

Title: The legacy of extinct Neotropical megafauna on plants and biomes

Authors: Vinicius L. Dantas^{1,*}, Juli G. Pausas²

Affiliations:

¹ Institute of Geography, Federal University of Uberlandia (UFU), Av. João Naves de Avila, 2121, 38400-902, Uberlandia, MG, Brazil.

² Centro de Investigaciones sobre Desertificación, Spanish National Research Council (CIDE-CSIC), Ctra. Naquera Km. 4.5 (IVIA), Montcada, 46113, Valencia, Spain.

* Correspondence to: viniciusldantas@gmail.com

Abstract

Large mammal herbivores are important drivers of plant evolution and vegetation patterns, but whether current plant trait and ecosystem geography reflect the historical distribution of extinct megafauna is unknown. We address this question for the Neotropical biogeographic realm (i.e. South and Central Americas) by relating plant defence trait information at the ecoregion scale to climate, soil, and fire data, as well as to the historical distribution of extinct megafauna. Here we show that megafauna history explains a substantial percentage of defence trait variability. We also identified three distinct regions in Southern America (called ‘antiherbiomes’) characterised by convergent plant defence strategies, environmental patterns, and megafauna history. We also identified ecoregions that experienced a biome shift, from grassy- to forest-dominated, following the Pleistocene megafauna extinction. These results suggest that extinct megafauna left a significant imprint on the current plant trait and ecosystems biogeography of the Neotropics.

Introduction

In pre-historical times, especially over 10,000 years ago, a large proportion of our planet was populated by large and even gigantic mammals: the megafauna. Whereas most of these animals became extinct during the Pleistocene and Early Holocene ¹, there are important exceptions where a great diversity of large mammals still roam – such as in Africa and Asia. These regions provide unique opportunities to understand megafauna ecology and its effects on ecosystems. Evidence suggests that consumption of plant biomass and related disturbances by African megaherbivores can drive and maintain woodlands in alternative grassland states ²⁻⁴. Moreover, large mammal herbivores impose limits on ecosystem susceptibility to fire (i.e. grazers) and can even influence soil fertility in the long term ⁴⁻⁶. Thus, large mammal herbivores create and maintain their own grassy ecosystems ². Given the key role of megafauna in maintaining grassy ecosystems, their extinction likely resulted in the replacement of many herbivory-maintained savannas by forests and woodlands, or by fire-maintained savannas, across the world ^{7,8}.

Woody plant species living in herbivory-maintained ecosystems are characterised by morphological and physiological adaptations that reduce damage caused by large herbivores (i.e. antiherbivory defence traits) ⁷. While defence traits that are disadvantageous under present condition were likely lost in many places after megafauna extinction (as ecosystems have changed), other traits may have persisted as anachronical features ^{9,10}. These anachronisms provide a valuable opportunity to understand plant-megafauna interactions and could provide insights on switches from open grassy ecosystems (with abundant megafauna and highly defended plants) to closed canopy ecosystems. There are multiple mechanisms by which plants defend themselves from large herbivores ^{5,7,11-13}, and these mechanisms differ with climate and

availability of soil resources¹²⁻¹⁸. As a result, two broad savanna regions (hereafter called ‘antiherbiomes’, as analogous to biomes) can be recognised in Africa based on the convergence of plant defence attributes and strategies, and on their typical association with specific environmental conditions^{13,14,19}: 1) dry nutrient-rich savannas dominated by small-leaved heavily armed plants, defended with thorns, densely branched crowns, and nitrogen-based chemical leaf defences; and 2) mesic nutrient-poor savannas dominated by broad-leaved plants that mainly rely on leaf defences (e.g. leaf spines, acid detergent compounds, and lignin). A third ecosystem type can be readily identified in tropical Africa, that is, forests, in which low levels of megafauna herbivory and high productivity enable quick canopy escape, and, therefore, plants are largely undefended^{7,12}. We hypothesised that the distribution of extinct megafauna in the Neotropical biogeographic realm (South and Central Americas) left an imprint on current patterns of plant functional trait geography. As a result, we should be able to recognise large regions, similar to those observed in Africa, in which plant assemblages converge in defensive strategies due to specific environment-megafauna interactions (i.e. antiherbiomes).

In order to test these hypotheses, we compiled data on defence traits such as spinescence (leaf and stem/branches), leaf size, and wood density for woody species, and then scaled up these traits to the ecoregion using species distribution data²⁰, and correlated this data with pre-historical extinct megafauna (from the last 130,000 years) and extant mammal herbivore richness, mean body mass, and dominant diet type. In this study, the term megafauna refers to large mammal herbivores over 50 Kg that became extinct in pre-historical times, whereas the term extant mammal herbivore refers to both extant and recently extinct (by modern men) species regardless of body mass. We also tested for relationships between defence traits, climatic and soil variables, as well as

with disturbances, such as fire and cyclones (see Methods for details). Moreover, we tested the hypothesis that ecoregions now dominated by forest were dominated by grassy ecosystems for some time before megafauna extinction using information on megafauna richness and dominant diet, antiherbivory, and published fossil data for Pleistocene/Holocene vegetation.

Results

We found diet information for 53 of the 66 Neotropical megafauna species included in the PHYACINE dataset ²¹. Based on the dietary preferences, we classified the megafauna species as browsers (22 species), grazers (16 species), and mixed feeders (15 species). The overall and per diet type megafauna richness were strongly positively correlated (Table S1), and were highest in south-central South America (e.g. Gran Chaco in northern Argentina/southern Paraguay, and grassland-forest mosaics in southern Brazil). Body mass was highest in Central America and in the western part of South America (Fig. 1a-d), in moist habitats, seasonal climates and fertile soils, and was extremely low on islands. The richness of megafauna species (overall and within each diet class) increased for nutrient-poor soils and in ecoregions experiencing frequent fires, as well as for low and less seasonal rainfall (Table S2-S5), and decreased in islands.

All of the studied plant traits had a large percentage of their variability explained by megafauna history and most traits (except leaf size) had more variability explained by megafauna than by any other factor (Fig. 2; Table S3, S5). Wood density and leaf spines increased, whereas leaf size decreased with megafauna-richness. Stem spines only increased with megafauna body mass. Woody density was especially high and leaf size

especially low where extinct large grazer (i.e. megagrazers) richness was greater than that of large browsers (i.e. megabrowsers; Fig. 2; Table S3, S5). For these two traits, replacing megafauna by megabrowser richness in the selected model slightly improved model fit (Table S6). In contrast, extant mammal herbivores did not explain any proportion of the variability observed in antiherbivory defence traits (Fig. 2; Table S5). In fact, for leaf traits and woody density, the relationship with extant herbivore richness indicated bottom up control of these traits on modern mammal distribution (i.e. a significant negative relationship; Table S7). As we were looking for top-down herbivore effects, such relationships were discarded (see Methods).

A cluster analysis on the trait by ecoregion matrix (see Methods for details) supported the hypothesis that three distinct antiherbiomes can be recognized in the Neotropics (Fig. 3 and 4): (1) arid nutrient-rich ecosystems with small-leaved and thorny woody species (SLT for small-leaved thorny); (2) mesic nutrient-poor ecosystems dominated by broad-leaved plants with leaf level defences and high wood density (LDW for leaf defences/dense wood); and (3) megafauna sensitive assemblages (BLS for broad-leaved sensitive). The SLT antiherbiome mostly occurred at the northern and southern extremes of the Neotropical region (Fig. 3), and was especially associated with very large megafauna species, seasonally dry and cold climates, nutrient-rich soils (high pH and CEC), as well as with ecoregions in which fire is currently intense (Figs. 4). The second antiherbiome (LDW) dominated in tropical latitudes (Fig. 3), and was associated with regions rich in megafauna, especially small and mixed-feeder species, high mean annual temperature and precipitation, high fire frequency, and extremely nutrient-poor soils (Fig. 4). The third antiherbiome (BLS) was associated with few, but large, megafauna species and was mostly related to productive conditions, low fire activity, and forest vegetation (also deserts, in some areas; Fig. 4). Excluding leaf spines (the

trait for which we had fewer data) from the trait matrix mainly changed the classification for some high-altitude ecoregions in the Andes and in south Chile (Fig. S1).

We identified 29 forest-dominated ecoregions with evidence of having been formerly dominated by grassy ecosystems (savanna or grassland) during the Pleistocene, covering an area of 6,911,248 Km² (Fig. 5), of which most are currently dominated by moist forest (4,321,374 Km²). These areas could be identified based on their extremely high richness of megafauna and megagrazer species (i.e. greater than the 75% quantile), and because they were either classified as part of the SLT or LDW antiherbiomes. From the 21 fossil sites with evidence of a past savanna state, 16 are currently forest ecosystems or oceanic areas adjacent to a currently forest-dominated ecoregions. Moreover, numerous savanna patches are presently found in south-eastern Amazon areas identified in our study as Pleistocene savannas. Most of the savanna-dominated ecoregions of South America seems to have been savanna in the past (2,762,311 km²), except for the Llanos ecoregion (358,657 km²), with fewer megafauna and megagrazer species than the 75% quantile.

Discussion

For most of the studied traits, megafauna history was the strongest factor predicting trait variability and, even when not (i.e. leaf area), it was still highly important. In all of the cases, relationships were consistent with our hypothesis of anachronistic patterns resulting from a past selective filter exerted by now extinct megafauna. This filter interacts with environmental conditions in such a way that produces three antiherbiomes whose distributions exceed continental limits (i.e. these antiherbiomes occur in both the

Neo- and Afro-¹⁹ tropics, although a detailed study of the latter is lacking), as is the case with actual biomes. Once assembled, these antiherbiomes persist for millennia, despite the cascading effects of megafauna extinction on ecosystems, suggesting that these traits are not under strong selection during the Holocene. In fact, we found that modern mammal herbivores in the Neotropics are currently bottom-up controlled by some of the studied plant defensive traits (Table S7). Overall, our results support the hypothesis that megafauna biogeography is a key driver of plant trait patterns⁷, generating convergent adaptive responses on plant species and producing globally distributed antiherbiomes. More generally, these findings highlight the key role of history, disturbance, and species interactions for understanding global patterns of plant functional trait variability.

By combining antiherbiome information with geographical patterns of megafauna richness and diet composition, we were able to identify forest ecoregions likely to have been dominated by grassy ecosystems before the Holocene. This does not imply, however, that forests were absent from these ecoregions during the Pleistocene, as the high woody plant diversity in these currently forest-dominated ecoregions suggests an ancient status. Yet, the distribution of these forests is likely to have been much more fragmented than in recent times. This interpretation is consistent with the fossil record of these ecoregions, as well as the distribution of savanna patches within these forest-dominated ecoregions (Fig. 5). In addition to Amazonian forest ecoregions (Fig. 5; ²²), savanna patches are also found in the rainforests of the Atlantic forest ²³, and in the drier forest/woodland regions of the Chaco and the Caatinga ^{24,25}. Paleoclimate alone is unlikely to explain the replacement of these forests by grassy ecosystems during the Pleistocene (except in some areas in the south-eastern Amazon where models predict the occurrence of savannas during the Last Glacial Maximum and mid-Holocene) ²⁶. Instead, megaherbivores must have played a key role in maintaining grassy biomes

during the Pleistocene. Thus, the past distribution of Neotropical savanna and grasslands was probably much broader than presently.

Megafauna richness was the main predictor of wood density (Fig. 2E; Table S3). Wood density is a widely studied functional trait whose variation is often attributed to climate and soil ^{7,27}. Yet, in broad-scale studies, individual climate and soil variables often predict a small proportion of the variability in this trait. For instance, although our results for mean annual temperature and soil moisture (explaining 13% of the variability in wood density) were consistent with a recent global study ²⁷, megafauna history was much more important in our models (explaining twice that variability, i.e., 30%). We also found that high hurricane recurrence is an important disturbance driver of wood density in plant species ^{7,28}. Thus, we suggest that wood density is an important general disturbance resistance trait that is especially useful for avoiding damage by megafauna. We also found that wood density was higher where grazers were relatively more abundant (Fig. 2), suggesting that this defence trait is especially adaptive under grazer-controlled fire regimes. This vision is consistent with evidence that wood density values are usually low in fire-maintained savannas in which grazers are absent⁷. This pattern is not very evident in our ecoregion level results because Neotropical savanna-dominated ecoregions include numerous forest areas. More studies are necessary to fully elucidate the relationships between wood density and grazer abundance.

Megafauna distribution significantly explained stem spine variation across ecoregions (Fig. 2F, Table S3). This variability was, however, explained by body mass rather than richness. In Africa, large browsers and mid-size social mixed-feeder species are important predictors of stem spines, although the latter are presumably more important ¹². In our study, including mixed-feeder did not improve model fit (not shown). This difference may be related to the fact that South America's extinct megafauna included

larger herbivores and fewer grazing species¹, as well as to the little available information about biomass per area and social behaviour for extinct Neotropical species. Like in Africa, stem spines were associated with dry areas (Fig. 2F; Table S3)^{12,13}, in which the selective pressure exerted by herbivores on individual plants is likely to be strong due to: 1) low plant density and productivity, and low plant stature; and 2) high soil fertility status (Table S5) that increases the consumption rate of scattered individual plants.

In a previous study²⁹, it was found that mean annual rainfall (MAR) was the most important climatic variable explaining variation in leaf size (explaining 22% of its variability). Yet, this study did not consider soil or disturbances as alternative predictors. Because climate, soil, and disturbances are often correlated, it is impossible to know the extent to which these previous results reflect direct or indirect effects. Here we show that soil pH, the most important variable predicting leaf size in our study (positive association), was strongly negatively correlated with mean annual precipitation ($r = -0.78$; see Table S5) and that the model including the former instead of the latter had a better fit (see Methods and Fig. 2). Consistently, we found that soil pH alone explains a much larger fraction than previously demonstrated for MAR (i.e. 32%), suggesting that the previously reported rainfall effect could be indirect. While we cannot rule out a direct effect of soil fertility in leaf size, the positive association between these two variables shown here was consistent with our theoretical predictions that leaf size is especially useful as a defence mechanism in nutrient-rich soils and drier climates. If this mechanism is correct, large herbivores could be even more relevant for leaf size than depicted in the results for the megafauna indicators.

Each of the detected antiherbiomes was associated with specific conditions (Fig. 4). As predicted, the SLT was associated with dry and nutrient-rich conditions, the LWD with

moist and nutrient-poor conditions, and the BLS with moist and intermediate soil fertility. In addition, fire frequency was highest in the LDW, whereas fire intensity was highest in the SLT antiherbiome (Fig. 4). This pattern was similar to that previously demonstrated for fire- and herbivory- maintained savannas in Africa², especially after acknowledging that grazers currently control fire intensity and frequency in Africa but not in the Neotropics^{2,5}. We also found that the megafauna assemblages of the three antiherbiomes were very distinct (Fig. 4). The SLT and the LDW concentrated a much greater number of large grazer and browser species than the BLS (Fig. 1 and 4). However, the SLT contained larger and fewer species than the LDW antiherbiome (Fig. 1 and 4). This was related to the smaller number of mixed-feeder species in the SLT (despite of the similar number of grazers and browsers), suggesting a higher tendency for diet specialization in this antiherbiome. There is also evidence of differences in species composition between ecoregions within the SLT and LDW antiherbiomes³⁰. Unfortunately, there are insufficient detailed faunal studies for these antiherbiomes in Africa, and this prevents a rigorous comparison with the patterns that we are reporting here.

That biomes changed after megafauna extinction implies a shift in the selective pressures experienced by plants. Whereas in grazer-dominated and less productive ecosystems megafauna extinction must have increased fire activity, in browser-dominated and/or productive ecosystems (e.g. current rainforests), it likely led to reduced light availability (woody encroachment)³¹. Functional trait trade-offs probably limited the accumulation of adaptations to both ancient and novel conditions. This process is likely to explain a substantial proportion of the residual variability in our models, that is, trait-megafauna associations were much stronger in the past. Moreover, the introduction of exotic megafauna species in many places (livestock) can replace the

ecosystem functions of extinct species³², which could shift the geography of defence traits away from past patterns. For instance, it was recently shown that the expression of leaf spines, the trait with the lowest percentage of variability explained by the selected model (but also with poorer data quality), is substantially affected by cattle density¹⁸. This emphasises the lability of such a trait, which is probably less pronounced for stem spines¹⁸. Finally, the limited availability of trait and distribution data for the Neotropics could also be a potential source unpredictability in our models. Yet, we believe that robust macroecological patterns can be depicted even with scattered data, as long as the data are representative and available in a spatial resolution that is adequate for investigating the geographical pattern under scrutiny (e.g. 28). While we attempted to compile data from studies from all across the Neotropics, we filtered our dataset to include only plant species native to South and Central Americas (see Methods). Thus, extrapolations for North American ecoregions should be interpreted cautiously. We also suggest that including data on herbaceous species traits could shed additional light on Neotropical antiherbiomes by evidencing the evolutionary and biogeographic influence of grazers in Neotropical vegetation.

Megafauna history explains a large fraction of the biogeographic variability in plant functional traits. This historical effect has been largely neglected in the ecological and biogeographical literature^{7,34}. Moreover, we found that the interplay among herbivory regimes, climate, soil, and plant traits in geological scales results in the emergence of globally distributed antiherbiomes characterised by convergent antiherbivory defence strategies. These antiherbiomes represent one of the most striking broad-scale biological anachronisms. Finally, by combining the observed trait patterns, megafauna history, and fossil data we were also able to show evidence of grassy- to forest- dominated ecosystem shifts after megafauna extinction in central South America. These shifts are

unlikely to have resulted from paleoclimate alone and suggest that megafauna developed a key role in controlling the distribution of grassy ecosystems in the continent up to the Holocene. These results highlight the importance of past and present megafauna distributions to understand plant and ecosystem biogeography.

Methods

Species level trait data

We compiled species level data on wood density (WD), spinescence, and leaf size for tropical and extra-tropical South and Central American woody species (the Neotropical biogeographic realm). WD was obtained for 2577 species from ³⁵. The presence or absence of stem (and branch) spines (mostly thorns, but also prickles) were obtained from ³⁶ for Neotropical savanna and forest species (1004 species) and complemented with other literature sources for other ecoregions (listed in the supplementary materials). Our final stem spine dataset included 2520 woody species. Leaf size data were obtained for 2660 woody species from ²⁹. For many species we had more than one trait value, so we computed the species trait value as the mean for quantitative traits (WD, leaf size), and as the maximum for binary traits (spinescence; 0 for absence and 1 for presence). This later decision was based in the assumption that omitting the presence of spine is more common than reporting that a spineless species has spines, and on the fact that, in some species, spinescence is inducible ¹⁸. Because of the lack of a large dataset of leaf level defences for woody species (e.g. chemical leaf defences, leaf spines), we used data on leaf spinescence of palm species as a proxy (694 species) from the global Palm Traits Database 1.0 ^{33,37}. In African savannas, leaf spines are found in resource

conservative plants and this trait is positively correlated with other leaf level defences, such as acid detergent fiber and lignin¹⁷.

From Species to Ecoregions

We obtained geographical distribution data (coordinates) from GBIF for all of the species in each species-trait dataset (Data available from GBIF; WD: <https://doi.org/10.15468/dl.3vua3x>; Stem spines: <https://doi.org/10.15468/dl.sj8hj5>; Leaf spines: <https://doi.org/10.15468/dl.vv8gw4>; Leaf size: <https://doi.org/10.15468/dl.k98nxc>). From these coordinates, we determined the ecoregion and biogeographical realm³⁸ and cropped out occurrences falling outside of the Neotropical realm. Based on the occurrence data, we built a species abundance (columns) by ecoregion (rows) matrix for each trait.

Using the respective matrices, we obtained ecoregion scale abundance weighted means for WD, Stem Spines, and Leaf Size by: 1) Multiplying species abundance in each grid cell of the ecoregion by the mean species value; 2) Summing up the row values; 3) dividing the resulting row sum by the total species abundance (row sum prior to trait multiplication), and 4) calculating the ecoregion mean (across all of the grid cells). This process resulted in weighted means for WD and stem spinescence for 173 ecoregions, and Leaf Size for 174, out of 179 Neotropical ecoregions. For leaf spinescence we used a similar approach, although, because of the fewer species, the abundance estimate from GBIF was unreliable. Thus, we transformed the ecoregion species abundance to presence/absence before multiplying the trait values (0/1 for absence/presence). We obtained leaf spinescence data for 159 out of the 179 Neotropical ecoregions.

Historical Megafauna and Herbivore Mammal Distributions

We obtained data on historical distribution of megafauna species from the MegaPast2Future/PHYLACINE_1.2 dataset ²¹, a dataset containing distribution maps (96.5 km of spatial resolution) and functional traits for mammal species of the last 130,000 years. Specifically, we extracted the potential present distribution of extinct mammal (coded “EP” for IUNC status) whose body mass was higher than 50 kg (megafauna), and for which at least 90 % of their diet consisted of plants (i.e. strict herbivores). The potential distribution of extinct species in this database was derived from the present distribution of extant species that are known, from the fossil record, to have coexisted with the extinct species. In this approach, an extinct species was considered to have been present in a given grid cell if at least 50% of the extant species found in the fossil (and subfossil) record coexisting with the extinct species currently occur in the cell ³⁹. This approach assumes that, since extant and extinct species coexisted in the same locations, they must have had similar ecological requirements. It also assumes that megafauna extinction had anthropogenic causes, instead of causes related to climate change ³⁹, which is largely accepted in the literature ^{40,41}.

For each Ecoregion, we calculated two megafauna-related metrics: extinct megafauna species richness (M_{rich}) and their mean body mass (M_{bm}). For this, we began by cropping the distribution maps of the mammal species to the Neotropical realm and assigned the corresponding dominant ecoregion to each of the grid cells of the megafauna map. Then, for each ecoregion, we built a grid cell (rows) by megafauna species (columns) matrix containing 1 for presence or 0 for absence for each megafauna species. For M_{rich} , we summed the matrix rows to obtain pixel level richness and

calculated the ecoregion level mean of each ecoregion. For M_{bm} , we multiplied the matrix columns (species names) by the body mass of the respective megafauna species before computing the grid cell mean, and, then, we calculated the ecoregion level mean. We also obtained diet preference information from the literature for most megafauna species that occurred in the Neotropical region in this dataset (references listed in the Supplementary Material). Based on this information, we calculated the richness of large grazer (MB_{rich} for megabrowser richness), grazer (MG_{rich} for megagrazer richness), and mixed-feeder (MMf_{rich} for mega mixed-feeder richness) species. We classified species based on the dominant diet component as grazer or browser, except when the proportion of browse and graze on the diet was very similar (difference in percentage of each component smaller than 10%), in which case the species was classified as mixed feeder.

We also compiled data on the distribution of extant and recently extinct (i.e. extinct during modern times) herbivore mammal species (for simplicity, called ‘extant’ species in this study). This data compilation had the objective of testing the alternative hypothesis that trait variation was explained by the selective pressure exerted by extant mammal herbivores. These data were obtained from the same dataset and represent the hypothesised original distribution of these species prior to anthropogenic reduction of their ranges. For this, the dataset was produced using the same approach described above for extinct megafauna species. Here, we also included strict herbivores (at least 90% of the diet constituted of plants), although we did not use a size threshold. We subsequently calculated the same metrics as for the extinct megafauna species (except for the richness of mixed-feeders as our source for diets⁴² labelled species according to dominant feeding pattern into grazer or browser).

Climate, Soil, Fire, Insularity and Hurricanes

For each Ecoregion, we obtained data on climate (mean annual precipitation and temperature, and rainfall seasonality) and soil (sand content, pH, and cation exchange capacity). Climate data was obtained from WorldClim 2.1 (10 minute spatial resolution) and was based on climate data from 1970 and 2000 ⁴³. Soil data were obtained from SoilGrids dataset (5 km of spatial resolution) ⁴⁴, and consisted of mean values for two depths, 0.05 and 2 m. We calculate Ecoregion level means for all of the soil and climate variables after intersecting the climate and soil grid maps with the ecoregion map.

We obtained the number (a proxy for frequency) and intensity of wildfires per ecoregion area using the MODIS active fire location product (MCD14ML) ⁴⁵. We only considered fires with detection confidence of 95% or higher occurring from November 2000 to December 2019 (both included). To ensure that only wildfires were considered, we associated each fire pixel with a land cover type (300 m of spatial resolution) from ⁴⁶ for a buffer area of 1000 m surrounding the fire pixel centroid. We excluded all of the fires occurring in areas in which more than 10% of the surrounding land cover pixels corresponded to agricultural, urban and water classes. We calculated the number of wildfires per ecoregion area by dividing the fire count of each Ecoregion by the ecoregion area, and multiplying the resulting value by the proportion of vegetated land cover pixels (same classes used to exclude fires in anthropogenic areas and water bodies above). Fire intensity was calculated as the average fire radiative power across all detected wildfires in the ecoregion. Ecoregions lacking large preserved vegetated areas (criteria above) were excluded from subsequent analyses.

Using the ecoregion map, we also classified ecoregions into insular (1), when most of the ecoregion area was located in islands, vs. continental (0), otherwise. This was

performed because island biogeography theory predicts that, in island, species richness should be low due to low colonization and high extinction rates. Insularity has also been shown to reduce megafauna body size (i.e. the island rule), even though the mechanisms are not fully understood⁴⁷. We also compiled data on hurricane activity, as woody density was suggested to confer resistance against this disturbance²⁸. We used data from 1990 to 2019 from the HURDAT2 dataset⁴⁸, containing six-hourly information about the location of all of the known tropical and subtropical cyclones (0.1° latitude/longitude). We used the sum of hurricane occurrences per ecoregions divided by ecoregion area as an indicator of hurricane activity.

Statistical Analyses

To understand megafauna patterns, we fit (multiple) regression models with habitat-related (fire, climate, soil) variables and insularity as predictors. We expected that megafauna richness in general was higher under Neotropical savanna conditions (frequent fire and nutrient-poor soils)¹. We also expected that megafauna richness and body mass were affected negatively by insularity (i.e. following the island biogeography theory and island rule). Before the analyses, we tested the correlation among all of the variables that would eventually be entered as predictors in the same model (Table S1), in order to avoid multicollinearity associated with highly correlated variables (i.e. $r \geq 0.60$). Since mean annual precipitation and pH were strongly positively correlated ($r = 0.78$), for all of the analyses (including the analyses with functional traits, described below), model selection was performed separately for these two variables (i.e. two different model selection procedures, one containing each of the two variables among the initial set of predictors). We selected the best among the two resulting models as

that with the lowest AIC (differences higher than two points in all of the cases). To make sure that no multicollinearity remained we also calculated the Variation Inflation Factor (VIF) for all of the predictor variables as $1/\text{tolerance}$, where tolerance is calculated as $1 - R^2$ of all of the model regressing a predictive variable against all of the other predictors. In all of the models, VIF was 3.33 or smaller (i.e. a tolerance of 0.30 or higher), indicating absence of multicollinearity.

Model simplification was carried interactively using stepwise (both forward and backward) searching for the model with the lowest AIC (using R's "step" function) and subsequently retaining only the significant variables ($p \leq 0.05$). We calculated the average contribution of each predictor variable in the selected model as the mean difference in R^2 before and after removing the target variable from models containing all of the possible subset combinations of the selected predictor variables, including the full selected model. This was carried out with the R package "dominanceanalysis".

For testing whether the studied plant functional traits were related to our megafauna indicators, we fit linear models to WD and leaf size, and generalized linear models (GLM; binomial family) for spinescence, using ecoregion as the unit. For spinescence, we used the matrix containing spiny and non-spiny species abundance (for stem spines) or number of species with or without spines (for leaf spines; see above) as response variables. The predictor variables included the animal indicators for the extinct megafauna species and for extant herbivores, as well as climate, soil, and fire predictors (and, for WD, the hurricane counts). Because total, as well as megagrazer, megabrowser, and mega mixed-feeder species richness were strongly positively correlated (Table S1), we only used megafauna richness, body mass, and the richness difference between grazers and browsers (Fig. S1). We used these same variables for extant species (i.e. mammal herbivore richness, body mass and the richness difference

between grazers and browsers). We did not identify strong correlations among extinct megafauna and extant herbivore indicators (Table S5); therefore, these variables were entered simultaneously in the same initial models. We also evaluated whether exchanging megafauna richness by megabrowser richness in the final model resulted in a significant increase in model fit.

As with the analyses of the megafauna indices, we also calculated the average predictor contribution in these models. For this, we used the MacFadden Pseudo- R^2 in the GLM models as implemented in the “pscl” and “dominanceAnalysis” packages for R, as this statistic is the most comparable with R^2 from linear multiple regression (Maximum Likelihood and Cragg and Uhler’s Pseudo- R^2 were also calculated for the logistic models). Standardized coefficients were also calculated for all models. Islands were not included in these models, as island plants were expected to respond differently due to the effects of insularity on animal species richness, precluding megafauna and extant mammal richness from being accurate proxies for consumer abundance.

For all of the general linear regression models, assumptions of normality and homoscedasticity in the residuals were checked using the Kolmogorov-Smirnov and Breusch-Pagan tests, respectively. In some cases, heteroscedasticity was detected and, thus, the significance of the coefficients was tested using heteroskedasticity-consistent covariance matrix estimation. Overdispersion in the generalized linear model was also detected and dealt with using overdispersed binomial logit models, as implemented in the “dispmod” package for R, in which weights are interactively calculated and used to maintain the residual deviance lower than the degrees of freedom.

To test the prediction that Neotropical ecoregions could be broadly classified into the three hypothesised antiherbiomes, we used hierarchical clustering on principal

component axes of the ecoregion by trait matrix (four traits, standardized to zero mean and unit variance). We selected the number of clusters associated with the highest loss of inertia (within group variability) when progressively increasing the number of clusters, using the R package “FactoMiner”. This procedure allowed the recognition of large regions characterised by specific patterns of defence strategies (‘antiherbiomes’). We subsequently tested for megafauna and environmental differences among the resulting antiherbiomes to verify whether trait, climate and soil patterns matched those described for African ecosystems, and to understand the megafaunal differences among the antiherbiomes. For these comparisons, we used Kruskal-Wallis and post-hoc pairwise Dunn tests, using the Benjamini & Hochberg (1995) correction of P-values for multiple comparisons in both cases, and exclusively included continental ecoregions. For spines, we used the proportion of spinescent plants/species (rather than the number of “yes” and “no” used on previous analyses). Because palms were missing from 20 ecoregions, we completed the values for these ecoregions using predicted model probabilities. To evaluate the sensitivity of the cluster result to this trait, we repeated the cluster analysis without leaf spines (see Fig. S2).

Finally, we identified forest ecoregions most likely to have experienced a biome shift after megafauna extinction. We considered that these ecoregions were those that: 1) are currently forest-dominated; 2) were classified as part of an antiherbiomes with highly defended species; 3) were megafauna- and, especially, megagrazer- rich during the Pleistocene (values equal or greater than the 75% quantile; 14 species for M_{rich} and 3 grazer species for MG_{rich}). We compared the distribution of these areas with fossil evidence from the Last Glacial Maximum and mid-Holocene^{26,50}. For this, we also used information about the present dominant vegetation type in the fossil sites to segregate savanna-forest shifts from data coming from savanna patches within forest or long-term

savannas. We also contrasted the patterns with the location of savanna patches within the Amazon forest region from ⁵¹.

All statistical analyses and data handling were carried out in the R environment, using the previously mentioned packages, as well as the packages FDA, gridExtra, grid, lattice, lmtest, latticeExtra, olsrr, pscl, raster, rgbif, rgdal, and sandwich.

Data availability statement

All the data supporting this study will be made available upon acceptance in a public repository

References

1. Owen-Smith, N. 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 (2013).
2. Dantas, V. L., Hirota, M., Oliveira, R. S. & Pausas, J. G. Disturbance maintains alternative biome states. *Ecol. Lett.* **19**, 12–19 (2016).
3. Pellegrini, A. F. A., Pringle, R. M., Govender, N. & Hedin, L. O. Woody plant biomass and carbon exchange depend on elephant-fire interactions across a productivity gradient in African savanna. *J. Ecol.* **105**, 111–121 (2016).
4. Staver, A. C. & Bond, W. J. Is there a ‘browse trap’? Dynamics of herbivore impacts on trees and grasses in an African savanna. *J. Ecol.* **102**, 595–602 (2014).

5. Dantas, V. L. & Pausas, J. G. The lanky and the corky: fire-escape strategies in savanna woody species. *J. Ecol.* **101**, 1265–1272 (2013).
6. Doughty, C. E. Herbivores increase the global availability of nutrients over millions of years. *Nat. Ecol. Evol.* **1**, 1820–1827 (2017).
7. Dantas, V. L. & Pausas, J. G. Megafauna biogeography explains plant functional trait variability in the tropics. *Glob. Ecol. Biogeogr.* 1–11 (2020)
doi:10.1111/geb.13111.
8. Doughty, C. E., Faurby, S. & Svenning, J. C. The impact of the megafauna extinctions on savanna woody cover in South America. *Ecography*. **39**, 213–222 (2016).
9. Janzen, D. H. & Martin, P. S. Neotropical anachronisms: The fruits the gomphotheres ate. *Science*. **215**, 19–27 (1982).
10. Galetti, M. *et al.* Ecological and evolutionary legacy of megafauna extinctions. *Biol. Rev.* **93**, 845–862 (2018).
11. Sebata, A. & Ndlovu, L. R. Effect of leaf size, thorn density and leaf accessibility on instantaneous intake rates of five woody species browsed by Matebele goats (*Capra hircus* L) in a semi-arid savanna, Zimbabwe. *J. Arid Environ.* **74**, 1281–1286 (2010).
12. Charles-Dominique, T. *et al.* Spiny plants, mammal browsers, and the origin of African savannas. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E5572–E5579 (2016).
13. Wigley, B. J., Fritz, H. & Coetsee, C. Defence strategies in African savanna trees. *Oecologia* **187**, 797–809 (2018).

14. Wigley, B. J., Bond, W. J., Fritz, H. & Coetsee, C. Mammal Browsers and Rainfall Affect Acacia Leaf Nutrient Content, Defense, and Growth in South African Savannas. *Biotropica* **47**, 190–200 (2015).
15. Goheen, J. R., Young, T. P., Keesing, F. & Palmer, T. M. Consequences of herbivory by native ungulates for the reproduction of a savanna tree. *J. Ecol.* **95**, 129–138 (2007).
16. Armani, M., Charles-dominique, T., Barton, K. E. & Tomlinson, K. W. Developmental constraints and resource environment shape early emergence and investment in spines in saplings. *Ann. Bot.* **124**, 1133–1142 (2019).
17. Armani, M. *et al.* Structural defense is coupled with the leaf economic spectrum across saplings of spiny species. *Oikos* (2020) doi:10.1111/oik.06960.
18. Göldel, B., Araujo, A. C., Kissling, W. D. & Svenning, J. C. Impacts of large herbivores on spinescence and abundance of palms in the Pantanal, Brazil. *Bot. J. Linn. Soc.* **182**, 465–479 (2016).
19. Huntley, B. J. Southern African Savannas. in *Ecology of Tropical Savannas* (eds. Huntley, B. J. & Walker, B. H.) 101–119 (Springer-Verlag, 1982).
20. Olson, D. M. *et al.* Terrestrial Ecoregions of the World: A New Map of Life on Earth. *Bioscience* **51**, 933–938 (2001).
21. Faurby, S. *et al.* MegaPast2Future/PHYLACINE_1.2: PHYLACINE Version 1.2.1 (Version v1.2.1). *Zenodo* (2020) doi:http://doi.org/10.5281/zenodo.3690867.
22. De Carvalho, W. D. & Mustin, K. The highly threatened and little known

- Amazonian savannahs. *Nat. Ecol. Evol.* **1**, 1–3 (2017).
23. Overbeck, G. E. *et al.* Brazil's neglected biome: The South Brazilian Campos. *Perspect. Plant Ecol. Evol. Syst.* **9**, 101–116 (2007).
 24. Grau, H. R. *et al.* Natural grasslands in the Chaco. A neglected ecosystem under threat by agriculture expansion and forest-oriented conservation policies. *J. Arid Environ.* **123**, 40–46 (2015).
 25. Dexter, K. G. *et al.* Inserting Tropical Dry Forests Into the Discussion on Biome Transitions in the Tropics. *Front. Ecol. Evol.* **6**, 1–7 (2018).
 26. Arruda, D. M., Schaefer, C. E. G. R., Fonseca, R. S., Solar, R. R. C. & Fernandes-Filho, E. I. Vegetation cover of Brazil in the last 21 ka: New insights into the Amazonian refugia and Pleistocenic arc hypotheses. *Glob. Ecol. Biogeogr.* **27**, 47–56 (2018).
 27. Bruelheide, H. *et al.* Global trait–environment relationships of plant communities. *Nat. Ecol. Evol.* **2**, 1906–1917 (2018).
 28. Curran, T. J., Gersbach, L. N., Edwards, W. & Krockenberger, A. K. Wood density predicts plant damage and vegetative recovery rates caused by cyclone disturbance in tropical rainforest tree species of North Queensland, Australia. *Austral Ecol.* **33**, 442–450 (2008).
 29. Wright, I. J. *et al.* Global climatic drivers of leaf size. *Science.* **12**, 917–921 (2017).
 30. Varela, L., Tambusso, P. S., Patiño, S. J., Di Giacomo, M. & Fariña, R. A. Potential Distribution of Fossil Xenarthrans in South America during the Late

- Pleistocene: co-Occurrence and Provincialism. *J. Mamm. Evol.* **25**, 539–550 (2018).
31. Johnson, C. N. Ecological consequences of Late Quaternary extinctions of megafauna. *Proc. R. Soc. B Biol. Sci.* **276**, 2509–2519 (2009).
 32. Lundgren, E. J. *et al.* Introduced herbivores restore late pleistocene ecological functions. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7871–7878 (2020).
 33. Lim, J. Y., Svenning, J. C., Gödel, B., Faurby, S. & Kissling, W. D. Frugivore-fruit size relationships between palms and mammals reveal past and future defaunation impacts. *Nat. Commun.* **11**, 1–13 (2020).
 34. Pausas, J. G. & Bond, W. J. Humboldt and the reinvention of nature. *J. Ecol.* **107**, 1031–1037 (2019).
 35. Zanne, A. E. *et al.* Global wood density database. *Dryad* (2009)
doi:<http://hdl.handle.net/10255/dryad.235>.
 36. Dantas, V. L. & Pausas, J. G. Data from: Megafauna biogeography explains plant functional trait variability in the tropics. *Dryad*, (2020).
 37. Kissling, W. D. *et al.* Data from: PalmTraits 1.0, a species-level functional trait database for palms worldwide, v4. *Dryad* (2019)
doi:<https://doi.org/10.5061/dryad.ts45225>.
 38. Dinerstein, E. *et al.* An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *Bioscience* **67**, 534–545 (2017).
 39. Faurby, S. & Svenning, J. C. Historic and prehistoric human-driven extinctions have reshaped global mammal diversity patterns. *Divers. Distrib.* **21**, 1155–1166

(2015).

40. Andermann, T., Faurby, S., Turvey, S. T., Antonelli, A. & Silvestro, D. The past and future human impact on mammalian diversity. *Sci. Adv.* **6**, eabb2313 (2020).
41. Sandom, C., Faurby, S., Sandel, B. & Svenning, J. C. Global late Quaternary megafauna extinctions linked to humans, not climate change. *Proc. R. Soc. B Biol. Sci.* **281**, (2014).
42. Kissling, W. D. *et al.* Data from: Establishing macroecological trait datasets: digitalization, extrapolation, and validation of diet preferences in terrestrial mammals worldwide. *Dryad*. doi: <https://doi.org/10.5061/dryad.6cd0v>
43. Fick, S. E. & Hijmans, R. J. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
44. Hengl, T. *et al.* SoilGrids1km — Global soil information based on automated mapping. *PLoS One* **9**, e105992 (2014).
45. Giglio, L., Schroeder, W., Hall, J. V. & Justice, C. O. MODIS Collection 6 Active Fire Product User’s Guide Revision B. NASA (2018).
46. Defourny, P. *ESA Land Cover Climate Change Initiative (Land_Cover_cci): Land Cover Maps, v2.0.7.* (Centre for Environmental Data Analysis, 2017).
47. Faurby, S. & Svenning, J.-C. Resurrection of the Island Rule: Human-Driven Extinctions Have Obscured a Basic Evolutionary Pattern. *Am. Nat.* **187**, 812–820 (2016).
48. Landsea, C. W. & Franklin, J. L. Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Weather Rev.* **141**, 3576–3592

(2013).

49. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
50. Mayle, F. E., Beerling, D. J., Gosling, W. D. & Bush, M. B. Responses of Amazonian ecosystems to climatic and atmospheric carbon dioxide changes since the last glacial maximum. *Philos. Trans. R. Soc. B Biol. Sci.* **359**, 499–514 (2004).
51. Devecchi, M. F. *et al.* Beyond forests in the Amazon : biogeography and floristic relationships of the Amazonian savannas. *Bot. J. Linn. Soc.* **20**, 1–26 (2020).

Author contributions: VLD conceived the idea, with inputs from JGP. VLD compiled and analyzed the data and wrote the first version of the manuscript. JGP contributed to revisions and authorized submission.

Competing interests: Authors declare no competing interests.

Materials & Correspondence: correspondence and material requests should be addressed to Vinícius L. Dantas

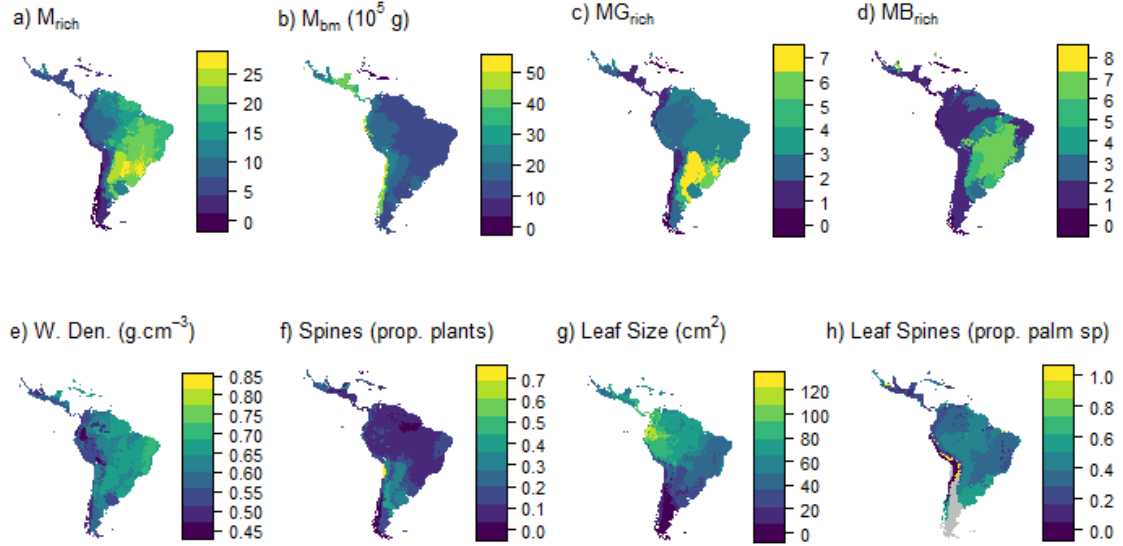


Fig. 1: Geographical variation in extinct megaherbivore species distribution (A-D) and antiherbivory resistance traits (E-H) among ecoregions of the Neotropical biogeographic realm. A): mean extinct megafauna species richness (M_{rich}); B): mean extinct megafauna species mean body mass (M_{bm}); C and D): mean extinct megagrazer (MG_{rich}) and megabrowser (MB_{rich}) species richness, respectively (mixed feeders excluded). Grey areas in H are ecoregions lacking data due to palm sensitivity to frost. prop.: proportion; sp: species.

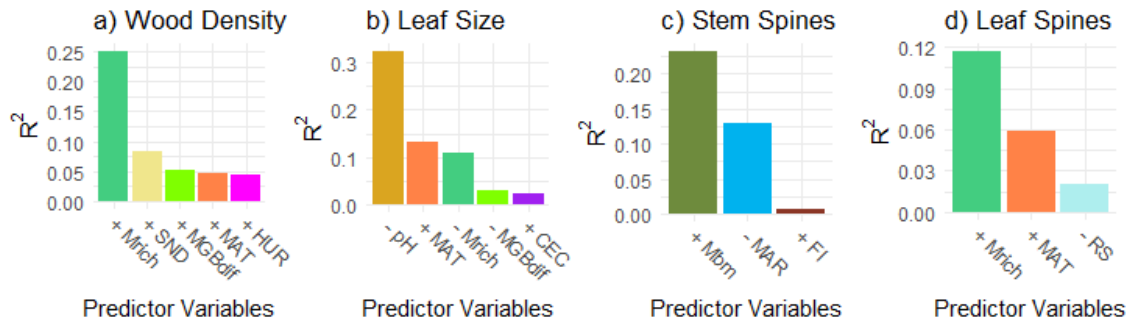


Fig. 2: Average contribution of predictor variables to the regression models of pant antiherbivory defence traits (a-d) (see Tables S1-S4 for details). Positive and negative signs preceding predictor variable abbreviations indicate positive and negative effects. Results were based on R^2_{adj} for (a) and (b) (multiple regressions), and McFadden's pseudo- R^2 for (c) and (d) (generalised linear models with binomial errors). Islands were excluded from the analyses. Mrich: mean extinct megafauna species richness; Mbm: mean extinct megafauna species mean body mass; MGBdif: richness difference between grazer and browser megafauna species (mixed feeders excluded); CEC: soil cation exchange capacity; FI: (wild) fire intensity; INS: insularity (categorical: 0/1); MAR: mean annual rainfall; MAT: mean annual temperature; pH: soil pH; RS: rainfall seasonality; SND: soil sand percentage; HUR: hurricanes.

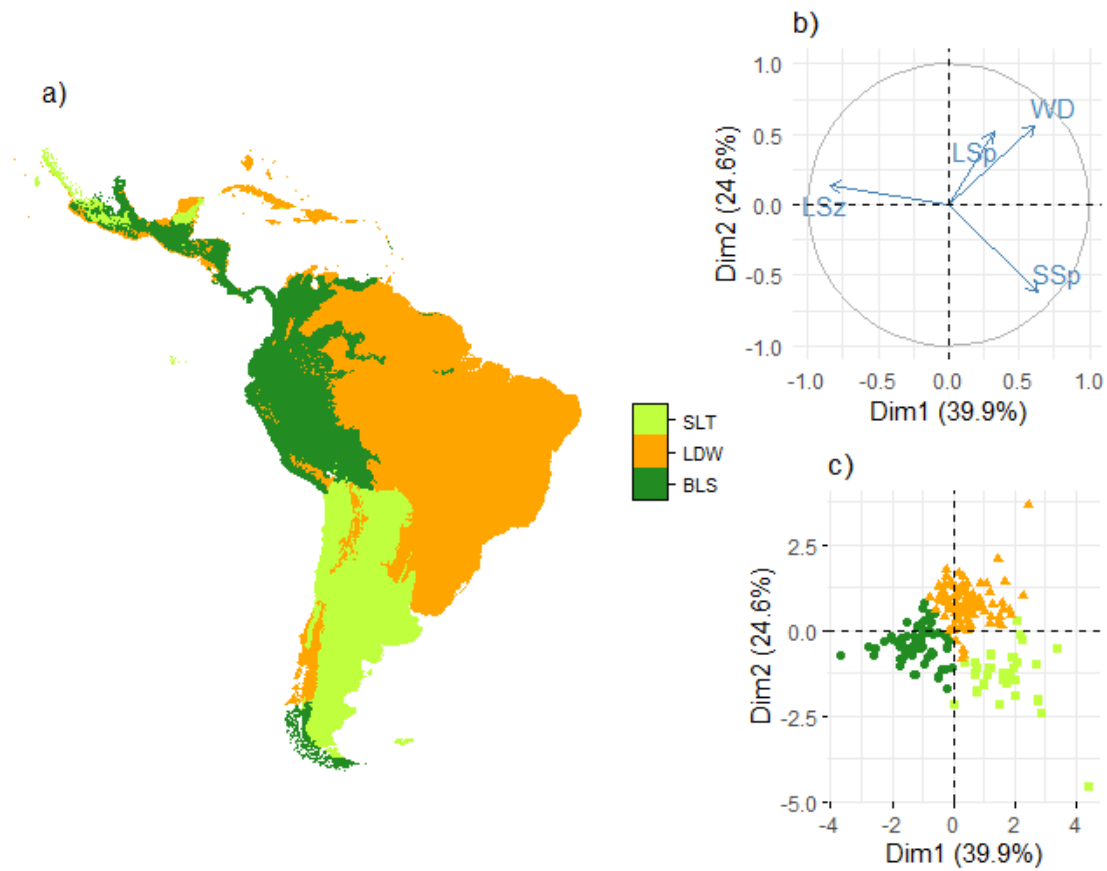


Fig. 3: Geographical distribution (a) and functional trait characterisation (b,c) of Neotropical antiherbivores. Antiherbivores were classified using hierarchical clustering on principal component axes of mean values of antiherbivory defence trait per ecoregion. BLS: broad-leaved sensitive; SLT: small-leaved thorny; LDW: leaf defences/dense wood; LSz: leaf size; SSs: stem spines; WD: wood density; LSp: leaf spines

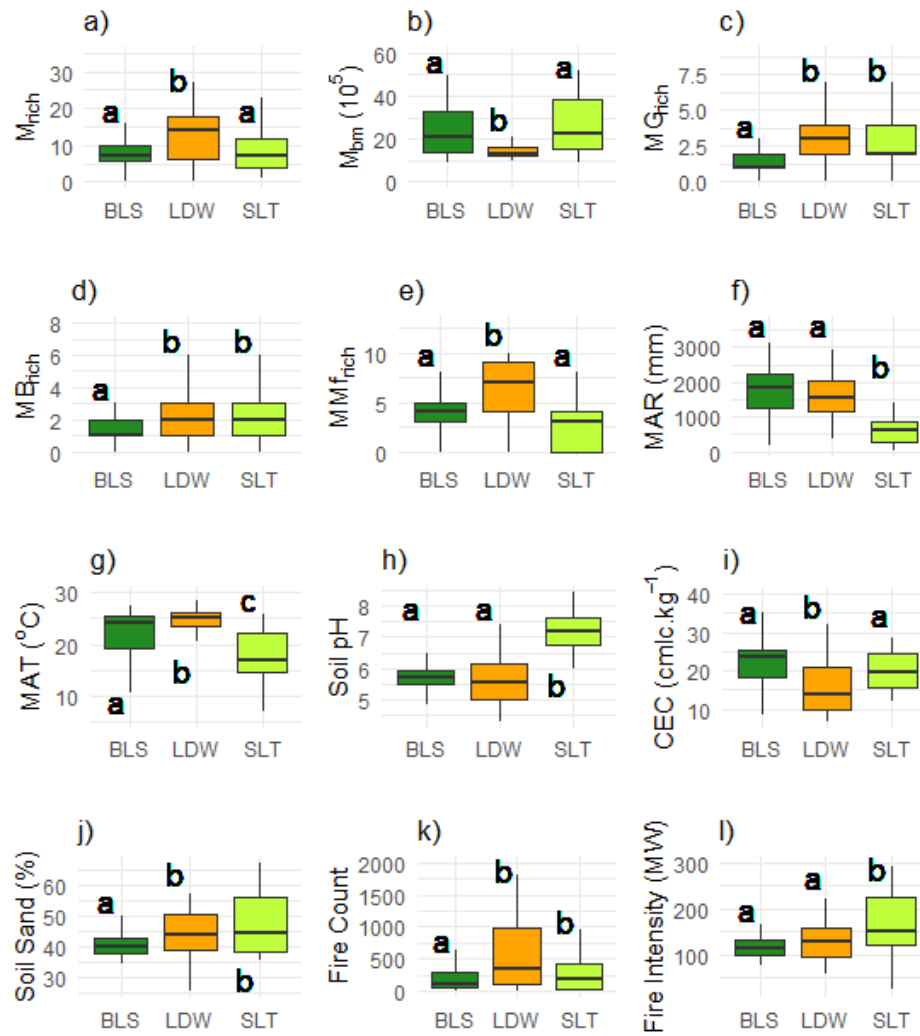


Fig. 4: Megafauna (a-e), environmental (f-j), and fire regime (k,l) differences among Neotropical antitherbiomes for continental ecoregions. Different letters indicate significant differences among antitherbiomes (P-values corrected using the Benjamini & Hochberg method). Only variables with significant differences among antitherbiomes are shown. M_{rich} : extinct megafauna species richness; M_{bm} : extinct megafauna species mean body mass; MG_{rich} : megagrazer richness; MB_{rich} : megabrowser richness; MMf_{rich} : Megamixed feeder richness; MAR: mean annual rainfall; MAT: mean annual temperature; CEC: soil cation exchange capacity.

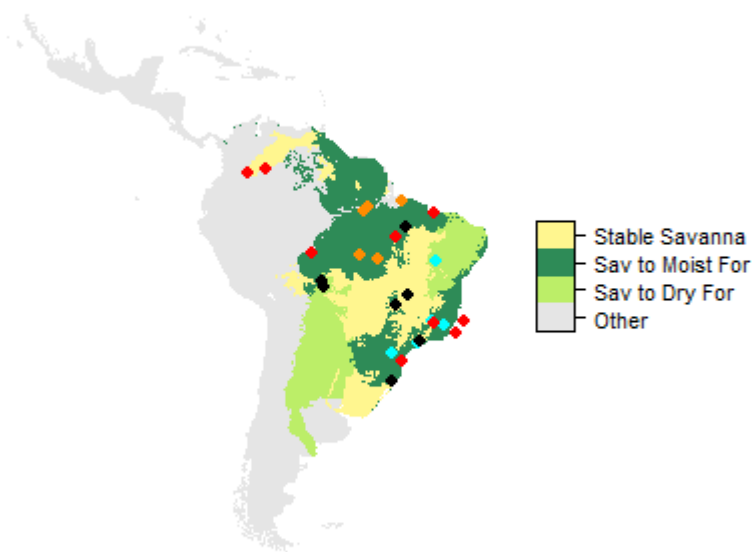


Fig. 5: Regions dominated by grassy ecosystems during the Pleistocene (coloured area), and fossil sites with evidence of grass-dominate ecosystems during the same period (diamonds; references in methods). Pleistocene grassy ecoregions are those in which a high richness of megafauna species, in general, and large grazers, in particular, used to roam (richness equal or greater to the 75% quantile), and in which plant assemblages were classified in the SLT or LDW antiherbiomes (see Fig. 3). This included ecoregions that are currently dominated by savannas (Sav; yellow area), dry forests and woodlands (Dry For; light green area) and moist forests (Moist For; dark green area). The megafauna- and large grazer- poorer Llanos savanna is also shown (northern yellow patch). Fossil sites supported grassy vegetation during the Last Glacial Maximum (red), Mid-Holocene (cyan) or during both periods (black diamonds). Only five out of the 21 fossil sites were in currently non-forest ecosystems (the four northwesternmost red points, and the leftmost black point in the central Cerrado ecoregion – the large central

yellow patch). Present Amazonian savannas are shown as salmon diamonds (reference in methods).