Corrigendum: Megafaunal isotopes reveal role of increased moisture on rangeland during late Pleistocene extinctions

M. Timothy Rabanus-Wallace, Matthew J. Wooller, Grant D. Zazula, Elen Shute, A. Hope Jahren, Pavel Kosintsev, James A. Burns, James Breen, Bastien Llamas and Alan Cooper


The Supplementary Information files that were originally published with the Article were not the intended final files. Both files have been updated.
Megafaunal isotopes reveal role of increased moisture on rangeland during late Pleistocene extinctions

M. Timothy Rabanus-Wallace1*, Matthew J. Wooller2,3, Grant D. Zazula4, Elen Shute1,5, A. Hope Jahren6, Pavel Kosintsev7, James A. Burns8, James Breen1,9, Bastien Llamas1 and Alan Cooper1*
Supplementary: Methods

SM Methods 1: Dataset

SM Figure 1 presents the $\delta^{15}$N measurements used in this study along with global region and source. New measurements (n = 63) are credited to ACAD (Australian Centre for Ancient DNA). Other data sources are detailed in SM References and SM Bibliography. Uncalibrated dates were calibrated using the calibration curve IntCal13 for all northern hemisphere data, and SHCal13 for South America. The calibrations were performed using the R package Bchron and represented as a single value by taking the median of the probability distribution over time (https://cran.r-project.org/package=Bchron; SM Code). Previously published calibrated dates were used from the studies indicated in SM Methods 1. Indirectly dated samples (n = 39 Bison and Equus samples from Drucker et al., 2003, and n = 5 sediment samples from Haile et al., 2009; see SM Data) were only used when multiple $^{14}$C dates from a dated sedimentary layer all fell within 750 yrs of one and another, and calibration was then performed using the mean of the $^{14}$C dates and the highest of the 95% error margins. All samples dating to between 45 and 1 ka (thousand years before present) were included in the main study and the SM investigations.
Samples were grouped by genus. *Bison* and *Equus* samples are abundant in the fossil record and their dietary preference for graminoids reduces the potential effects of diet selection on δ¹⁵N values [1-4]. *Equus* includes specimens identified as *E. caballus* and *E. ferus*, while *Bison* includes *B. bison*, *B. bonasus*, *B. antiquus*, *B. priscus*, along with samples identified only to genus level. *Lama* comprises *L. gracilis* and *L. guanicoe*. For the SM analyses, *Mammuthus* includes only specimens of *M. primigenius* while the carnivorous felids are described in the text accompanying SM Figure 11. Similarly, sediment δ¹⁵N values are given in SM Figures and Investigations (Figure 10).

The samples used in the main study (*Equus*, *Bison*, and *Lama*) were then grouped into one of the following bins: European, Siberian (including all Eurasian samples east of Central Europe), North American, or South American (see SM Figure 1). The North Slope of Alaska was separated from the remainder of the North American continent since ecological continuity between the two is broken by the Brooks Range, and the fossil-rich North Slope provides adequate data to independently characterise trends in δ¹⁵N values.

**SM Methods 2: Data handling and the RMAC algorithm: motivation, description, implementation**

We developed a novel approach to combine all available data in each region to yield a proxy for the relative changes in Moisture Availability (MA) over time, based on grazer collagen δ¹⁵N values. This method produces a Relative Moisture Availability Curve (RMAC) based on five steps: normalisation, exclusion of data-poor regions, smoothing, transformation, and scaling. We aimed to estimate changes in relative moisture in a way that accounted appropriately for the reliability of the data at different points in time. To achieve this we established four guidelines: 1) In periods where data were too sparse to adequately represent δ¹⁵N variation then no estimate of moisture should be made; 2) for any given time point, relative MA should only be estimated if data occur nearby both before and after that time point; 3) at time points where the data are dense, it is more reasonable to reconstruct relative MA based upon the abundant data immediately surrounding the time point. Where data become sparse, the fairest estimation of relative MA should be more heavily influenced by more distant data. 4) Some measure
of the variability of the relative moisture estimate at different points in time should be included. Guidelines 1–3 are illustrated in SM Figure 2;

To make data from different taxa comparable, δ^{15}N values are normalised separately for each genus and the resulting standard scores (Z-scores) are then combined to give a dataset \( \bar{x} = (x_1, x_2, \ldots, x_n) \), with \( n \) data points, mean equal to 0, and standard deviation equal to 1. This step is intended to eliminate the differences in mean δ^{15}N values, as well as heteroscedasticity between regions and taxa due to idiosyncracies in diet, habitat, and physiology. Each data point in \( \bar{x} \) is associated with a date, listed in \( \bar{t} = (t_1, t_2, \ldots, t_n) \). The RMAC reconstruction is based upon the central tendency of \( \bar{x} \) over time, in accordance with the properties listed above. The smoothing step, as illustrated below, accounts for guidelines 1–3 (see also SM Methods 3).

![SM Figure 2. Schematic illustration of the smoothing method used in the RMAC. Blue dots: data. Black line: the smoothed average that is later transformed to give the RMAC. Blue curves: weighting functions at selected time points, illustrating the relationship between the variance of the weighting function and the local density of data points.](image)

In practice, the smoothed value at a time point is given by the function \( s(T) \) for each time point \( T \). To account for guidelines 1 and 2, \( s \) is undefined at any point in time where fewer than \( m \) entries in \( \bar{t} \) fall within the range \([T - r, nT + r]\), or when no entries in \( \bar{t} \) fall within either \([T - r, T]\) or \([T, T + r]\), for some constant \( r \). A flag variable \( f \) is
given the value 1 when these conditions are met, and 0 otherwise. Where these conditions are not met, \( s(T) \) is given as the weighted mean of the surrounding datapoints, with the contribution of each point weighted using the Gaussian probability density function \( G(x, \mu, \sigma) \). In practice, a Gaussian kernel smoother:

\[
ns(T) = \begin{cases} 
\text{undefined if } f = 1, & \text{and} \\
\frac{\sum_{i=1}^{n} G(T - t_i, 0, \mu_T) x_i}{\sum_{i=1}^{n} G(T - t_i, 0, \mu_T)} & \text{otherwise.}
\end{cases}
\]

The variable \( \mu_T \), the standard deviation of the weighting function, is included to account for guideline 3, declining in magnitude as data points become more sparse in the vicinity of \( T \) (see diagram above). To set \( \mu \) appropriately at each time \( T \), we require some measure of the density of data in the vicinity of \( T \). We use the score \( d_T = \sum_{i=1}^{n} G(T - t_i, 0, a) \), which is proportional to the Gaussian kernel density at \( T \) with bandwidth \( a \). The standard deviation of the weighting function, \( \mu_T \), is then assigned using the relationship \( u_T = be^{-cd_T} \), which allows for the standard deviation to approach zero as the density increases (see also SM Methods 3). The parameters \( b \) and \( c \) influence how the density affects the weighting function. The parameters used in this study are \( na = 1500, nb = 2000, c = 100, r = 3800, m = 3 \), with the exception of Europe where \( b \) is set to 10000, to prevent overfitting caused by the high density of datapoints in this group.

Smoothed values \( s(T) \) are the basis of the lines shown over the data displayed in the inset graphs in Main Text Figure 1.

The deterministic component of the relationship between measures of Moisture Availability (\( MA \)) and \( \delta^{15}N \) values in arid and semi-arid environments (annual precipitation < ~500–800 mm; see SM References 5 and SM Figure 14) has been described as \( \delta^{15}N \sim \frac{y}{MA} + z \), where \( y \) and \( z \) are constants specific to the system under study [5]. Since there is little such information about the paleoecosystems or species in this study, and we are pooling normalised data from multiple different systems (e.g.
taxa and regions), we have assigned the relative $MA$ vs. $\delta^{15}N$ relationship to provide a reasonable reconstruction of the relative moisture changes over time. To make our $MA$ vs. $\delta^{15}N$ curve compatible with the results of empirical studies (see SM References 5) we ensured that, at the upper range of values, a change in $\delta^{15}N$ values corresponds to a change in $MA$ perhaps five to ten times smaller than a similar change at the lower end of the range. Given the variation in published studies across different systems this approach seems a reasonable compromise until appropriate empirical data are gathered. The transformation is effected using the function:

$$R(T) = \frac{1}{(s(T) + \min_{T=0}^{45,000} (s(T)) + 1)}.$$ 

Terms added to $s(T)$ in the denominator upshift the data to avoid undefined values of $R$. The output of $R$ is in arbitrary units, best interpreted as approximate changes in the relative availability of moisture over time, allowing the values to be scaled for plotting and comparison to other proxies (e.g. Main Text Figure 1). For each study group, we took 1000 bootstrap samples from the data and for each evaluated $R$ at 200 points spaced equally along the range of $[0,45000]$. At each point, a bootstrap 95th percentile confidence interval is calculated and plotted to yield the RMACs shown in Main Text Figure 1.

The RMAC reconstruction was implemented in R and the package ggplot2 (see SM Code). Deglaciation proxies were reproduced directly from the source publications (listed in SM References 2). Peatland initiation was represented by applying R’s density function with bandwidth 1000 on peatland initiation dates grouped by region (North Slope $n = 70$, other North America $n = 919$, Siberia $n = 363$, Europe $n = 518$, South America $n = 54$; sources listed in SM References 3). The peatland initiation dates were calibrated using the calibration curves and methods as described for the $\delta^{15}N$ values (SM Methods 1 and SM Code).

**SM Methods 3: The effects of parameters $r$, $a$, $b$, $c$, $m$, on the RMAC**
To illustrate how the parameters of the RMAC reflect the reconstruction of relative moisture, we performed demonstrations of the smoothing step of the algorithm using a subset of the North American data.

In the current study, the parameters \((a,b,c,r,m)\) were set as \((1500,1500,100,3800,3)\). With these parameters, the smoothed values appear as follows:

**SM Figure 3.** The RMAC smoother plotted over North American data and the parameters \((a,b,c,r,m) = (1500,1500,100,3800,3)\). The marked decrease in relative \(\delta^{15}N\) characterising the Late Glacial Nitrogen Excursion is evident at 12–15 ka. **Point Colours.** Mammoth: cyan; Bison: red; Horse: green; Sediment: purple.

The parameters \(r\) and \(m\) control the requirements for plotting. With \(r = 1000\) and \(m = 6\), the RMAC only plots at time points where 6 or more data points fall within 1000 yrs of the time point, with at least one data point falling on each side:

**SM Figure 4.** The RMAC smoother plotted over North American data, with parameters as per SM Figure 3, but with \(r = 1000\) and \(m = 6\), causing the curve to be undefined where data are sparse. **Point Colours.** Follows SM Figure 3.
The parameter $a$ controls how the density of data is measured at each point in time. A low value means the density is estimated based largely on the density of data very close to the estimation point, while a higher value means the estimate gives more weight to datapoints further from the estimation point. Ideally, the density score will change quickly where an obvious change in density occurs, such as before and after $\sim 27$ ka in the examples here (SM Figures 3–8). The effect of altering $a$ can be observed by plotting the normalised density scores, to assess intuitively how well the height of the line reflects changes in the density of data.

Pale: Normalised density with $a = 5000$
Medium: Normalised density with $a = 1500$
Dark: Normalised density with $a = 500$

**SM Figure 5.** The effect of parameter $a$ upon RMAC density scores corresponding to North American data, showing how low values of $a$ produce density scores that are more sensitive to changes in the density of data. Parameters besides $a$ as per SM Figure 3. **Line Colours.** Dark: $a = 500$; Medium: $a = 1500$; Pale: $a = 5000$. **Point Colours.** Follows SM Figure 3.

The density score $d$ is used to set the standard deviations of the weighting distributions for the smoother. As shown in SM Figure 2, the standard deviation is small where the data are dense, and large where the data are sparse. The relationship between the kernel density and the standard deviation is controlled by the parameters $b$ and $c$: 
Red: Normalised Density (with $a = 1500$)
Pale: Standard deviation with $b = 1500$
Medium: Standard deviation with $b = 1000$
Dark: Standard deviation with $b = 200$

**SM Figure 6.** The effect of parameter $b$ upon the relationship between RMAC density scores, and the SD of the weighting function using the North American dataset. Parameters besides $b$ follow SM Figure 3. **Red Line.** RMAC density score. **Blue Lines.** SD of weighting distribution. **Blue Line Colours.** Dark: $b = 200$; Medium: $b = 1000$; Pale: $b = 1500$. 

Red: Normalised Density
Pale: Standard deviation with $c = 50$
Medium: Standard deviation with $c = 250$
Dark: Standard deviation with $c = 500$
SM Figure 7. The effect of parameter $c$ upon the relationship between RMAC density scores, and the SD of the weighting function using the North American dataset. Parameters besides $c$ follow SM Figure 3. **Red Line.** RMAC density score. **Blue Lines.** SD of weighting distribution. **Blue Line Colours.** Dark: $c = 500$; Medium: $c = 250$; Pale: $c = 50$.

To demonstrate the effect of density-dependent weighting, the results of the smoothing step can be compared with and without this feature. The red curve in SM Figure 8 represents the RMAC smoother with density-dependent weighting ($b = 3000, c = 300$). The blue curves are density-independent ($c = 0$) with narrower (dark blue; $b = 1000$) or broader (light blue; $b = 2000$) weighting functions. Note that the light blue curve performs well in the low-density regions <27 ka, but underfits in the dense regions around 14–5 ka. The dark blue curve performs well in this high-density region but overfits in the lower-density areas. The red RMAC smoother appears to capture the central trends well in both regions:

SM Figure 8. A comparison of the performance of the RMAC smoother with and without density-dependent weighting, using North American data, showing how density-dependent weighting prevents the under- and
overfitting of the smoother without this feature. **Line Colours.** Red: Smoother with density-dependent weighting, \((b = 3000, c = 300)\); Dark blue: Smoother without density-dependent weighting, with parameters \((b = 1000, c = 0)\); Light blue: Smoother without density-dependent weighting, with parameters \((b = 2000, c = 0)\). **Point Colours.** Follows SM Figure 3.
Supplementary: Text

SM Text 1: Proxy choice and sources of variation in the data

Since the current study compares RMACs derived from data spanning broad geographical ranges with proxies from a similarly broad range, the results should be considered an integration over continental/regional-scale trends. Similarly, the trends are generalised since they rely on a geographically and temporally heterogeneous sample of rangeland fauna. However, since the variable driving the signal is argued to be moisture, and moisture regimes act over broad spatial scales, some generalisation ought to be possible. For example, the volume of δ¹⁵N data appears adequate to demonstrate that the Late Glacial Nitrogen-15 Excursion (LGNE) falls well outside the range of pre-LGM (Late Glacial Maximum) variation. The magnitude of the δ¹⁵N shift and the correlation between different regions and taxa suggest a biological event associated with the climatic and environmental events of the Last Glacial-Interglacial Transition (LGIT; Main Text Fig. 2).

The RMAC reconstructions are contrasted with deglaciation proxies for environmental change in Main Text Figure 1. While the mass and area of ice sheets is a product of complicated dynamic inputs (precipitation, condensation, surging) and outputs (melting, sublimation, calving), glaciation was favoured as a proxy for large scale climate-driven changes at the landscape level because factors causing an ice sheet to retreat are also likely to be significant to other systems. Furthermore, the retreat of ice sheets is itself expected to strongly influence the surrounding environments via mechanisms including isostatic/eustatic sea-level change, the discharge of meltwater, altered wind and thermohaline circulation, and the retreat of periglacial environments.

Peatland formation is also plotted in Main Text Figure 1 as a generalised proxy for environmental change at continental scales over time due to the extensive datasets available for peatland initiation over a broad spread of locations within the study regions, each represented by a single date (rather than a series of abundance measures over time). The discussion in the Main Text assumes that grazers preferentially aggregated in flourishing grasslands where moisture accumulated and drove down the
$\delta^{15}N$ values of graze. This would specifically require prolonged increased moisture: in arid grasslands, the immediate effect of increased moisture (for instance, the anomalous rainy season reported by Arranibar, Otter et al. in 2004 [6]) can in fact increase foliar $\delta^{15}N$ values in the short term. Therefore, $\delta^{15}N$ values only normally decrease along with averaged long-term (annual/decadal) moisture levels. Peatland formation has the same requirement for persistent moisture in localised areas, just in a larger degree and for a longer time [7-9].

The $\delta^{15}N$ values of natural samples can be influenced by a number of biological, geological, and chemical processes. For instance, salinity, slope, pH, temperature, and depth have all been linked to soil and plant $\delta^{15}N$ values [10-13]. Moisture levels have predictable effects on several aspects of the nitrogen cycle. It has been suggested that generally dry environments have proportionally more inorganic nitrogen, which is more amenable to $^{15}N$ enrichment via evaporation, while wet environments have more organic nitrogen, which may suffer greater $^{15}N$ depletion via leaching [14]. Such interactions ultimately lead to the strong statistical $\delta^{15}N$-moisture correlations that have been reported from field studies in arid and semi-arid regions (see SM References 5). The abundance of influences on $\delta^{15}N$ values can complicate the interpretation of palaeo-isotopic data, however it is likely that moisture exerted the primary influence on the assembled dataset as has been observed in modern global studies of ecosystem $\delta^{15}N$ values (SM References 5).

The LGIT warming trend is known to have produced landscape moisture through several means. Increased precipitation in many regions was caused by changes in global wind currents and thermohaline circulation, with storm tracks and low pressure regions moving moisture inland in Europe [15], Siberia [16], and North America [17-19], and the Inter-Tropical Convergence Zone shifting southward over South America [20]. Higher temperatures increased evaporation, and rising sea levels brought the influence of maritime climate regimes further inland. Direct evidence for enhanced LGNE moisture input are provided by LGIT records of stream incision detected in the study regions of Europe [21, 22], the North Slope [23], Alaska [24], Siberia [25, 26], and South America [27, 28]. This is reinforced by changes in certain plant communities, which are more directly linked to the soil moisture that exerts strong control over
grazer $\delta^{15}N$ values. LGIT pollen profiles collected from across the study regions show transitions of steppe vegetation (graminoid/forb dominance) to forest with early-establishing trees (often deciduous shrubs or small trees), being replaced with larger forest evergreens. North of the treeline the transition is to tundra vegetation, and where the water balance allows prolonged accumulation, peatlands form [20, 29-32] (see also SM References 3). Within these robust general relationships, the correlation between local proxies to the soil moisture that influences graminivore collagen $\delta^{15}N$ values are usually indirect, and heterogeneous within regions with respect to both timing and extent.

**SM Text 2: Carbon stable isotopes**

For the collagen samples analysed for $\delta^{15}N$ values, we also compared $\delta^{13}C$ values with data from previous studies (see SM References 1). The $\delta^{13}C$ values remain relatively stable during the LGNE, and the most striking changes are elevated post-LGNE $\delta^{13}C$ values in North American bison between ~10 and 6 ka, and a slight decrease in Holocene European horses (< 10 ka). These excursions are in opposite directions and are therefore likely to be controlled by variables acting at local scales [33]. The North American $\delta^{13}C$ increase is first observable in horses, but becomes most apparent in bison south of the ice sheets (see SM Figures 1), and is most likely the result of changing proportions of C3 vs. C4 plants in southern areas [34], though increasing atmospheric CO2 may also have played a part [35]. The slight decrease in post 10 ka European horse $\delta^{13}C$ values (~2 per mil) could reflect a decreased need for Water Use Efficiency (WUE) by C3 plants, changes in atmospheric CO$_2$, or an increasing canopy effect [36]. Foliar $\delta^{13}C$ values also responds to changes in WUE in plants, which might well be expected to alter with changes in moisture.

While $\delta^{15}N$ values are most sensitive under arid conditions, $\delta^{13}C$ values are more affected by moist environments and remain far less sensitive to changes in moisture under comparatively dry regimes such as those known to sustain extensive rangelands in the Pleistocene and today [37]. However, the lack of change in Siberian and North American $\delta^{13}C$ values during the LGNE indicates that forage composition remained...
consistent with a primarily C3, graminoid diet and rules out major shifts in forage, e.g. in favour of herbs [38]. Current palaeoecological evidence also indicates that C4 plants did not expand into Beringia in the Late Pleistocene (e.g. [39, 40]).

Palaeo-isotope studies that investigate food webs generally assume that $\delta^{15}N$ in a species at a given time is stabilised by that species’ tendency to consume the same kinds of forage across broad geographic ranges [41]. Tooth wear, isotopic studies of plant type, and the diets of modern analogues have been used to argue for this stabilisation [3, 41]. Overall then, a purely diet-based explanation for the LGNE pattern seems unlikely, and would need to posit a near-simultaneous transition of multiple grazing taxa on different continents to an alternative food source that affected $\delta^{15}N$ in the same manner, and in all cases was not reflected in $\delta^{13}C$ values. This is unlikely, although moderate changes in the ratio of plant types consumed is possible since moisture is likely to encourage the rangeland flora to alter in similar ways, for instance, by increasing the proportion of $^{15}N$-depleted forbs and pioneer rudderals in the diet.

**SM Text 3: Megafaunal community collapse**

It is important to emphasise why the Pleistocene rangelands were critical in the structure of the megafaunal mammal communities. A key question about this environment is the so-called *productivity paradox* which asks how an environment that was presumed to be largely tundra was able to support such a large biomass of mammals. Guthrie [3] has suggested that rangelands were in fact some of the most productive environments, with a warm growing season thaw providing ample moisture for deep-rooted grasses, while clear skies and long days allowed for fast growth and uptake of mobile nutrients. During the cold season, dead superterranean foliage was rapidly composted, and the nutrients redistributed by megafauna. Cold-arid grasslands may also favour large grazers in particular: large-bodied endothermic animals, having a large surface area-to-volume ratios and thermal inertia, may incur a selective benefit in cold environments. Large bodies also allow large digestive systems for the extraction of energy from sizeable amounts of low-quality forage such as the graminoids that dominate arid regions. Also, large bodies may allow longer survival when resources are
scarce. The selective advantage conferred by the latter two factors may be amplified by strong seasonal shifts in polar regions [3].

The main text includes only cursory discussion of the mechanisms by which altered moisture regimes may lead to enhanced extinction pressure on rangeland megafauna, focusing on habitat loss as climatic variables alter the environment. Previous work has explored the interdependence of megafauna and vegetation, along with other factors such as fire, and offered insights into mechanisms that directly link megafaunal extinction to moisture via the nitrogen cycle [42-45]. For instance, increased moisture (alongside temperature and atmospheric CO₂) may cause nitrogen to become a limiting factor in the nutrient cycle. Herbivore behaviour and physiology changes significantly to compensate for inadequate foliar nutrition [45], and by retaining more nitrogen and focusing deliberately on locations containing high-nitrogen forage, the redistribution of nutrients by herbivores is impaired. This can further exacerbate the nitrogen shortage creating a feedback cycle [42].

**SM Text 4: Non-graminivore herbivore δ¹⁵N values.**

Published studies [46-49] have demonstrated that browsing deer probably responded to some influences not affecting the obligate browsers used to investigate the rangeland habitats in this study. A gradual increase in δ¹⁵N values for North Slope caribou (*Rangifer tarandus*) has been noted from ~40 ka into the LGM [48], which is not obviously reflected in grazer values from the region. While the facultative consumption of grass, and some shared influences of moisture on both graze and browse probably explains the late glacial depletion (and, in the European case of both caribou and red deer, *Cervus elephas*, subsequent recovery), the North Slope caribou record from Mann *et. al.* (2013) [1] possibly reflects cooling and a decrease in moisture availability leading to the LGM, which, based on horse, bison, and mammoth data from the same study, appears to have influenced the ecology of rangelands less than the biomes browsed by caribou. The pre-LGM increase may therefore reflect a greater reliance by deer on graminoid material in cold and dry times. Most importantly, however, the decoupling of the North Slope caribou and graminivore records demonstrates that while ecological
changes were occurring before the LGIT, they did not significantly impact the entire herbivore community until the late glacial moisture increases [48].

**SM Text 5: South American rangelands were analogous to those on the mammoth steppe**

Landscape-scale interpretations of late-glacial environmental change predict that a similar decrease in δ¹⁵N values should be found wherever arid rangelands grazed by mega-herbivores were subject to a sustained increase in plant-available moisture. Southern South American rangelands, and their graminivore [50] community, make a good analogue for the mammoth steppe, sharing similar histories with respect to changes in the environments and the extinctions/range shifts of the grazers [28, 51, 52]. Following LGIT population decline and range contraction, *Lama gracilis* became confined to highland steppes in the central west, while *L. guanicoe* populations also underwent range shifts during the glacial period [53], and occupied similar steppe-like habitats along the central and southern west coast [54]. Pleistocene rangelands were replaced by various biomes including woodland, wetland, forest, and desert, and the permafrost that underlay much of the continent’s southern half during the LGM is now only found at high altitudes along the western cordillera [28, 55]. Previous studies have shown South American camelid δ¹⁵N values correlate with precipitation and to vary over time [56, 57]. Overall then, the llama isotopic signals record an analogous pattern to that seen for horse, bison, and mammoth across the holarctic rangelands, suggesting a similar process of moisture change associated with megafaunal extinctions.

**Supplementary: Figures and Investigations**

SM Figure 9 displays all herbivore δ¹⁵N data compiled for the study prior to normalisation. Mammoth δ¹⁵N values were not used in the main text RMAC reconstruction, owing to their low density and poor temporal coverage. While the RMAC is changed very little by inclusion of these values, the normalisation step has the effect of over-emphasising noise in the data when few data points are available. Where data points overlap with the LGNE, the mammoth δ¹⁵N values do indeed suggest a decline, which is expected given that digestive physiology, dental morphology, modern analogues [4], ancient DNA [58], and stable isotope analysis [1] indicate that mammoth
were rangeland graminivores that subsisted primarily on large quantities of low-quality graminoid forage complemented by forbs.

**SM Figure 9.** Complete unnormalised herbivore $\delta^{15}$N data compiled for the study, showing varying onset of LGNE pattern in different regions. Total sample numbers for each region are given in each panel. North American samples are separated into those falling in Eastern Beringia (Alaska, USA, and Yukon Territory, Canada; $n = 92$), and south of the North American ice sheet complex (Alberta and British Columbia in Canada; Minnesota, Nevada, and Wyoming in the USA; $n = 62$). Siberian samples are separated into those from the Ural Mountains ($n = 33$) and other regions (including the Russian Plains, Taymyr, and Western Beringia; $n = 56$). South American llamas are separated by...
species into *Lama gracilis* (n = 15) and *L. guanicoe* (n = 30). Mammoth δ¹⁵N values from the literature (n = 93; See SM Data and SM Bibliography) are also shown.

Modern field studies show that the moisture-δ¹⁵N relationship is detectable at trophic levels from soil upwards. Our moisture-based interpretation of δ¹⁵N changes therefore predicts that the LGNE will be reflected in soil δ¹⁵N values. In addition to data presented in the main text, we collated δ¹⁵N measurements from permafrost sediment samples (n = 5) from a river embankment at Steven’s Village, Alaska taken from two points within a transect reported by Haile et al., 2009 [59] (SM Figure 10). The age of the samples is estimated to be ca. 10.8 and 10.2 ka respectively, based upon the radiocarbon dates made along the transect. While these few soil measurements represent only a preliminary investigation, the low δ¹⁵N values match the faunal signals, and are consistent with the return phase of the LGNE. While foliar nitrogen may show a depletion in ¹⁵N compared with soil [60, 61], herbivore δ¹⁵N values are typically ~3–4‰ higher than the forage they consume. The soil δ¹⁵N values appear to fall in a range that would be unexpectedly low if they had not experienced LGIT depletion.

**SM Figure 10.** Sediment δ¹⁵N data (n = 4) from Steven’s Village, Alaska North America juxtaposed against herbivore data from the same region, showing that values consistent with the LGNE can be detected in soils dating from the same period.
SM Figure 11 shows the raw (unnormalised) South American camelid $\delta^{15}$N data, and a similar, but elevated, pattern of rapid increase in a set of carnivores from the same area. It has been shown by Bump et al. [33] that organisms at high trophic levels better integrate stable isotope signatures over the surrounding area. Hence the carnivore $\delta^{15}$N values indicate that a large proportion of the herbivores consumed by these felids (including taxa other than llamas) exhibited the same general trend, reinforcing the view that the effects of the late glacial moisture increase were large enough to affect the megafaunal community beyond just grazers. Compared to the llamas, the carnivores show $\delta^{15}$N values 3–4‰ higher, consistent with the $\delta^{15}$N value increase that typically occurs between predators and their prey.

**SM Figure 11:** South American unnormalised $\delta^{15}$N data including carnivores, showing a rapid increase in values from ~16 ka consistent with the LGNE recovery phase, suggesting a major moisture spike around the rapid deglaciation of ~17–16 ka. The X-axis has been limited to between ~15 and 17.5 ka, to emphasise the strong upward $\delta^{15}$N trend at this time. Data from South American felid carnivores ($n = 13$; *Puma concolor*, *Smilodon populator*, and *Panthera onca*) demonstrate that the $\delta^{15}$N signals were transmitted to the upper trophic levels of the food chain.
SM Figure 12: δ¹³C values for original data used in this study showing the overall lack of significant changes during the LGNE period. The causes and implications of δ¹³C changes over time (in particular those that occur after the LGNE, such as in North America), is discussed in SM Text 2.
The tests presented in SM Figures 13–16 aim to investigate the statistical significance of changes in Z-normalised $\delta^{15}N$ values. Data were grouped into points falling within a particular window, and those falling outside. For each window, a Student’s t-test (two-tailed) was used to generate a p-value on the null hypothesis that the $\delta^{15}N$ values within the window were drawn from the same distribution as the $\delta^{15}N$ values outside the window. Each field in the plot represents a window, with the colour corresponding to the p-value. For instance, the darkest field in the European example below (SM Figure 13) represents a window beginning at 11 ka, and lasting for 3 ky. This indicates that the period 11–14 kya contained $\delta^{15}N$ values that were most significantly different from those before and after this window ($p \leq 1E-50$ in this case).

Diagonal striping is a consequence of $\delta^{15}N$ excursions: long windows containing the excursion may still have elevated significance despite including some datapoints from the regular (non-excursion) range of values. This is especially prevalent where the excursion abuts a period of sparse or absent data, for instance, after 10 ka in the North Slope (see SM Figure 9). Fields are coloured grey when fewer than 5 datapoints exist within the window. This test is not informative on the South America dataset, which shows a strong, single trend in a very short period.

**SM Figure 13.** Student's t-tests for significant variation in z-normalised $\delta^{15}N$ values in Europe, showing highly significant $\delta^{15}N$ excursions around 11–14 ka.
**SM Figure 14.** Student's t-tests for significant $\delta^{15}$N variation in North America showing highly significant $\delta^{15}$N excursions around 10–17 ka.

**SM Figure 15.** Student's t-tests for significant $\delta^{15}$N variation on the Alaskan North Slope, showing a highly significant $\delta^{15}$N excursion at around 10 ka. Note the diagonal striping caused by the lack of data after 10 ka.
SM Figure 16. Student’s t-tests for significant δ^{15}N variation in Siberia, showing a significant δ^{15}N excursion around 13–20 ka.

At large geographic scales, variation in modern δ^{15}N values has been investigated with respect to many inter-correlated variables, such as moisture, temperature, altitude, and latitude (See SM Text 1). To justify the assumption that the relationship between δ^{15}N values and moisture holds between rangeland environments in different continents we compare a modern δ^{15}N value dataset from Australia with another from Alaska, demonstrating the consistency between the results.

SM Figures 17. Investigation of the global applicability of the moisture-δ^{15}N relationship, using modern grass data (n = 35) collected in Alaska [62], which is one of our study areas, and Australia [5]. Error bars associated with the Alaska
\( \delta^{15}N \) values = 1 standard deviation. These two geographically disparate regions show compatible results, with respect to both the direction of the relationship and the raw \( \delta^{15}N \) values at certain moisture levels (approximated here using mean annual precipitation).

We chose a subset of the data to investigate whether any source of variation in the data could be easily explained by other factors besides moisture (SM Figures 18–20). Bison from North America (Eastern Beringia, Canada, and Minnesota; \( n = 17 \)) were chosen from the post-LGNE “stable” period where some spread in the data are still evident (5–10 ka). It was assumed that post-LGNE climate might bear better correlation with modern climate than data dating earlier, and only inland sites were considered to remove the major effects of marine transgression and coastal climate regimes. The \( \delta^{15}N \) values were plotted against present-day decadal averages for annual temperature and precipitation for nearby locations (accessed via http://www.wrcc.dri.edu/cgi-bin/cliMAIN.pl?mn5325 and http://www.eldoradocountyweather.com/canada/climate2/Calgary.html). The results suggest that modern precipitation is a better predictor for palaeo \( \delta^{15}N \) variation between regions than temperature or latitude, even suggesting the inverse-convex relationship seen in modern studies of this kind, especially in arid regions.

**SM Figure 18.** Investigating possible confounding factors in the climate-\( \delta^{15}N \) relationship. This representative subset of the data suggests latitude bears no obvious direct relationship to \( \delta^{15}N \) values.
SM Figure 19. Investigating possible confounding factors in the climate-$\delta^{15}$N relationship. This representative subset of the data suggests Mean Annual Temperature (MAT) bears no obvious direct relationship to $\delta^{15}$N values.

SM 20. Investigating climate-$\delta^{15}$N relationships: Mean Annual Precipitation (MAP) appears to obey the inverse-convex relationship to $\delta^{15}$N seen in empirical studies.

Supplementary: References

SM References 1: Dated $\delta^{15}$N measurements
Publications supplying data for this study are listed by region and taxon in SM Methods. The data are available in Supplementary Data. The following institutions and collectors kindly supplied samples for this publication, which we gratefully acknowledge:

USA: American Museum of Natural History, New York, USA; University of Alaska, Fairbanks, USA; University of Kansas Museum of Natural History, Lawrence, KS, USA; University of Alaska Museum of the North, USA; Alaska Department of Fish and Game, Fairbanks, USA; Bell Museum of Natural History, University of Minnesota, USA

Canada: Canadian Museum of Nature, Ottawa, Canada; Canadian Museum of Civilisation, Gatineau, Canada; Provincial Museum of Alberta / Royal Alberta Museum, Edmonton, Canada; Simon Fraser University, Vancouver, Canada; University of Victoria, Victoria, Canada; Yukon Heritage Centre, Whitehorse, YT, Canada; Yukon Palaeontology Program, Whitehorse, YT, Canada; M. C. Wilson, private collection, Canada

Russia: Institute of Plant and Animal Ecology, Ekaterinburg, Russia; Paleontological Institute, Moscow, Russia; Zoological Institute, St. Petersburg, Russia; Northern Research Station, Cherskii, Russia; Laboratory of Prehistory, St Petersburg, Russia; Local Museum, Chersky, Russia

Europe: Malmo Museum of Natural History, Sweden; Zoological Museum Amsterdam, Netherlands; Natural History Museum, London, UK
South America: Instituto de la Patagonia, Universidad de Magallanes, Magallanes, Chile; La Plata Museum, Universidad Nacional de Cuyo, Mendoza, Argentina; Museo de Historia Natural de San Rafael, Mendoza, Argentina; Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, Peru

SM References 2: Glaciation chronologies
Eurasian Ice Sheet: Van den Berg et al. (2008) [63], Figure 10; Northern Urals: Svendesen et al. (2014) [64], Figure 14; North American Ice Complex: Tarasov et al. (2003) [65], Figure 18; Patagonian Ice Sheet: Boex et al. (2013) [66], Figure 3.

SM References 3: Peatland Initiation Dates
Northern Hemisphere (n = 1870):[7-9]; South America (n = 54):[9].

SM References 4: Climatic factors/proxies
NGRIP: [67]; EPICA Dome C: [68]; Insolation: [69](http://CRAN.R-project.org/package=palinsol).

SM References 5: Selected studies demonstrating the climate-δ^{15}N link in modern arid and semi-arid environments

<table>
<thead>
<tr>
<th>Study Title</th>
<th>Reference</th>
<th>Type</th>
<th>Region</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convergence of soil nitrogen isotopes across global climate gradients</td>
<td>Craine 2015 [70]</td>
<td>Global</td>
<td>Multiple</td>
<td>Mixed, arid less well-represented</td>
</tr>
<tr>
<td>Kangaroo metabolism does not cause the relationship between bone collagen δ^{15}N and water availability</td>
<td>Murphy 2006 [71]</td>
<td>Arid</td>
<td>Australia</td>
<td>Grass, Kangaroo</td>
</tr>
<tr>
<td>Study Title</td>
<td>Author(s)</td>
<td>Region</td>
<td>Species</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>----------------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>Lack of precipitation effects on δ¹⁵N for animals consuming low amounts of C3 plants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altitudinal gradients of grassland carbon and nitrogen isotope composition are recorded in the hair of grazers</td>
<td>Mannel 2007[74]</td>
<td>Semi-arid–Mesic</td>
<td>Europe Sheep, Cattle, Goats, Alpaca, Horse</td>
<td></td>
</tr>
<tr>
<td>Quaternary record of aridity and mean annual precipitation based on δ¹⁵N in ratite and dromornithid eggshells from Lake Eyre, Australia</td>
<td>Newsome 2011[75]</td>
<td>Arid–Semi-arid</td>
<td>Australia Moa Eggshell, Grass</td>
<td></td>
</tr>
<tr>
<td>Annual rainfall and nitrogen-isotope correlation in macropod collagen: application as a paleoprecipitation indicator</td>
<td>Grocke 1997[77]</td>
<td>Arid, Coastal</td>
<td>Australia Possum, Kangaroo</td>
<td></td>
</tr>
<tr>
<td>Climatic influence on the</td>
<td>Heaton 1986[78]</td>
<td>Arid</td>
<td>Africa Human, Elephant,</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>----------------</td>
<td>------</td>
<td>--------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>
Supplementary: Code (in R v3.2.2)

Subroutines

# z-transform a vector
ztransform <- function(data)
{
  mu <- mean(data)
  sd <- sd(data)
  (data - mu)/sd
}

# residuals from the mean, not used in final paper but possibly useful.
residuals <- function(data)
{
  mu <- mean(data)
  data - mu
}

# weighting values based on the Gaussian. Scaled to between 0 and 1.
gaussian_weights <- function(calc_point, data_points, stdev)
{
  dnorm(data_points, mean = calc_point, sd = stdev)/dnorm(0, 0, sd = stdev)
}

# z-transforms by levels of a factor. For big datasets, rewriting this using apply-family functions will be essential.
factorwise_ztransform <- function(x, values_colname, factor_colname)
{
  for(lev in levels(x[[factor_colname]]))
  {
    x[values_colname][x[factor_colname] == lev & !is.na(x[factor_colname]) & !is.na(x[values_colname])] <-
    ztransform(x[values_colname][factor_colname] == lev & !is.na(x[factor_colname]) & !is.na(x[values_colname]))
  }
  out <- x[values_colname]
  colnames(out) <- paste(sep = "", values_colname, "_Z_trans_", factor_colname, "_wise")
  out
}

# not used in final paper but possibly useful
factorwise_residuals <- function(x, values_colname, factor_colname)
{
  for(lev in levels(x[[factor_colname]]))
  {
    x[values_colname][x[factor_colname] == lev & !is.na(x[factor_colname]) & !is.na(x[values_colname])] <-
    residuals(x[values_colname][factor_colname] == lev & !is.na(x[factor_colname]) & !is.na(x[values_colname]))
  }
  out <- x[values_colname]
  colnames(out) <- paste(sep = "", values_colname, "_residuals_", factor_colname, "_wise")
  out
}

# give it a vector of datapoints
bootstrap <- function(x)
{
  x[floor(runif(nrow(x)) * nrow(x)) + 1]
}

RMAC_unscaled <- function(x, df, dd_sd, sd_a, sd_b, top, bottom, ...)
{
  df <- df[!is.na(df[1]) & !is.na(df[2]),]
  if(!is.na(df[3]))
  {
    df <- df[match(df[3], list(...))]
  }
  if(sum(x <= 0) > 0) (df[1] - df[1]) # reversing the x-axis just feeds stat_function the x-values as negative ...
  # so we deal with it like this
  df[2] - min(df[2]) + 1000000 - df[2] # this is a relative scale and it's just easier to work with positive values here ...
  # rather than messing with the function. the result is exactly the same.

  outvector <- numeric() # define output storage
  for (t in x) # t is the time at which we want to know the RMAC_unscaled
  {
    num <- 0 # numerator
den <- 0 # denominator
    l_in <- 0 # number within 2sd of mean on left
    r_in <- 0 # on right
    # for this s (aka p), scan through time/d15N pairs, apply contribution to numerator (N_density(p-time,0,sd)*d15N) and denominator (N_density(p-time,0,sd))
    print("t"); print(t)
  }
}
densityscore <- 0
for (row in 1:nrow(df)) # for each datapoint (time, measurement) pair
{
  dist <- t - df[row, 1] # distance of datapoint from t
  if (dist >= 0) & (dist <= 2 * dd_sd) [l_in <- l_in + 1] # l_in: points to the (l)eft of t that are (in) the threshold
  else if (dist < 0) & (dist >= -2 * dd_sd) [r_in <- r_in + 1] # r_in: points to the (r)ight of t that are (in) the threshold
  densityscore <- densityscore + dnorm(dist, 0, dd_sd) # a score to characterise how dense the datapoints are around this t.
}
  #print(paste("lin/rin", l_in, "/", r_in))
  #print(paste("density score:", densityscore))
  sd_wt <- sd_a * exp(-sd_b * densityscore) # define an appropriate sd for the weighting distribution
  #print("sd this calc point:", sd_wt) # really useful for choosing meaningful params
for (row in 1:nrow(df)) # for each datapoint (time, measurement) pair
{
  dist <- t - df[row, 1] # dist of point from t
  weight <- dnorm(dist, 0, sd_wt) # using the defined weighting dist
  num <- num + (weight * df[row, 2])
  den <- den + weight;
}
mu <- numeric()
if (!(l_in > 0) || !(r_in > 0) || (l_in + r_in < 6)) {mu <- NA} # limits based on lin and rin
else {mu <- 1 / (num / den)} # no limits
outvector <- c(outvector, mu)
}
  #print(outvector)
  #(((outvector - min(outvector, na.rm = T))/(max(outvector, na.rm = T) - min(outvector, na.rm = T)) *(top-bottom))+bottom) -> outvector # to scale between top and bottom
return(outvector)
}
# returns the mean or 95% CIs at each time. "What" can be "mu", "upper" or "lower".
RMAC_unscaled_bootstrap_intervals <- function(x, df, dd_sd, sd_a, sd_b, top, bottom, what,...)
{
  n_bootstraps <- 1500
  x -> times
do.call(rbind, lapply(1:n_bootstraps, function(x) RMAC_unscaled(times, bootstrap(df, dd_sd, sd_a, sd_b, top, bottom,...))) -> bstraps
  # rescale here
  (bstraps - min(bstraps, na.rm = T)) / (max(bstraps, na.rm = T) - min(bstraps, na.rm = T)) * (top-bottom) + bottom -> bstraps
return(do.call(rbind, apply(bstraps, 2, function(x)
{
  strap <- x[!is.na(x)]
  n_out_each_end <- round((.05 * length(strap)) / 2), digits = 0)
  if (n_out_each_end > 0)
  {
    return(data.frame('mu' = median(strap), 'upper' = strap[order(strap)][(length(strap) - n_out_each_end)], 'lower' = strap[order(strap)][n_out_each_end + 1])
  } else
  {
    return(data.frame('mu' = NA, 'upper' = NA, 'lower' = NA))
  }
}
), what) # delete [what] to get the whole data frame, i.e. what is plotted in the ms fig 1.
}

Data Handling and Calibration

The column “notes” here was used as a convenient informal repository of relevant metadata concatenated into as-long-as-necessary character strings and used to group or extract information for samples. It is not included in the supplementary data but is available upon request.

x <- read.csv(file="data.csv") # columns named "study", "genus", "region", "date", "cal_error", "d13C", "d15N", "dating_record", "notes", "locality", "extra", "diet", "d15N_Z_trans_genuswise", "cal_date"
as.character(x$xnotes) -> x$notes
as.character(x$extra == "") <- NA
as.character(x$dating_record) -> x$dating_record
as.character(x$locality) -> x$locality
as.numeric(sub("\", ",", x$d15N, fixed = T)) -> x$d15N
as.numeric(sub("\", ",", x$d13C, fixed = T)) -> x$d13C
as.numeric(sub("\", ",", x$date, fixed = T)) -> x$date

Data Handling and Calibration

The column “notes” here was used as a convenient informal repository of relevant metadata concatenated into as-long-as-necessary character strings and used to group or extract information for samples. It is not included in the supplementary data but is available upon request.

x <- read.csv(file="data.csv") # columns named "study", "genus", "region", "date", "cal_error", "d13C", "d15N", "dating_record", "notes", "locality", "extra", "diet", "d15N_Z_trans_genuswise", "cal_date"
as.character(x$xnotes) -> x$notes
as.character(x$extra == "") <- NA
as.character(x$dating_record) -> x$dating_record
as.character(x$locality) -> x$locality
as.numeric(sub("\", ",", x$d15N, fixed = T)) -> x$d15N
as.numeric(sub("\", ",", x$d13C, fixed = T)) -> x$d13C
as.numeric(sub("\", ",", x$date, fixed = T)) -> x$date

Data Handling and Calibration

The column “notes” here was used as a convenient informal repository of relevant metadata concatenated into as-long-as-necessary character strings and used to group or extract information for samples. It is not included in the supplementary data but is available upon request.

x <- read.csv(file="data.csv") # columns named "study", "genus", "region", "date", "cal_error", "d13C", "d15N", "dating_record", "notes", "locality", "extra", "diet", "d15N_Z_trans_genuswise", "cal_date"
as.character(x$xnotes) -> x$notes
as.character(x$extra == "") <- NA
as.character(x$dating_record) -> x$dating_record
as.character(x$locality) -> x$locality
as.numeric(sub("\", ",", x$d15N, fixed = T)) -> x$d15N
as.numeric(sub("\", ",", x$d13C, fixed = T)) -> x$d13C
as.numeric(sub("\", ",", x$date, fixed = T)) -> x$date

Data Handling and Calibration

The column “notes” here was used as a convenient informal repository of relevant metadata concatenated into as-long-as-necessary character strings and used to group or extract information for samples. It is not included in the supplementary data but is available upon request.

x <- read.csv(file="data.csv") # columns named "study", "genus", "region", "date", "cal_error", "d13C", "d15N", "dating_record", "notes", "locality", "extra", "diet", "d15N_Z_trans_genuswise", "cal_date"
as.character(x$xnotes) -> x$notes
as.character(x$extra == "") <- NA
as.character(x$dating_record) -> x$dating_record
as.character(x$locality) -> x$locality
as.numeric(sub("\", ",", x$d15N, fixed = T)) -> x$d15N
as.numeric(sub("\", ",", x$d13C, fixed = T)) -> x$d13C
as.numeric(sub("\", ",", x$date, fixed = T)) -> x$date

Data Handling and Calibration

The column “notes” here was used as a convenient informal repository of relevant metadata concatenated into as-long-as-necessary character strings and used to group or extract information for samples. It is not included in the supplementary data but is available upon request.

x <- read.csv(file="data.csv") # columns named "study", "genus", "region", "date", "cal_error", "d13C", "d15N", "dating_record", "notes", "locality", "extra", "diet", "d15N_Z_trans_genuswise", "cal_date"
as.character(x$xnotes) -> x$notes
as.character(x$extra == "") <- NA
as.character(x$dating_record) -> x$dating_record
as.character(x$locality) -> x$locality
as.numeric(sub("\", ",", x$d15N, fixed = T)) -> x$d15N
as.numeric(sub("\", ",", x$d13C, fixed = T)) -> x$d13C
as.numeric(sub("\", ",", x$date, fixed = T)) -> x$date

Data Handling and Calibration

The column “notes” here was used as a convenient informal repository of relevant metadata concatenated into as-long-as-necessary character strings and used to group or extract information for samples. It is not included in the supplementary data but is available upon request.
as.numeric(as.character(x$d15N)) -> x$d15N

# added stuff for exploration
levels(x$extra) <- c(levels(x$extra), 'EB', 'NonEBCan', 'USA')
x$extra$xregion == 'NorthAmerica' & grepl('[hicken[ukon][YT][laska]AI[trish Gulch][quartz', x$notes] <- 'EB'
x$extra$xregion == 'NorthAmerica' & grepl('[herta][Aljokubia][monon][Chall]Buffalo', x$notes] <- 'NonEBCan'
x$extra$xregion == 'NorthAmerica' & grepl('[hicken[ukon][YT][laska][minnesota][Trap][48 States', x$notes] <- 'USA'
x$extra$xregion == 'NorthAmerica' & grepl('[azula[obsb[za pak', x$study] <- 'EB'

# separate llamas by species
levels(x$extra) <- c(levels(x$extra), 'guanaco', 'gracilis', 'unidentified')
x$extra$genus == 'Lama' & grepl('[ExinctG', x$extra] <- 'guanaco'
x$extra$genus == 'Lama' & grepl('[Gracil', x$extra] <- 'gracilis'
x$extra$genus == 'Lama' & grepl('[Holocene', x$extra] <- 'gracilis'
x$extra$genus == 'Lama' & grepl('[UnID', x$extra] <- 'unidentified'
x$extra$xregion == 'USA' & is.na(x$extra] <- 'EB'
levels(x[grepl(['USA', x$extra] & !grepl('unidentified', x$extra)],) -> x

# add siberia separation
levels(x$extra) <- c(levels(x$extra), 'Urals', 'Other')
x$extra[grepl('rals', x$notes] & x$region == 'Siberia' <- 'Urals'
x$extra[grepl('rals', x$notes] & x$region == 'Siberia' <- 'Other'

# do z-transformations
droplevels(x) -> x
rm(b,t)
for (i in levels(x$region))
{
  x[x$region == i] -> t
  cbind(factorwise_ztransform(L,'d15N','genus')) -> t
  if(exists('b'))
  {
    t -> b
  }
  else
  {
    rbind(b,t) -> b
  }
}
b -> x
rm(b,t)
x <- x[
  x$date >= 45000 &
  x$date <= 1000 &
  x$d15N_Z_trans_genuswise > -3 & x$d15N_Z_trans_genuswise < 3 &
  1
]
x <- x[
  ls.na(x$date] & ls.na(x$d15N] &
  1
]
x <- x[x$genus != 'Mammuthus']

calibrate
x$cal_date = x$date
library(Bchron)
length(x$date[ls.na(x$date] & ls.na(x$cal_error)]) -> l
calcurves <- ifelse(x$region[ls.na(x$date] & ls.na(x$cal_error]) == 'SouthAmerica', 'shcal13', 'intcal13')
BchronCalibrate(ages=x$date[ls.na(x$date] & ls.na(x$cal_error], ageSds=x$cal_err[ls.na(x$date] & ls.na(x$cal_error], calCurves=calcurves) -> calcs
x$cal_date[ls.na(x$date] & ls.na(x$cal_error]] <- as.numeric(lapply(calcs,function(i)
  range(i$AgeGrid[1]+sum(i$densities)/2)*(range(i$AgeGrid[2]-range(i$AgeGrid[1])))
))

droplevels(x) -> x
# tidy up and have a look
str(x)
table(x$genus,x$region)
table(x$study,x$region)
Insolation/Peatland Initiation

```r
# Insolation
library(palinsol)
insolation <- function(times, astrosol=la04, ltd, ..., 
sapply(times, function(tt) Insol(orbit=astrosol(tt, long=ltd), lat=65*pi/180))
tts <- seq(from = -45000, to = 0, by = 100)
isl_S <- insolation(tts, ber78, ltd=pi/2)
isl_W <- insolation(tts, ber78, ltd=pi/2)
plot(tts, isl, type='T')

ggplot(data=data.frame(isl_S=isl_S, isl_W=isl_W, time=-tts)) + geom_line(mapping=aes(x=time, y=isl_S)) + geom_line(mapping=aes(x=time, y=isl_W))

# Peatlands
# Read in
read.csv(file="data.csv")[,1:7] -> x # columns include "lat", "long", "date", "region"
# A little cleaning
sub("\", "", x$date) -> x$date
# Assign regions by latitude
c(levels(x$region), "Siberia", "Europe", "NA", "NorthSlope") -> levels(x$region)
cat(x$long)[startsWith(x$region, "MacDonnell_06")].breaks <- c(-180, -169.02, -54.99, -13.9, 58.68, 180).labels = c("Siberia", "NorthAmerica", "NA", "Europe", "Siberia")
x$region[x$ref == "MacDonnell_06"]
x$region[x$lat > 68.26 & x$long > -166.49 & x$long < -140.96] <- "NorthSlope"
x < droplevels(x)
x$date <- as.numeric(as.character(x$date))

# T-tests (SM Figure 13–16)

# Do p-values over time function (x-time point, df = data frame, will give you p-value using t-test for all d15N vals pre-point vs. after point)
pval_split <- function(from, to, df)
# Requires columns of data frame 'df' labeled 'date' and 'd15N_Z_trans_generuswise'.
{
  # print(paste("to:	", to, ",	from:	", from))
  if (from > 45000) {return(NA)} # window beyond end
  if (to < from) {return(NA)} # window beyond < df$d15N_Z_trans_generuswise[(df$date > from | df$date < to) & !is.na(df$d15N_Z_trans_generuswise)] # before + after
  within <- df$d15N_Z_trans_generuswise[df$date < from & !is.na(df$d15N_Z_trans_generuswise)] # between
  if (length(between) < 5 || length(within) < 5) {return(NA)}
  t.test(x = between, y = within)["p.value"]
}

# Make matrix # Each cell will be the p-value when excluding the region [from, to]
inc <- 1000
loc <- "NorthSlope"
make.m <- data.frame where rows are the startpoints of each window, and cols are the winsize
m <- sapply(seq(1000, 20000, by = 1000), function(winsize, Ax[x$region == loc]))
m <- as.data.frame(m)
dimnames(m) <- list(seq(0, 45000, by = inc), seq(1000, 20000, by = 1000))
df <- data.frame()
for(i in colnames(m))
  for(j in rownames(m))
  {
    df <- rbind(df, data.frame(t = j, w = m[j, i]))
  }

# A typical plot
library(ggplot2)
ggplot(data = df, aes(x = t, y = w, fill = log(p))) + geom_tile() + scale_fill_gradient(low = "steelblue", high = "white", guide = guide_legend(title = "log(p-value)")) + scale_x_discrete(breaks = seq(0, 45000, 5000)) + xlab("Window begins (earliest) (kya)") + ylab("Window size (ky)") + ggtitle(paste(loc, "d15N z-scores (two-sample t-test, values in window vs other) log p-values"))
```

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
Supplementary: Bibliography


68. Jouzel, J., EPICA Dome C Ice Core 800KYr Deuterium Data and Temperature Estimates, IGBP PAGES/World Data Center for Paleoclimatology. 2007, Data Contribution Series.