation (1782 ± 603 m). This fauna has the second highest number of endemic species (n=56). There are gaps between 110,000-240,000 g and 18,000-47,000 g. Subterranean-terrestrial (51%) and terrestrial (22%) species are the most common locomotor categories in this fauna, the subterranean-terrestrial percentage being the highest and terrestrial percentage being the lowest of any. The percentage of subterranean species (6%) is highest among the faunas.

Eastern US (Cluster 4 at *k*=8 and Cluster 7 at *k*=11) is found in the Hot Continental and Hot Continental Mountains divisions (71% of its total area). It is the only fauna that substantially occupies the Prairie division (22% of its area). There are gaps between 240,000-625,000 g, 110,000-240,000 g and 30,000-75,000 g. Subterranean-terrestrial (35%) and terrestrial (32%) species are the major locomotion groups.

Northern Mexico (Cluster 2 at *k*=8 and Cluster 3 at *k*=11) is found in the Tropical/Subtropical divisions (87 % of its total area). This fauna and **Great Basin** have similar precipitation, elevation, number of species, and a similarly high number of endemic species. However, mean temperature differs significantly (6.8 ± 2.7 °C in **Great Basin** and 18.2 ± 3.1 °C in **Northern Mexico**). There are gaps between 240,000 - 625,000 g, 110,000 - 240,000 g, and 21,000 - 47,000 g. Subterranean-terrestrial (49%) and terrestrial (23%) species are the major locomotion groups. The percentage of terrestrial-aquatic species (2%) is the lowest of any of the faunas.

Table 2. Summary of eight functionally distinct North American faunal clusters. (*No=cluster number in supplementary material at k=8; n*=total number of species; *E*=number of endemic species (species not found in any other faunal cluster); *U*=number of ubiquitous species (species that are found in every grid point of the cluster); A=arboreal; *AT*=arboreal–terrestrial; *S*=subterranean; *ST*=subterranean–terrestrial; *T*= terrestrial; *TA*=terrestrial–aquatic*; SD*=standard deviation).

Cluster		Species			Locomotor groups (%)						Body mass (g)			body mass categories (%)		
Name and abbreviation	**No**	**n**	**E**	**U**	**A**	**AT**	**S**	**ST**	**T**	**TA**	**mean**	**SD**	**median**	**< 500**	**500 – 8000**	**> 8000**
High Arctic Canada (HC)	5	58	8	0	6.9	8.6	0.0	34.5	44.8	5.2	25222	77201	933	43.1	32.8	24.1
Eastern Beringia (EB)	8	57	2	10	1.8	12.3	0.0	33.3	45.6	7.0	24038	68951	286	50.9	22.8	26.3
Northern High Canada (NC)	7	49	0	5	2.0	14.3	0.0	30.6	44.9	8.2	41161	118484	747	49.0	24.5	26.5
Southern Canada (SC)	1	81	1	5	3.7	14.8	4.9	33.3	38.3	4.9	27962	92261	286	53.1	25.9	21.0
Great Basin (GB)	3	150	56	8	2.0	15.3	6.0	51.3	22.0	3.3	12904	67009	183	65.3	25.3	9.3
Eastern US (EU)	4	78	11	11	5.1	15.4	5.1	34.6	32.1	7.7	15355	75921	156	60.3	29.5	10.3
Northern Mexico (NM)	2	140	33	8	2.9	17.9	5.0	49.3	22.9	2.1	11144	58256	122	67.9	22.9	9.3
Mesoamerica (MA)	6	248	175	4	11.7	21.8	4.8	35.5	23.0	3.2	3584	20603	73	72.6	22.6	4.8

Mesoamerica (Cluster 6 at *k*=8 and Cluster 4 at *k*=11) is found in the Humid Tropical domain (99 % of the units grid points). It is the warmest and wettest fauna (23.8±3.4 °C; 1737 ± 786 mm) and has the greatest number of species (248) and endemics (175). Median body mass is smaller than any other fauna (73 g). There is a gap between 84,000 - 295,000 g. Subterranean-terrestrial (36% of the community composition) and terrestrial (23%) species are the most common locomotor types, and arboreal species are more common than in any other fauna (22%).

Description of climatically distinct faunas that are not functionally distinct

British Columbia (Cluster 10 at *k*=11) has 95% of its area spread over four mountain divisions: Subarctic, Marine, Warm Continental, and Temperate Steppe Mountains. Mean annual temperature is 1.4±3.0 °C, and annual precipitation is 772±451 mm per year.

Northern Rocky Mountains (Cluster 1 at *k*=11) is located in the Temperate Steppe division (90% of its total area), has a mean annual temperature of 5.6±1.7 °C, and an annual precipitation of 361±47 mm per year.

Southeastern US (Cluster 5 at *k*=11) is located in the Subtropical division (70% of its total area), has a mean annual temperature of 17.0±2.5 °C, and an annual precipitation of 1294±160 mm per year.

Description of European climatically distinct clusters

Northern Scandinavia – Finland (Cluster 3 at *k*=5), is the only climatically distinct fauna found primarily in the Polar domain (Subarctic division, 54% of the total area). It is the coldest European fauna (1.1±2.4 °C, Table 1) and has the fewest species (62, Appendix S4).

The remaining climatically distinct European faunas belong to the Humid Temperate domain. Three of these form a stepwise temperature-precipitation continuum: **Central Europe and The Baltic countries** (Cluster 1 at *k*=5) has similar mean annual temperature as **Southern Scandinavia – UK** (Cluster 4 at *k*=5; 8.1 ± 2.0 °C, 8.2 ± 2.6 °C, respectively). However, **Central Europe and The Baltic countries** have a lower mean annual precipitation than the **Southern Scandinavia – UK** (678 ± 172 mm year-^{1}, 837 ± 284 mm year-^{1}, respectively). **France** (Cluster 2 at *k*=5) has mean annual precipitation similar to the **Southern Scandinavia – UK** (839 ± 187 mm year-^{1}) but a higher mean annual temperature (9.7 ± 2.5 °C). These three climatic units have quite similar numbers of species (between 83 to 87, Appendix S4) with few endemics (0 to 2, Appendix S4). **Southern Scandinavia – UK** and **France** are found in the Marine division (46% and 45% of their total areas), while the **Central Europe and The Baltic countries** are found in the Hot and Warm Continental division (29% of the total area).
Mediterranean (Cluster 5 at *k*=5) belongs to the Mediterranean domain. It has the highest mean temperature of 12.4 ± 3.7 °C and the highest number of species (111), of which ten are endemic. This unit also includes Ireland, which in Bailey's system belongs to Marine division. The European areas which were never clustered (k=3 to 21) were Iceland and the Faroe Islands (*Icecap* and *Tundra Divisions* in Bailey's system, Appendix S4).

4 Discussion

Why are European faunas not differentiated by functional traits?

One of the most intriguing results is the lack of trait differentiation among European faunas. European faunas are climatically differentiated at a similar spatial scale as North America. North America has more climatically differentiated faunas (*k*=11) than Europe (*k*=5), but that can be explained by the difference in continental size because the average area of the faunas is, in fact, statistically equal (ANOVA F(1,14)=0.51, p < 0.49). Therefore, one might expect as much trait differentiation in Europe as in North America albeit spread over fewer clusters.

Europe's narrower climate range probably explains why its faunas do not show statistically significant differentiation in body mass. Despite climatic differentiation at similar

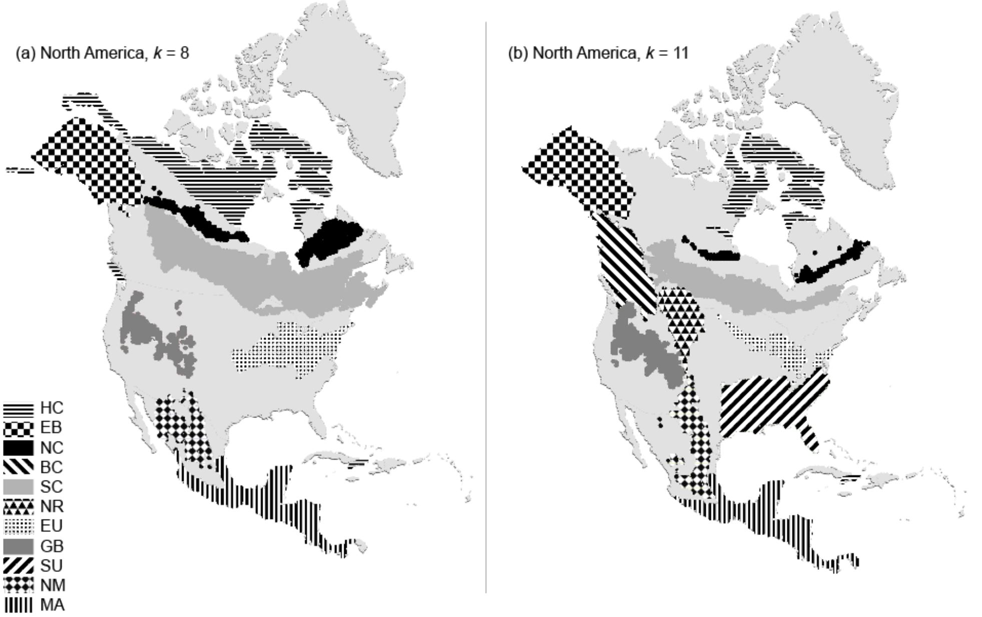


Figure 1. North American mammal community clusters based on k-means clustered species. (a) Functionally distinct faunas defined by North American community clusters at *k*=8. Each fauna differs statistically in body mass, locomotion, and climate (HC, High Arctic Canada; EB, Eastern Beringia; NC, Northern High Canada; SC, Southern Canada; EU, Eastern US; GB, Great Basin; NM, Northern Mexico; MA, Mesoamerica.). (b) Climatic units defined by North American community clusters at *k*=11. Each unit differs statistically by mean annual precipitation and mean annual temperature. (HC, High Arctic Canada; EB, Eastern Beringia; NC, Northern High Canada; BC, British Columbia; SC, Southern Canada; NR, Northern Rocky Mountains; EU, Eastern US; GB, Great Basin; SU, Southeastern US; NM, Northern Mexico; MA, Mesoamerica).

spatial scales, North American faunas are dispersed more widely in climate space (Fig. 3). North America has a wider range of mean annual temperature (-26°C to 29°C) and annual precipitation (54mm to 4860mm) and fills a larger climate space than Europe (-9.7°C to 18.2°C, 242mm to 2331mm). The only three North American faunas that overlap climatically with European ones are the **Eastern US,** which overlaps with **France** and **Southern Scandinavia–UK**, and **Southern Canada** and **British Columbia**, which overlaps with **Northern Scandinavia–Finland.** The remaining eight North American faunas lie outside the climatic boundaries of the European faunas, forming three unique groups: warm and moist, dry, and cold. The greater climate gradient of North America may be necessary to sort species into faunas based on body mass. Added to the reduced climate range, the smaller number of species in Europe reduces statistical power to detect body mass differences.

Similarly, a narrower range of habitats in Europe probably does not facilitate locomotor sorting among the faunas. The range of environments in North America is greater, with 28 ecoregions compared to only 15 in Europe (Bailey and Hogg, 1986). Tropical, desert, and basin and range environments are missing entirely from Europe. The dispersion of North American faunas in climate space is also greater in North America, suggesting that even though the climatic differences in the European faunas are statistically significant, they are smaller (Fig. 3). Therefore, the strong sorting effects imposed by extremely different North American habitats like dense tropical forests, open desert and scrubland, grasslands, and large expanses of taiga may be absent in Europe, thus explaining the lack of locomotor diversification among faunas in the latter.

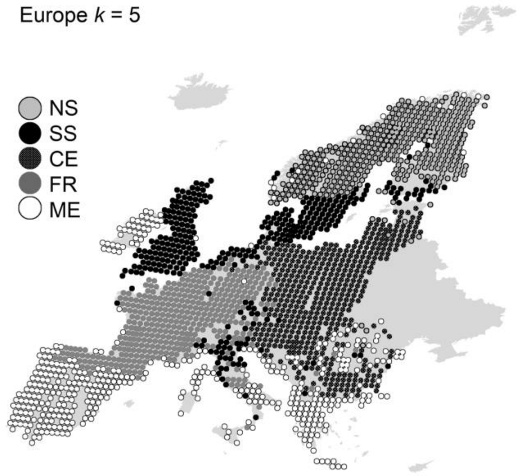


Figure 2. Climatically distinct European faunal clusters at *k*=5. Each unit differs statistically by mean annual precipitation and mean annual temperature. (NS, Northern Scandinavia – Finland; SS, Southern Scandinavia – UK; CE, Central Europe and The Baltic countries; FR, France; ME, Mediterranean).

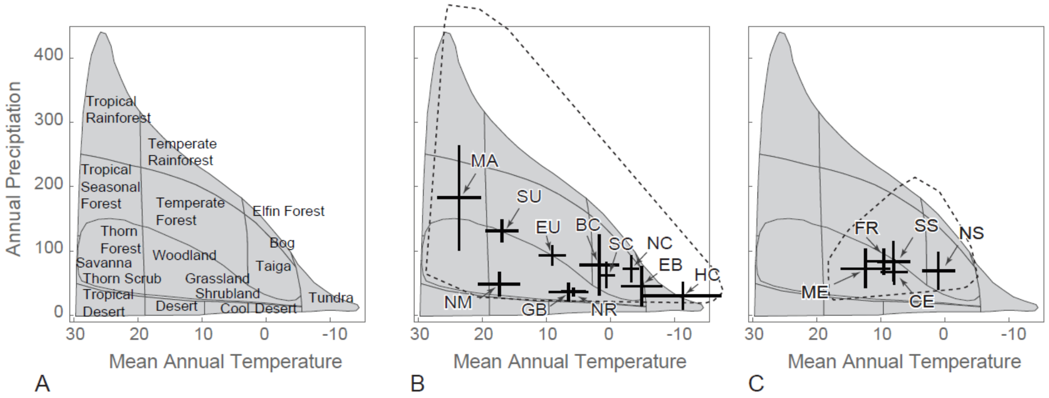


Figure 3. The faunal clusters plotted in a climate space defined by mean annual temperature and annual precipitation. (a) The climatic space that Whittaker, 1975 used to categorise vegetative biomes. North American (b) and European (c) faunal clusters shown with double box plots, corresponding to plus and minus one standard deviation of precipitation and temperature values of the grid points. The dashed areas provide a layer of continental climate ranges, the "climatic spaces", in which all the clusters grid points belong. (BC, British Columbia; CE, Central Europe and The Baltic countries; EB, Eastern Beringia; EU, Eastern US; FR, France; GB, Great Basin; HC, High Arctic Canada; MA, Mesoamerica; ME, Mediterranean; NC, Northern High Canada; NM, Northern Mexico; NR, Northern Rocky Mountains; NS, Northern Scandinavia – Finland; SC, Southern Canada; SS, Southern Scandinavia – UK; SU, Southeastern US).

Another reason for lack of trait differentiation among European faunas could be the long-term impact of humans, who have occupied Europe for more than 780 thousand years (e.g., Ashton et al., 2014), compared to North America, where human occupation is probably less than 25 thousand years (e.g., Bourgeon et al., 2017). Human impacts of hunting and landscape change could well affect trait composition, as has been shown to be the case for locomotor traits in carnivores (Polly and Head, 2015). To investigate this further is beyond the scope of this paper, however.

Heikinheimo et al. (2007 and 2012) argued that major physiographic features, such as rivers and mountains, defined the faunal clusters that they identified. However, their clusters were on a small spatial scale (*k*=12) than the climatically distinct ones that we recovered (*k*=5). That scale transposed into North America would be approximately *k*=21, which would be consistent with physical barriers of the same type in North America (c.f., the 28 ecoregions in North America).

Regional species pools and the hierarchy of faunal sorting in North America

Regional species pools are groups of species that inhabit large areas that have similar climate and physiography and which are capable of coexisting in local communities (Zobel, 1999). To form a regional pool, species must cohabit the same region, be able to reach local habitats, and have traits that allow them to coexist within the physical and biotic context of the local community (Zobel et al., 1998; Zobel, 1999). For this reason, we examined faunal clusters in the context of climate, ecoregion (which is a type of physiographic categorisation), and functional traits that are relevant to climate and physiography, namely body mass and locomotor category. Interestingly, these factors are differentiated at different spatial scales that form a hierarchy (Fig. 4).

Locomotor categories differ at comparatively small spatial scales in North America, similar to the physiographic scale of ecoregions. In fact, locomotor differences appear to form a hierarchy that creates significant differences between faunas at small scales from *k*=21 to large scales at *k*=4. Bailey's ecoregion hierarchy is based on a combination of physiography, vegetation cover, and climate. At the lowest division scale, North America is divided into 60 local vegetative regions (e.g., dry steppes and high forest-meadows), 28 local physiographic divisions (e.g., temperate desert and savanna mountains divisions), and four climate domains (e.g., polar and humid tropical domains). At the lower end of the scale, the faunas at k=21 are divided into patches of similar size to the ecoregion divisions, which are fundamentally physiographic entities (Appendix S3). At the upper end, the faunas at k=5 correspond almost exactly to the main ecoregion climatic domains (Appendix S3: cluster 1 corresponds roughly to the humid tropical domain, cluster 2 to the humid temperate domain, cluster 3 to the dry domain, and clusters 4 and 5 to the polar domain).

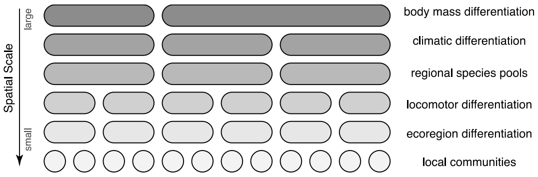


Figure 4. Diagram showing the spatial hierarchy of faunal differentiation. Local communities are composed of subsets of the regional species pool. Regional species pools are differentiated by climate, which occurs at a larger spatial scale than locomotor differences in faunas, but a smaller scale than body mass differences.

This hierarchy suggests that the frequency of different locomotor types is structured at a high level by climate and associated vegetative cover and at a lower level by physiographic variation, which echoes the finding by Polly et al. (2017) that locomotor proportions in North American Carnivora are sorted by vegetation openness and physiography. Body mass differed at a comparatively large spatial scale where *k*=8. The proportion of large (> 8,000 g) species varied substantially among these functionally distinct faunas, being lowest (5% of the faunal composition) in the south (**Mesoamerica**) and highest in the north (**High Arctic Canada**, **Eastern Beringia**, and **Northern High Canada**, >24%; Table 2). This trend was same in the median body masses, varying from 73 g in the southern fauna, 123 g - 183 g in the central faunas to 286 g - 933 g in the northern faunas. These patterns at the faunal level generally parallel Bergmann's rule (Ashton et al., 2000; Badgley and Fox, 2000; Meiri & Dayan, 2003; Blackburn & Hawkins, 2004; Rodríguez et al., 2006) and latitudinal and altitudinal biodiversity gradients (c.f., Badgley and Fox, 2000; Brown, 2001; Hillebrand, 2004). Cenograms showed that body mass gaps were primarily in the large size range (20,000-75,000 g, 110,000-240,000 g, and 240,000-625,000 g) and were found primarily in the mid-latitude and southern faunas. All northern community clusters have a gap at 500 g, which is consistent with open environments (Legendre 1986).

Faunas were climatically differentiated at an intermediate spatial scale of *k*=11 (Fig. 1b). We purposefully limited our consideration of climate to mean annual temperature and annual precipitation because of the link between these variables and vegetative biomes (Whittaker, 1975). Note, however, that these variables do not directly capture variables that are known to influence mammalian diversity, including seasonal temperature extremes, evapotranspiration, or elevation, which might be significantly different between faunas at smaller spatial scales (Badgley and Fox, 2000).

Diet did not differentiate faunas at any scale. This may be because the dietary categories were too fine (Lintulaakso & Kovarovic, 2016), but is more likely because all types of diet are likely to be mixed local communities whereas body mass and locomotor specialisations have a functional relationship to climate or landscape conditions that vary geographically.

These findings suggest a hierarchy of processes involved in the formation of regional species pools and local community assembly (Fig. 4). If we define regional species pools as those faunas that are differentiated by climatic conditions, which is only one aspect of Zobel's (1999) definition, then we find that functional locomotor traits associated with mobility and thus the ability to colonise local communities are differentiated at a smaller scale that is subequal to physiographic differences. However, body mass, which is associated more with temperature and openness of habitat, is differentiated at a larger scale. These results imply a series of hierarchical filters operating across the breadth of the North American continent. The lack of body mass differentiation in Europe is consistent with this hypothesis because the scale of climatic differentiation is less there. However, the absence of locomotor differentiation among European faunas is puzzling since in North America that differentiation is found at almost all scales.

Implications for interpreting paleontological faunas

Recovering complete local communities is notoriously problematic in palaeontology because of taphonomic filters and biases (e.g., Kidwell and Flessa, 1995; Kowalewski and Bambach, 2008; Miller et al., 2014). However, delimiting regional faunas, especially ones that persist through time as chronofaunas, is arguably a more reliable enterprise in the fossil record than in the extant world because of the same spatial and time averaging affects that help mask local community compositions (e.g., Woodburne, 1987; Eronen et al., 2009; Figueirido et al., 2012).

The hierarchical distribution of faunas, climate, and functional traits provides a framework for interpreting paleontological faunas in terms of climate. If our North American results are typical, faunal clusterings based on a combination of species occurrences, body size, and locomotor traits should correspond to some level of climatic and environmental differentiation. Spatial or temporal turnover in faunas should indicate some level of climatic and environmental turnover, a hypothesis that has been borne out in the fossil record in several studies (e.g., Fortelius et al., 2002; Eronen et al., 2009; Figueirido et al., 2012; Polly and Head, 2015). This hypothesis is not necessarily contradicted by the lack of functional trait differentiation in the European faunas since they were climatically distinct at a certain level; however, the lack of functional differentiation urges caution in interpreting paleontological faunas based on taxonomic similarity alone.

Our results suggest that the frequency of locomotor types in a fauna may be a guide to paleoenvironmental interpretation. Purely terrestrial locomotion dominates the northern faunas (38%-46% of the fauna), while subterranean-terrestrial species dominate the mid-latitude and southern faunas (35-51%). The northern faunas of **Northern High Canada, High Arctic Canada,** and **Eastern Beringia** lack subterranean species entirely, perhaps because of permafrost conditions that are found in these areas (Brown et al., 1997). Pocket gophers (*Thomomys*), for example, are known to be range-limited by soil conditions, especially hardness (Marcy et al., 2013) and the origin of digging adaptations has been associated with the onset of drier and more open environments from the Oligocene through Pliocene (Cook et al., 2000; Vizcaíno et al., 2016; Blois and Hadly, 2009). Subterranean species are found in the **Southern Canada** fauna, and even more frequently in the **Great Basin** and **Northern Mexico** faunas, that have varied soils associated with high topographic relief and variable conditions, both diurnally and seasonally. This combination of conditions may favour subterranean and subterranean-terrestrial species that look for shelter and food storage underground. **Mesoamerica**, with its tropical and subtropical forests, has a high proportion of arboreal and arboreal-terrestrial species (12% and 22% respectively). Arboreality is generally associated with dense tree cover, while arboreal-terrestrial species are associated with savanna and woodland environments (Reed 1997, 1998; Louys et al., 2011; Lintulaakso & Kovarovic, 2016). This distribution of locomotor categories among all mammals at the faunal level parallels sorting at the community level as measured by ratios of the hind limb in Carnivora at the local community level (Polly, 2010; Polly and Sarwar, 2014), a pattern that arose by trait-based clade sorting (Polly et al., 2017).

Our results confirm previous studies that showed that standing diversity and body size distributions are related to climate and could thus be useful for palaeoclimatic reconstruction (e.g., Legendre, 1986; Rosenzweig, 1995; Badgley and Fox, 2000). Cold regions (mean annual temperatures < -5 °C) have a low number of mammals (between 49 to 58 species). The proportion of large (> 8000 g) species is over 24% of the community composition, and the proportion of small species (< 500 g) is less than 51%. Relatively wet and mild regions (700 - 1050 mm per year; 0 - 11 °C MAT) have a moderate number of species (≈ 80) with large species making up between 10 - 21% of the fauna and small species between 53 - 61%. Dry areas with low precipitation, moderate temperatures, and high elevations (300 - 500 mm per year; 6 - 20 °C; > 1400 m) have a high number of species (140 - 150) with few large species (9%) and many small ones (65 - 68%). Warm and humid areas (> 23 °C, > 1700 mm per year) have a very high number of species (>240) with fewer than 5% large species and more than 70% small species.

5 Conclusions

Species pools are a key concept in functional trait ecology, providing the source pool for community assembly processes that are guided by trait-environment relationships (Zobel, 1999; Weiher and Keddy, 2001; McGill et al., 2006). We found that in North American mammals, the factors that influence the formation of regional species pools are themselves hierarchically distributed: faunas are differentiated by locomotor traits at fairly small scales, by climate at middling scales, and by body mass at larger scales. Interestingly, however, European mammal faunas are not differentiated by functional traits even though they are climatically differentiated at approximately the same scale as North American ones. We attribute this difference to the narrower European climate space and the possible imprint of anthropogenic effects on mammalian functional diversity. Paradoxically, these findings support Heikenheimo et al.'s (2007, 2012) hypothesis that faunal clusters are likely to have functional trait differentiation, but only in North America, not in Europe where Heikenheimo's study was based. The processes that result in functional, taxonomic, and climatic differentiation between faunas support the idea that clustering methods applied to taxa and traits in the fossil record can be used to measure palaeoclimatic and palaeoenvironmental differentiation through time and across space.

Acknowledgements

Comments from David Fox and two anonymous reviewers substantially helped us with the writing of the manuscript. We thank Joonas Lintulaakso who helped us with finalising the figures. Funding for KL came from the Emil Aaltonen Foundation and Nordenskiöld Society in Finland. Funding for PDP was provided by US National Science Foundation grant EAR 1338298 and by the Prepared for Environmental Change grand challenge initiative at Indiana University.

Data Accessibility

The distribution ranges for mammals in North America can be accessed via NatureServe (http://www.natureserve.org/, Patterson et al., 2003) and for Europe via Societas Europaea Mammalogica (http://www.european-mammals.org/, Atlas of European mammals, EMMA; Mitchell-Jones et al., 1999). Point sampled data using the equidistant 50 km grid are available at http://mypage.iu.edu/~pdpolly/Data.html for climate (mean annual temperature and annual precipitation; Willmott and Legates, 1988), elevation (Hastings and Dunbar, 1998), and biomes (Bailey & Hogg, 1986; Bailey, 1989). Species trait data in additional supporting information may be found in the online version of this article at the publisher's web-site:

Biosketch

Kari Lintulaakso is a PhD student at the University of Helsinki. He specialises in recent mammals and his main interest is linking current mammalian community structures with key environmental factors that can be used in palaeoclimatological and environmental studies.

P. David Polly is a vertebrate palaeontologist and evolutionary biologist. He is interested in mammalian evolution and the responses of both species and communities to large-scale environmental and climatic changes. His specialities are functional morphology, morphometrics, quantitative evolution, spatial analysis, and carnivores.

Jussi Eronen is investigating how humankind and society are capable of solving the looming environmental and climate crisis. He has researched how past climates have developed and what are the driving mechanisms, as well what controls the terrestrial biodiversity and ecosystems structures through time.

Author contributions: The study was conceived by JTE. Data were collected by KL and JTE, and analysed by PDP and KL. The results were interpreted by all authors. Writing the article was done by all authors with the main responsibility on KL and PDP

References

Ashton, K. G., Tracy, M. C., & Queiroz, A. D. (2000). Is Bergmann's rule valid for mammals?. *American Naturalist*,156, 390-415.

Ashton, N., Lewis, S.G., De Groote, I., Duffy, S.M., Bates, M., Bates, R., Hoare, P., Lewis, M., Parfitt, S.A., Peglar, S. & Williams, C. (2014). Hominin footprints from early Pleistocene deposits at Happisburgh, UK. *PLoS One*, 9, e88329.

Badgley, C. & Fox, D. L. (2000) Ecological biogeography of North American mammals: species density and ecological structure in relation to environmental gradients. *Journal of Biogeography*, 27, 1437–1467.

Badgley, C., Damuth, J., Theodor, J. & Werdelin, L. (2001). Global variation in ecological structure of mammalian faunas and paleoenvironmental inferences. *Journal of Vertebrate Paleontology*, 21 (suppl.), 30A.

Bailey, R. G. (1989). Explanatory supplement to ecoregions map of the continents. *Environmental Conservation*, *16*(04), 307–309.

Bailey, R. G. & Hogg, H. C. (1986). A world ecoregions map for resource reporting. *Environmental Conservation*, *13*(03), 195–202.

Blackburn, T.M. & Hawkins, B.A., 2004. Bergmann's rule and the mammal fauna of northern North America. *Ecography*, *27*(6), pp.715-724.

Blois, J. L. & Hadly, E. A. (2009). Mammalian response to Cenozoic climatic change. *Annual Review of Earth and Planetary Sciences*, 37, 181-208.

Bourgeon, L., Burke, A. & Higham, T. (2017). Earliest human presence in North America dated to the last glacial maximum: new radiocarbon dates from Bluefish Caves, Canada. *PLoS One*, 12, e0169486.

Brown, J.H. (2001). Mammals on mountainsides: elevational patterns of diversity. *Global Ecology and Biogeography*, 10, 101-109.

Brown, J., Ferrians, Jr., O.J, Heginbottom, J.A. & Melnikov, E.S., eds. (1997). Circum-Arctic map of permafrost and ground-ice conditions. Washington, DC: U.S. Geological Survey in Cooperation with the Circum-Pacific Council for Energy and Mineral Resources. Circum-Pacific Map Series CP-45, scale 1:10,000,000, 1 sheet.

Connor, E.F. and Simberloff, D., 1979. The assembly of species communities: chance or competition? Ecology, 60(6), pp.1132-1140.

Cook, J. A., Lessa, E.P., & Hadly, E. A. (2000). Paleontology, phylogenetic patterns, and macroevolutionary processes in subterranean rodents. In *Life Underground. The Biology of Subterranean Rodents*, eds. E. A. Lacey, J. L. Patton, & G. N. Cameron, pp. 332–69. Chicago, IL: Univ. Chicago Press

Damuth, J., Fortelius, M., Andrews, P., Badgley, C., Hadly, E.A., Hixon, S., Janis, C., Madden, R.H., Reed, K., Smith, F.A., Theodor, J., Van Dam, J.A., Van Valkenburgh, B. & Werdelin, L. (2002). Reconstructing mean annual precipitation based on mammalian dental morphology and local species richness. *Journal of Vertebrate Paleontology.* 22 (suppl.), 48A.

Eronen, J.T., Ataabadi, M.M., Micheels, A., Karme, A., Bernor, R.L. & Fortelius, M. (2009). Distribution history and climatic controls of the Late Miocene Pikermian chronofauna. *Proceedings of the National Academy of Sciences*, 106, 11867-11871.

Figueirido, B., Janis, C.M., Pérez-Claros, J.A., De Renzi, M. & Palmqvist, P. (2012). Cenozoic climate change influences mammalian evolutionary dynamics. *Proceedings of the National Academy of Sciences*, 109, 722-727.

Fox, B.J. & Brown, J.H., 1993. Assembly rules for functional groups in North American desert rodent communities. Oikos, pp.358-370.

Fortelius, M., Eronen, J., Jernvall, J., Liu, L., Pushkina, D., Rinne, J., Tesakov, A., Vislobokova, I., Zhang, Z. & Zhou, L. (2002). Fossil mammals resolve regional patterns of Eurasian climate change over 20 million years. *Evolutionary Ecology Research*, 4, 1005-1016.

Gotelli, N.J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, 81, 2606-2621.

Hall, E. 1981. *The mammals of North America. Vols. 1 and 2.* John Wiley, New York.

Hastings, D. & Dunbar, P. (1998). Development and assessment of the Global Land One-km Base Elevation digital elevation model (GLOBE). *IAPRS*, 32, 218–221.

Heikinheimo, H., Fortelius, M., Eronen, J. & Mannila, H. (2007). Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. *Journal of Biogeography*, *34*(6), 1053–1064.

Heikinheimo, H., Eronen, J.T., Sennikov, A., Preston, C.D., Oikarinen, E., Uotila, P., Mannila, H. & Fortelius, M. (2012). Convergence in the distribution patterns of Europe's plants and mammals is due to environmental forcing. *Journal of Biogeography*, *39*(9), 1633–1644.

Hillebrand, H., 2004. On the generality of the latitudinal diversity gradient. *The American Naturalist*, *163*(2), 192-211.

Holdridge, L. R. (1967). *Life Zone Ecology*. San Jose, Costa Rica: Tropical Science Center.

Jones, K.E., Bielby, J., Cardillo, M., Fritz, S.A., O'Dell, J., Orme, C.D.L., Safi, K., Sechrest, W., Boakes, E.H., Carbone, C. & Connolly, C. (2009). PanTHERIA: a species- level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, 90, 2648–2648.

Kidwell, S.M. & Flessa, K.W. (1995). The quality of the fossil record: populations, species, and communities. *Annual Review of Ecology and Systematics*, 26, 269-299.

Kowalewski, M. & Bambach, R.K. (2008). The limits of paleontological resolution. Pp. 1-48 in Harries, P. J. (ed.), *High-Resolution Approaches in Stratigraphic Paleontology*. Springer, Dordrecht.

Kraft, N.J. & Ackerly, D.D. (2010). Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecological Monographs*, 80, 401-422.

Lawing, A. M., Eronen, J. T., Blois, J. L., C. Graham, C. & Polly, P. D. 2016. Community functional trait composition and the effects of non-ecological processes. *Ecography*, 39, 1-13.

Legendre, S. (1986). Analysis of mammalian communities from the late Eocene and Oligocene of southern France. *Palaeovertebrata*, 16, 191-212.

Linder, H.P., de Klerk, H.M., Born, J., Burgess, N.D., Fjeldså, J. & Rahbek, C. (2012). The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography*, *39*(7), 1189–1205.

Lintulaakso, K. (2013). MammalBase — database of recent mammals. http:// www.mammalbase.net.

Lintulaakso, K. & Kovarovic, K. (2016). Diet and locomotion, but not body size, differentiate mammal communities in worldwide tropical ecosystems. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *454*, 20–29.

Louys, J., Meloro, C., Elton, S., Ditchfield, P. & Bishop, L. C. (2011). Mammal community structure correlates with arboreal heterogeneity in faunally and geographically diverse habitats: implications for community convergence. *Global Ecology and Biogeography*, *20*(5), 717–729.

Marcy, A. E., Fendorf, S., Patton, J. L. & Hadly, E. A. (2013). Morphological adaptations for digging and climate-impacted soil properties define pocket gopher (*Thomomys* spp.) distributions. *PloS One*, 8, e64935.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21, 178-185.

Meiri, S. & Dayan, T. (2003). On the validity of Bergmann's rule. *Journal of Biogeography*, 30, 331–351.

Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mücher, C.A. & Watkins, J.W. (2005). A climatic stratification of the environment of Europe. *Global ecology and biogeography*, *14*, 549–563.

Miljutin, A. (2009). Substrate utilization and feeding strategies of mammals: description and classification. *Estonian Journal of Ecology*, *58*, 60–71.

Miller, J.H., Behrensmeyer, A.K., Du, A., Lyons, S.K., Patterson, D., Tóth, A., Villaseñor, A., Kanga, E. & Reed, D. (2014). Ecological fidelity of functional traits based on species presence-absence in a modern mammalian bone assemblage (Amboseli, Kenya). *Paleobiology*, 40, 560-583.

Mitchell-Jones, A.J., Mitchell, J., Amori, G., Bogdanowicz, W., Spitzenberger, F., Krystufek, B., Vohralík, V., Thissen, J., Reijnders, P., Ziman, J.M. & Stubbe, C.M. (1999). *The atlas of European mammals* (Vol. 3). London: Academic Press.

Olson, E.C. (1952). The evolution of a Permian vertebrate chronofauna. *Evolution*, 6, 181-196.

Patterson, B., Ceballos, G., Sechrest, W., Tognelli, M., Brooks, T., Luna, L., Ortega, P., Salazar, I. & Young, B. (2003). *Digital distribution maps of the mammals of the western hemisphere, ver. 3.0.* NatureServe, Arlington, http://www.natureserve.org.

Polly, P.D. (2010). Tiptoeing through the trophics: geographic variation in carnivoran locomotor ecomorphology in relation to environment. *Carnivoran evolution: new views on phylogeny, form, and function*, (ed. by A. Goswami, and A. Friscia), pp. 374–401. Cambridge University Press, Cambridge.

Polly, P. D. & Sarwar, S. (2014). Extinction, extirpation, and exotics: effects on the correlation between traits and environment at the continental level. *Annales Zoologici Fennici,* 51, 209-226.

Polly, P. D. & Head, J. J. (2015). Measuring Earth-life transitions: ecometric analysis of functional traits from late Cenozoic vertebrates. In: P. D. Polly, J. J. Head, and D. L. Fox (eds.), *Earth-Life Transitions: Paleobiology in the Context of Earth System Evolution*. The Paleontological Society Papers, 21: 21-46. Yale Press, New Haven, CT.

Polly, P. D., Fuentes-Gonzales, J., Lawing, A. M., Bormet, A. K., & Dundas, R. G. (2017). Clade sorting has a greater effect than local adaptation on ecometric patterns in Carnivora. *Evolutionary Ecology Research*, 18, 61-95.

Reed, K. E. (1997). Early hominid evolution and ecological change through the African Plio-Pleistocene. *Journal of human evolution*, *32*(2), 289–322.

Reed, K.E. (1998). Using large mammal communities to examine ecological and taxonomic structure and predict vegetation in extant and extinct assemblages. *Paleobiology*, 384–408.

Rodríguez, M.Á., López- Sañudo, I.L. & Hawkins, B.A., 2006. The geographic distribution of mammal body size in Europe. *Global Ecology and Biogeography*, *15*(2), pp.173-181.

Rosenzweig, M.L. (1995). *Species Diversity in Space and Time*. Cambridge, United Kingdon: Cambridge University Press.

Shinichiro, T. (2013). boxplotdbl: Double Box Plot for Two-Axes Correlation. R package version 1.2.2. https://CRAN.R-project.org/package=boxplotdbl

Steinhaus, H. (1956). Sur la division des corp materiels en parties. *Bulletin of the Polish Academy of Sciences and Mathematics*, 4, 801–804.

Valverde, J.A. (1964). Remarques sur la structure et l'évolution des communautés de Vertébrés terrestres. *Revue d'Écologie (La Terre et La Vie)*, 111, 121–154.

van Dam, J.A., Andrews, P., Badgley, C., Damuth, J., Fortelius, M., Hadly, E.A., Hixson, S., Janis, C., Madden, R.H., Reed, K., Smith, F.A., Theodor, J., Van Valkenburgh, B. & Werdelin, L. (2001). Within-habitat mammal diversity and productivity and their recent patterns across latitude. *Journal of Vertebrate Paleontology*, 21 (suppl), 43A.

Vizcaíno, S. F., Bargo, M. S., Kay, R. F. & Milne, N. (2006). The armadillos (Mammalia, Xenarthra, Dasypodidae) of the Santa Cruz Formation (early-middle Miocene): an approach to their paleobiology. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 237, 255–69

Weiher, E. & Keddy, P. eds. (2001). *Ecological Assembly Rules: Perspectives, Advances, Retreats*. Cambridge, United Kingdom: Cambridge University Press.

Whittaker R. H. (1975). *Communities and Ecosystems*. New York, NY: MacMillan Publishing.

Willmott, K.M. & Legates, D.R. (1988). Global air temperature and precipitation: regridded monthly and annual climatologies (version 2.01). Newark, Delaware: Center for Climatic Research, University of Delaware.

Woodburne, M.O. ed. (1987). *Cenozoic Mammals of North America: Geochronology and Biostratigraphy*. Berkeley, California: University of California Press.

Zobel, M. (1999). The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence? *Trends in Ecology and Evolution*, 12, 266-269.

Zobel, M., van der Maarel, E., & Dupré, C. (1998). Species pool: the concept, its determination and significance for community restoration. *Applied Vegetation Science*, 1, 55-66.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site:
Appendix S1. Mammal species list with traits (body mass, locomotion, diet)
Appendix S2. Number of grid points in clusters and species trait statistics between clusters
Appendix S3. North American Core Clusters 3-21
Appendix S4. European Core Clusters 3-21
Appendix S5. Cenogram of North American Core Clusters at *k*=8