SUPPLEMENTARY INFORMATION

doi:10.1038/nature10574

SECTION OVERVIEW

SECTION S1: Species distribution modelling

1.1 Introduction

Species distribution models, or SDMs, have been developed over the last three decades to address our incomplete knowledge of species distributions, a challenge described as the "Wallacean shortfall"[1]. While primarily developed to estimate the current distribution of species for which we have incomplete sampling, SDMs have also been heavily utilized over the last decade to forecast species future distributions due to modern climate change[2]. They may also be a promising tool for reconstructing the distribution of species in past time periods[3] for which varying sampling intensity and bias in the fossil record are more significant problems than in the distribution data from current ecological sampling schemes.

Species distribution models are deeply rooted in niche theory[4,5] (see also Soberón[6] for a recent study on niche theory and SDMs) and gradients analysis[7,8]. They link ecological theory and statistics under the principle that species abundance and population performance, which control species distributions, change across environmental gradients[9,10], by relating the distribution of species and the environmental conditions in which they occur in an *n*-dimensional environmental space, in which each dimension is an environmental variable, to statistically describe the environmental niche of a species (or the climatic niche if only climatic variables are used). The modelled species niche can be transferred into geographical space where each grid cell (or unit of space) is assigned specific values of the environmental parameters used to define the species niche. The methodological approach, which transfers the species niche from environmental to geographical space is rooted in the duality between Hutchinson's "niche" and "biotope"[11]. Under climate change, the spatial extent of suitable climatic conditions for a given species can increase or decrease, driving changes in the distribution of that species. For example, a large reduction in the availability of suitable climate conditions would be expected to cause a reduction in a species' realised distribution, thus contributing to a reduction of population size and a potential increase in extinction risk[12].

To relate changes in the megafauna species' (woolly rhinoceros (*Coelodonta antiquitatis*), woolly mammoth (*Mammuthus primigenius*), horse (wild *Equus ferus* and living domestic *Equus caballus*), reindeer/caribou (*Rangifer tarandus*), bison (*Bison priscus/Bison bison*) and musk ox (*Ovibos moschatus*)) distributions against estimates of effective population size from the Bayesian skyride models (see Supplementary Information section S3), we used SDMs to estimate range sizes for each

species through the late Quaternary. In practice, SDMs reconstruct species' geographic distributions by relating species' presence records (in this case, fossil locality data) to a set of environmental predictors (e.g., temperature and rainfall) to map a species' geographic range using a geographic information system (GIS)[13]. Strong enthusiasm for incorporating SDMs in a variety of biological studies has resulted in intense scrutiny of the method's theoretical assumptions[14]. Paramount is recognising the difference between a species' fundamental niche, the full set of conditions in which a species can survive long-term, and the realised niche, the subset of the fundamental niche that is actually occupied at a given time[5] and upon which SDMs are based. SDMs are generated using climatic data[15], but a species' realised niche is also determined by other factors (such as barriers to dispersal). Projecting an SDM onto past or future climate surfaces, as is common in climate change studies, may ignore those limits while assuming a species will exist in all places with favourable climatic conditions, and that the niche is static through evolutionary time—assumptions which need to be explored for many species[14]. Further, combinations of climatic variables with no analogues in other time periods may result in underestimation of a species' ecological and geographic range in past or future projections[16].

Therefore, range size estimates to be compared with the genetic data (results in Fig. 2 in main text) were modelled using only locality and climate data from the same time periods (42, 30, 21 and 6 kyr BP); SDMs from one time period were not projected onto earlier or later periods, and range measurements were restricted to regions for which fossils were used to build the models, rather than all potentially suitable Holarctic area. This approach allowed us to circumvent assumptions regarding climatic niche stasis through time, as well as the effects of dispersal limitations which might have prevented species from reaching areas of otherwise suitable habitat.

1.2 Palaeoclimate data

Late Quaternary climatic conditions are simulated using Atmospheric-Ocean coupled General Circulation Models (hereafter AOGCMs). An AOGCM is a set of equations simulating the dynamics of the ocean and the atmosphere under certain environmental conditions (i.e, CO^2 concentration, ice sheet extent) to provide estimates of past climatic parameters (e.g., rainfall or temperature). Each AOGCM differs slightly in both the absolute values of estimated climatic conditions and in the geographical distribution of those conditions. To assess the effect of AOGCM choice on species' distributions and the subsequent relationship between range size and effective population size, we simulated past climatic conditions using two different AOGCMs: GENESIS2 and HadCM3.

1.2.1 GENESIS2

Four GENESIS2 simulations were used: two for Marine Isotope Stage 3 (MIS 3), one for the Last Glacial Maximum (LGM; ~21 kyr BP) and one for the mid-Holocene (~6 kyr BP). The Marine Isotope Stage 3 (MIS 3) simulations represent the warmer middle part (~42 kyr BP) and colder later part (~30 kyr BP) of MIS 3. Carbon dioxide levels were specified at 200 ppm for the MIS 3 and LGM simulations[17] and 280 ppm for the mid-Holocene simulation[18]. Sea surface temperatures (SSTs) for the MIS 3 and LGM simulations were taken primarily from CLIMAP[19], with modifications from GLAMAP-2000 and other sources[20]. SSTs for the mid-Holocene simulation were prescribed at present-day values[21]. In all cases, insolation was calculated using orbital parameters[22,23]. All simulations were spun up to equilibrium; results are 10-year averages.

Comparison of GENESIS2 model output to proxy data shows that temperatures in Europe are accurate to within ± 1°C for the mid-Holocene and ± 2°C for the LGM[24]. For Oxygen Isotope Stage 3, [25] found that GENESIS2 temperatures in southern Europe agree well with proxy data but were 3-4°C too warm in Northern Europe. Comparison of present day GENESIS2 model output with observations suggests that variability in Europe can be extrapolated to all of northern Eurasia and North America[26]. We are unaware of any GENESIS2 model-data comparisons for palaeoprecipitation. However, for present-day northern Eurasia and North America the model is accurate to within ± 1 mm day^{-1} when compared to observations[26]. Atmospheric carbon dioxide boundary conditions are well-constrained[27,28].

1.2.2 HadCM3

A second set of climate model outputs for the same Quaternary periods, the warmer middle part (~40 kyr BP) and colder later part (~32 kyr BP) of MIS 3, LGM and the Mid-Holocene were derived from the Hadley Centre Coupled Climate Model Version 3 (HadCM3). The simulations form part of the ensemble presented in [29]. The particulars of HadCM3 are well documented[30]. HadCM3 was one of the first coupled atmosphere-ocean climate models which required no flux corrections, even for simulations of a thousand years or more[31]. The climate model consists of a linked atmospheric model, ocean model and sea ice model. In HadCM3 the horizontal resolution of the atmosphere model is 2.5 degrees latitude by 3.75 degrees longitude. This gives a grid spacing at the equator of 278 km in the north-south direction and 417 km east-west and is approximately comparable to a T42 spectral model resolution. The atmospheric model consists of 19 layers. The spatial resolution over the ocean in is 1.25˚ × 1.25˚ and the model has 20 layers. The atmospheric

model has a time step of 30 minutes and includes a radiation scheme that can represent the effects of minor trace gases[32]. A parameterization of simple background aerosol climatology is also included[33]. The convection scheme is that of [34]. A land-surface scheme includes the representation of the freezing and melting of soil moisture. The representation of evaporation includes the dependence of stomatal resistance on temperature, vapour pressure and CO^2 concentration[35].

The ocean model includes the use of the Gent-McWilliams mixing scheme[36]. There is no explicit horizontal tracer diffusion in the model. The horizontal resolution allows the use of a smaller coefficient of horizontal momentum viscosity leading to an improved simulation of ocean velocities. The sea ice model is a simple thermodynamic scheme and contains parameterizations of ice drift and leads (Polynyas[37]).

For the MIS 3, LGM and mid-Holocene, orbital parameters are taken from [23]. Atmospheric concentrations of CO_2 were taken from the Vostok ice core record[38] and CH_4, and N_2O were taken from EPICA[39]. All ice-core data were on the same EDC3 timescale[40].

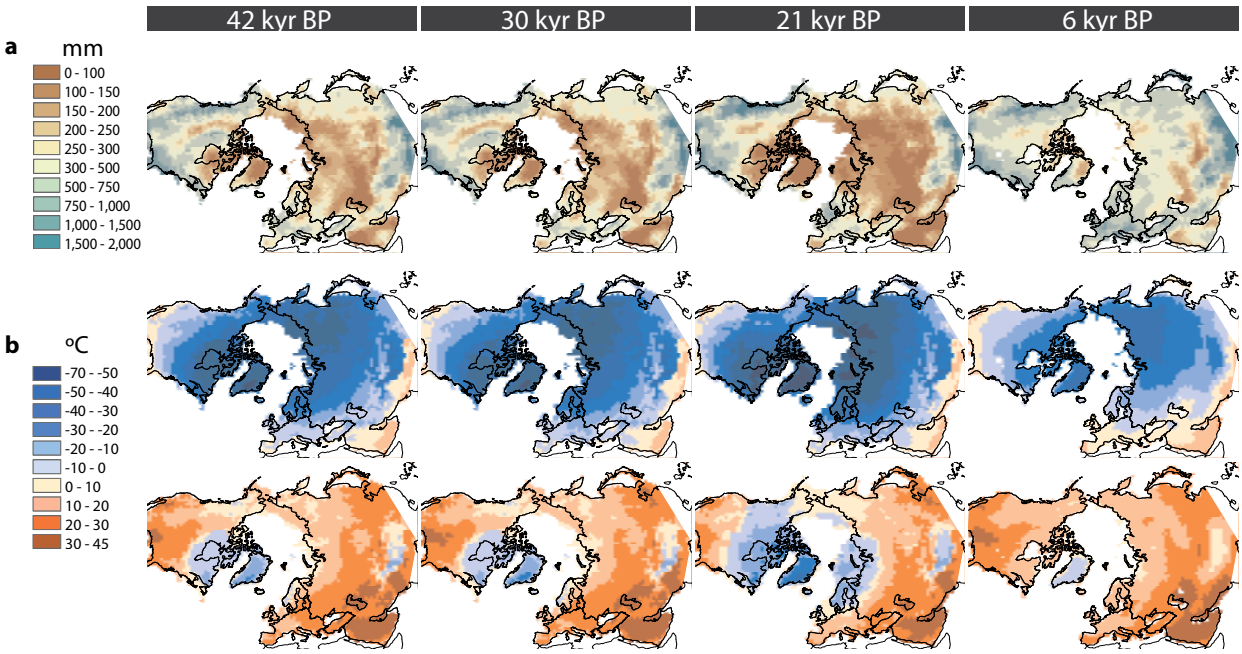
Ice-sheet reconstructions are developed from the ICE5G model[41], which includes a detailed evolution of the ice thickness, extent and continental isostatic rebound for the whole period from the LGM to the modern at 500-year intervals. Using standard linear interpolation techniques, this dataset was used to calculate, at the scale of the climate model, the total continental elevation (including the direct thickness of the ice sheets plus the effects of isostatic adjustment), bathymetry (including isostatic changes), ice-area extent, and land sea mask for the LGM and mid-Holocene. To ensure consistency with pre-industrial boundary conditions an anomaly-based method was used to calculate palaeogeographic boundary conditions. In this method, for anomalies of a particular time-slice, palaeogeography minus pre-industrial ICE-5G data are then added to our model pre-industrial geographical boundary conditions. The geographical extent and heights of the major ice sheets prior to the LGM were based on [42], which included a calculation of the pre- and post-glacial ice. Singarayer and Valdes[29] used the SPECMAP[43] record of $d^{18}O$ history to constrain the evolution of the volume of land ice from the last interglacial up to the LGM.

Each simulation was integrated for approximately 200 years, a sufficient period of time to bring the surface climatology to equilibrium, with the final 30 years used to calculate the required climatological mean. The version of HadCM3 used does not include interactive vegetation, so all simulations use the same pre-industrial vegetation boundary condition. Similarly, aerosol loading in the model is unchanged and does not account for changes in dust during the cycle.

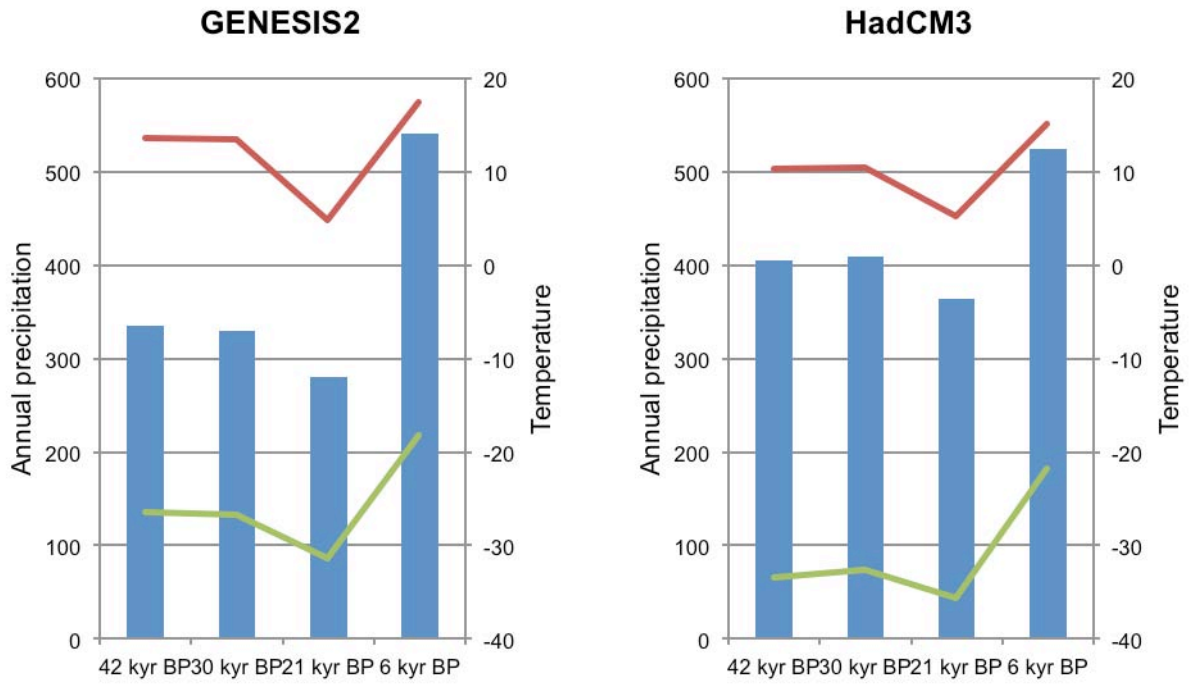
Using HadCM3, we obtain reasonable estimates of the global temperature glacial/interglacial range as well as trends in polar regions, although the magnitude of change at high latitudes is underestimated, as other similar models have found for the LGM[29]. HadCM3 produces a pattern of cooling for the LGM which are broadly consistent with the findings from simpler models and palaeoclimatic data[44]. HadCM3 performance in simulating both the LGM and mid-Holocene have been evaluated and are recognised as being either good or generally comparable with other climate models run for the same time intervals[45].

1.2.3 Palaeoclimate variables

From both AOGCM datasets, three variables (Supplementary Figure S1.1) were selected to describe the potential distribution of each species: mean temperature of the coldest month (°C), mean temperature of the warmest month (°C), and annual precipitation (mm). Temporal trends in these climatic variables are similar between the two AOGCMs (Supplementary Figure S1.2). Palaeoclimatic simulations for GENESIS2 are at a 2×2-degree spatial resolution; palaeoclimatic simulations for HadCM3 are resampled at a 2×2-degree resolution. This conservative variable set was selected to balance the number of species occurrences versus the number of climatic variables used to calibrate species' realised climatic niches. Using many climatic variables to model the potential distribution of a species using few presence records (e.g., fossil localities) is likely to lead to model over-fitting, yielding a misrepresentation of the geographical distribution of the modelled species[46]. Given the constraint of using only a small number of climatic variables, we aimed to capture the upper and lower thermal limits of each species, as well as a moisture variable. Previous applications of SMDs have used similar limited sets of climatic variables as predictors of megafauna distributions[3]. All figures presented in the main text were generated using climate variables from GENESIS2.



Supplementary Figure S1.1. Palaeoclimatic data (GENESIS2) for annual precipitation (mm), average temperature of the warmest month (°C), average temperature of the coldest month (°C) and the fossil record were used to estimate species potential ranges.



Supplementary Figure S1.2. Average climatic temporal trends across the Holartic based on GENESIS, left, and HadCM3, right, AOGCMs. Green (lower) lines indicate mean temperature of

the coldest month (°C), red (upper) lines indicate mean temperature of the warmest month (°C), and blue bars indicate annual precipitation (mm).

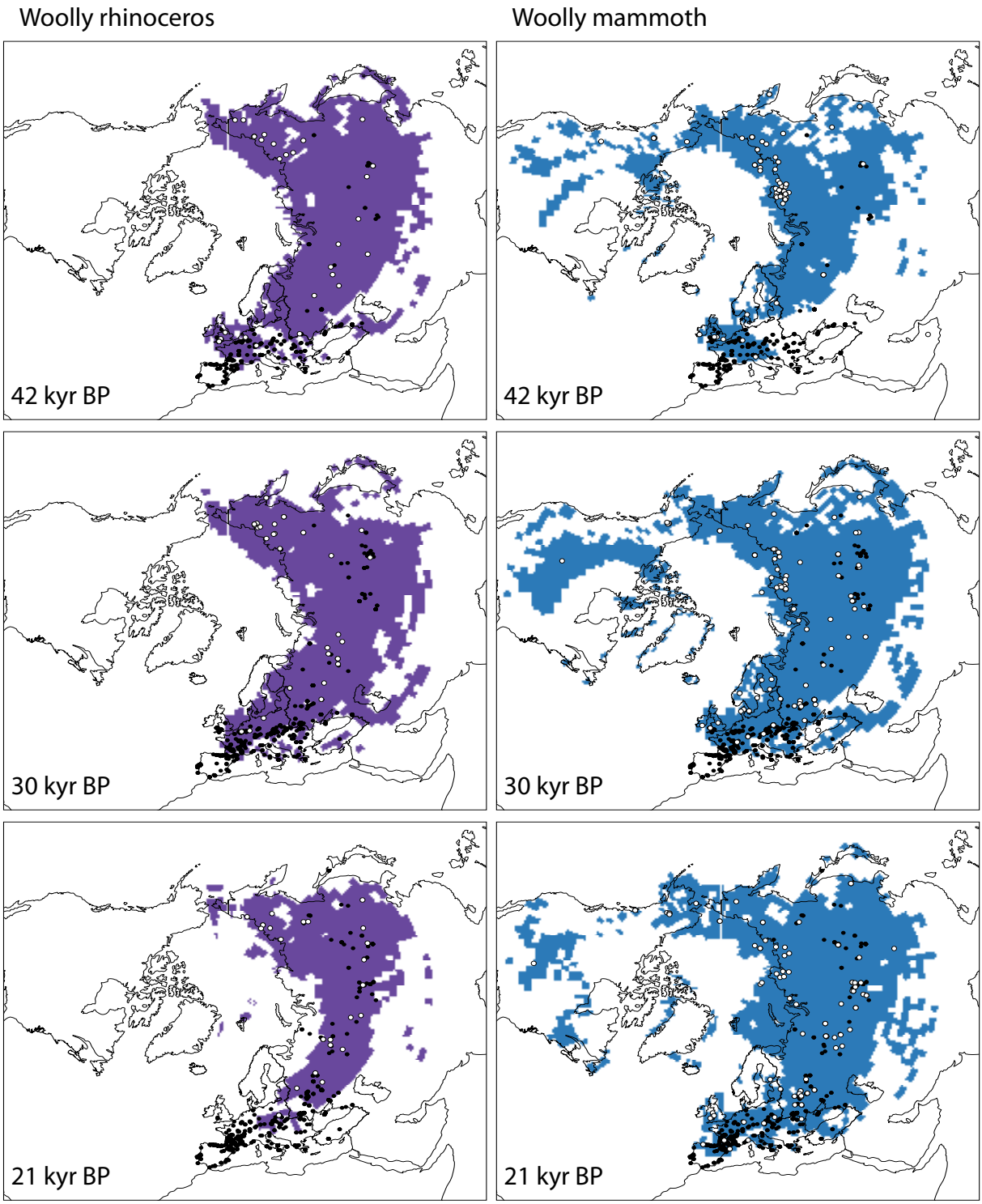
1.3 Megafauna locality data

For each species, ^{14}C-dated fossil localities from Eurasia and North America were obtained for the following calendar time intervals: 45–39, 33–27, 24–18 and 9–3 kyr BP. Radiocarbon dates (uncalibrated ^{14}C dates) were calibrated into calendar years using the IntCal09 calibration curve[47] using the OxCal 4.1 online calibration resource (https://c14.arch.ox.ac.uk). The 829 fossil localities included data from sequenced specimens (Supplementary TableS6.2-6.4), supplemented with fossil localities from the literature (Supplementary Table S6.1); The majority of the 829 localities used for the Species Distribution Models (SDMs) were compiled from synthetic works, rather than the original papers in which the dates were first presented, so we were unable to evaluate our data by fossil context, dated material or dating method[48]. However, the vast majority (98%) of the 829 localities used were from directly dated (standard or AMS) animal remains; thus, if we assume that most dates were from bone collagen, dung, hide or hair, we anticipate that most of our directly dated samples would rank an 11 or 12 on the scale proposed by [48] . Given the breadth of the time bins used—6,000 calibrated years—the difference between dating methods should have little impact on our results, as this is significantly greater than the differences between conventional and AMS dates observed by [48].

The list of localities is not and was not intended to be exhaustive, but was meant to cover at least those regions with genetic data to enable comparison of the estimates of potential range size and estimated effective population size. Therefore, modelled species ranges are not intended to fully represent past species distributions in great detail, and may under-represent the species' actual range in areas for which we have little data. Fossil localities are indicated in Supplementary Fig. S1.3. Data were only included where the literature contained explicit geographic coordinates or detailed site descriptions which could be located at http://toolserver.org/~geohack/. As each set of geographical coordinates relates to a specific dated fossil, localities are duplicated where more than one dated fossil has been found.

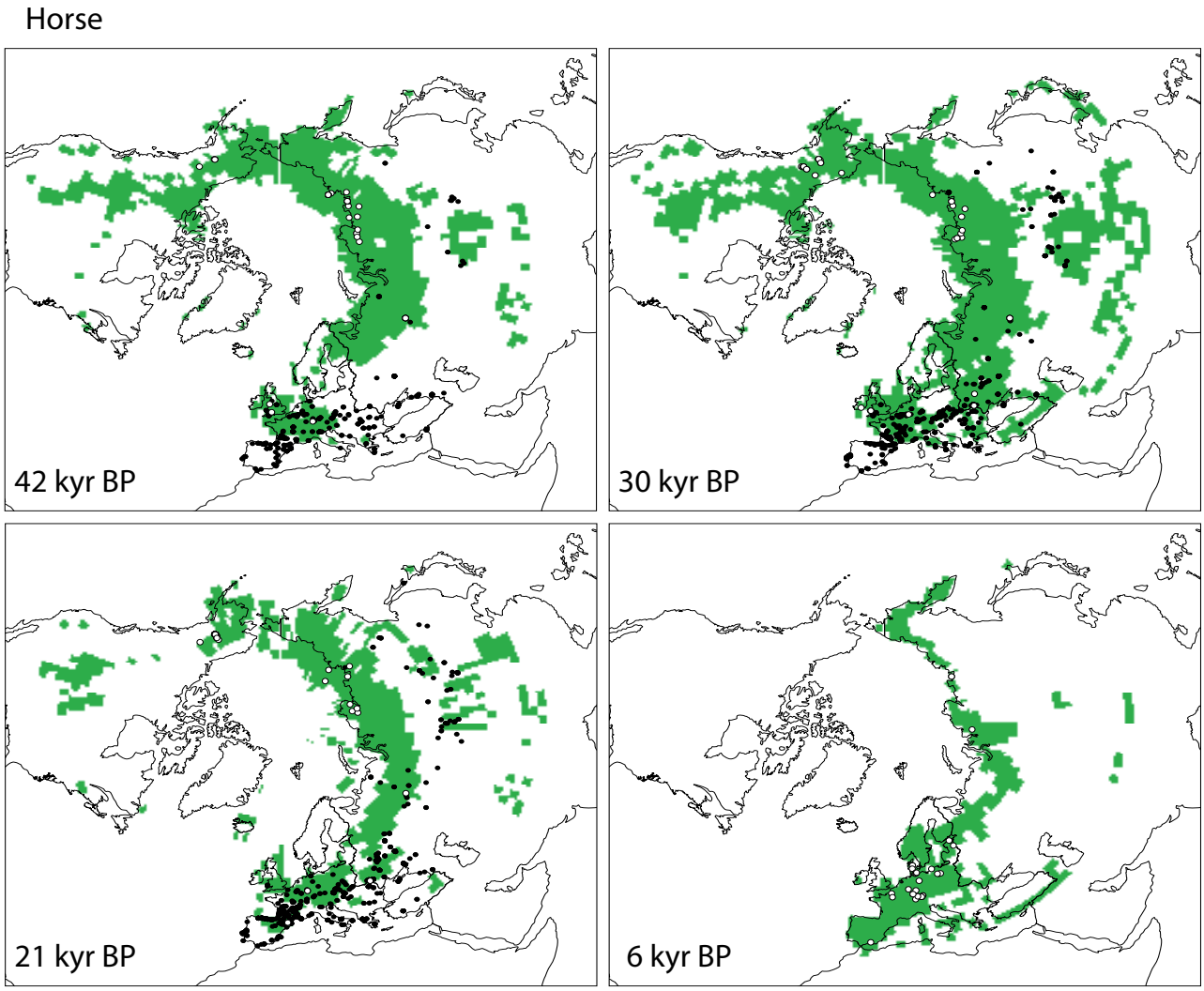
For woolly mammoth, all known fossils dating from 9–3 kyr BP are from Wrangel Island[49,50], with the exception of one known specimen from St. Paul Island[51]. This fossil distribution, consisting of two unique localities, is insufficient data with which to generate species distribution models, and so

no distribution is presented for the woolly mammoth at 6 kyr BP, even though the species was not yet extinct.

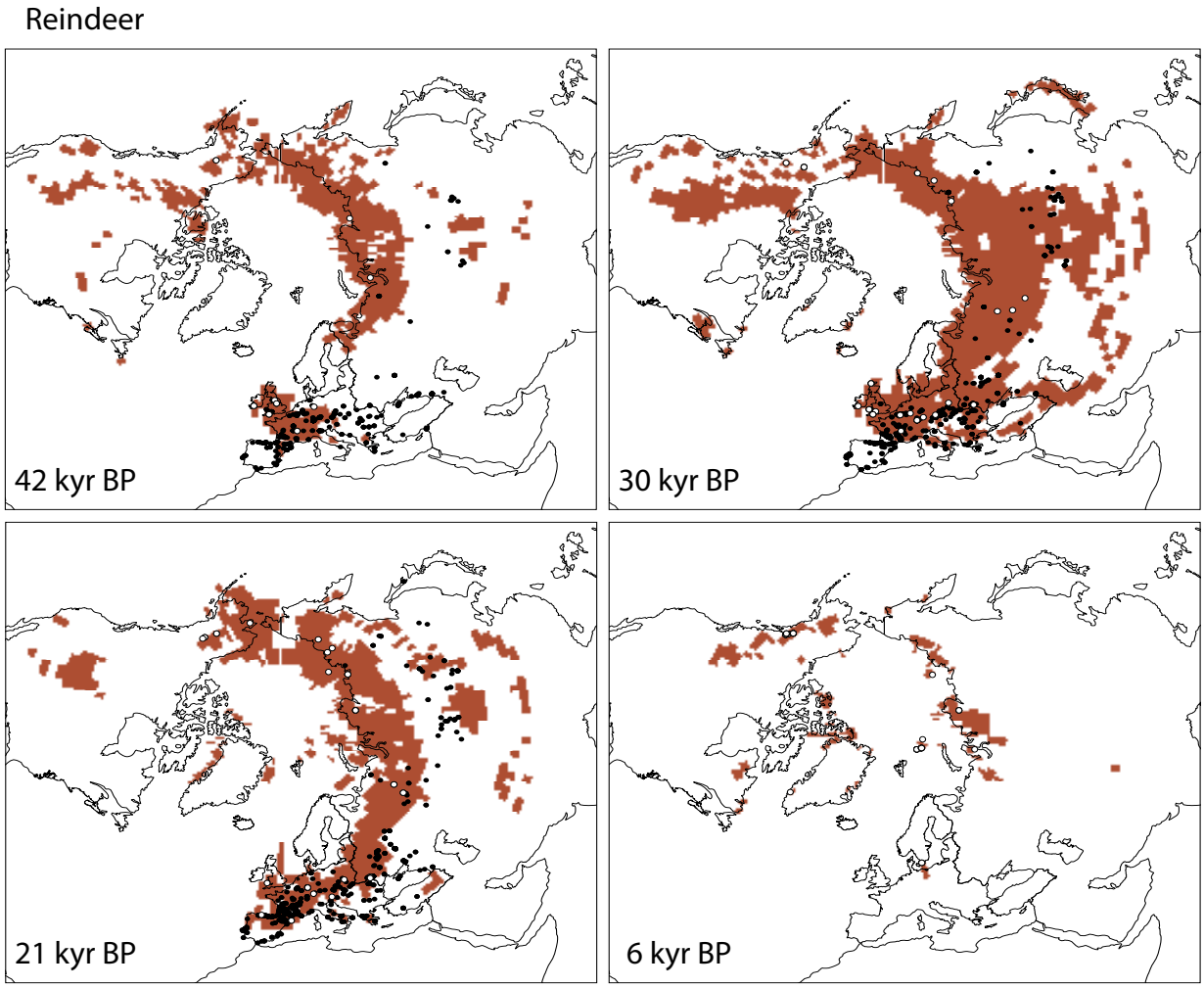


Supplementary Figure S1.3. Megafauna potential range at 42, 30, 21 and 6 kyr BP estimated from palaeoclimatic data (GENESIS2) and dated fossils for each species, represented by white dots.

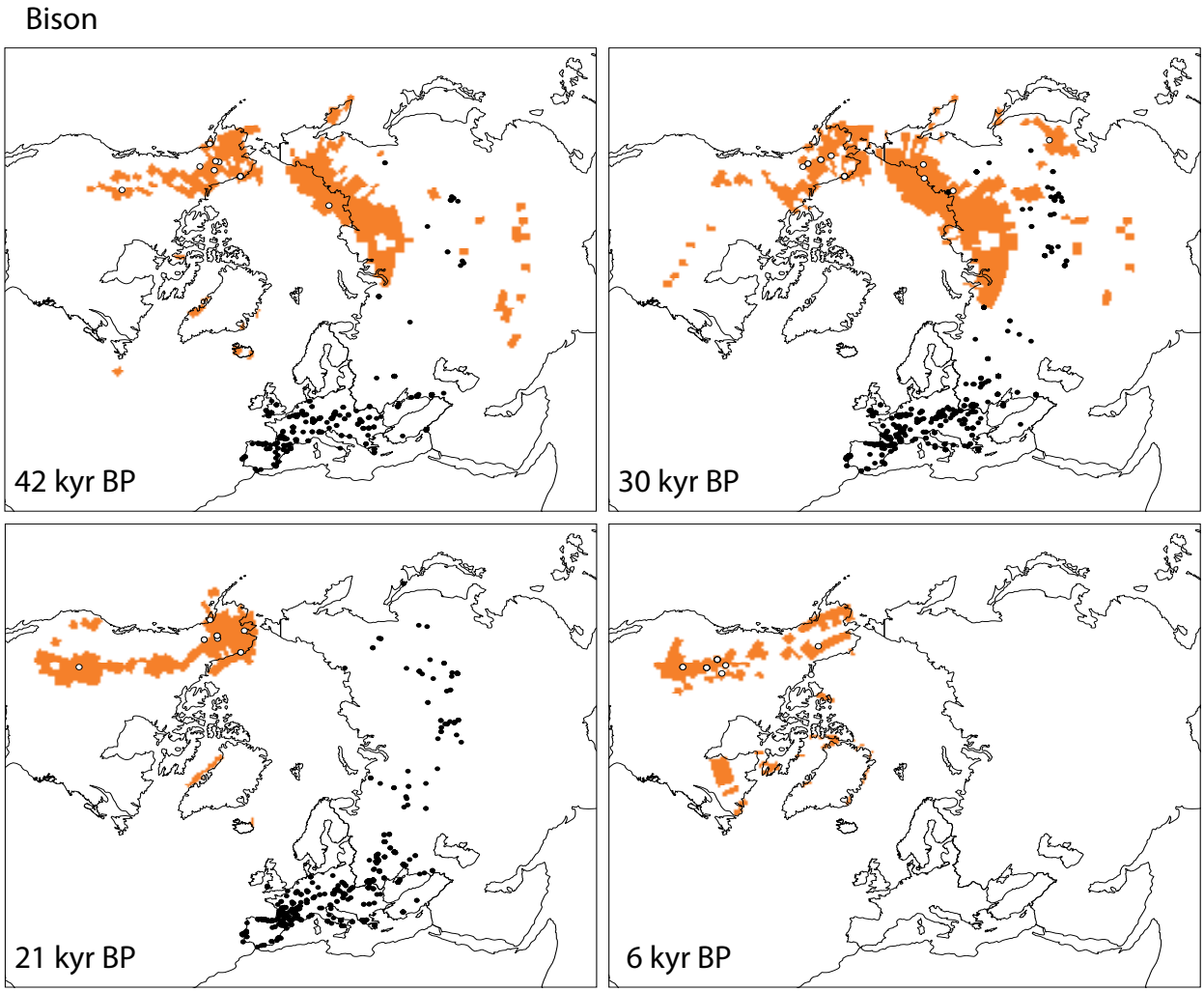
Range measurements were restricted to regions for which fossils were used to build the models, rather than all potentially suitable Holarctic area. Contemporaneous Palaeolithic human sites for each period are represented by black dots. No or too few fossils were available for woolly rhinoceros and mammoth to estimate their ranges at 6 kyr BP.



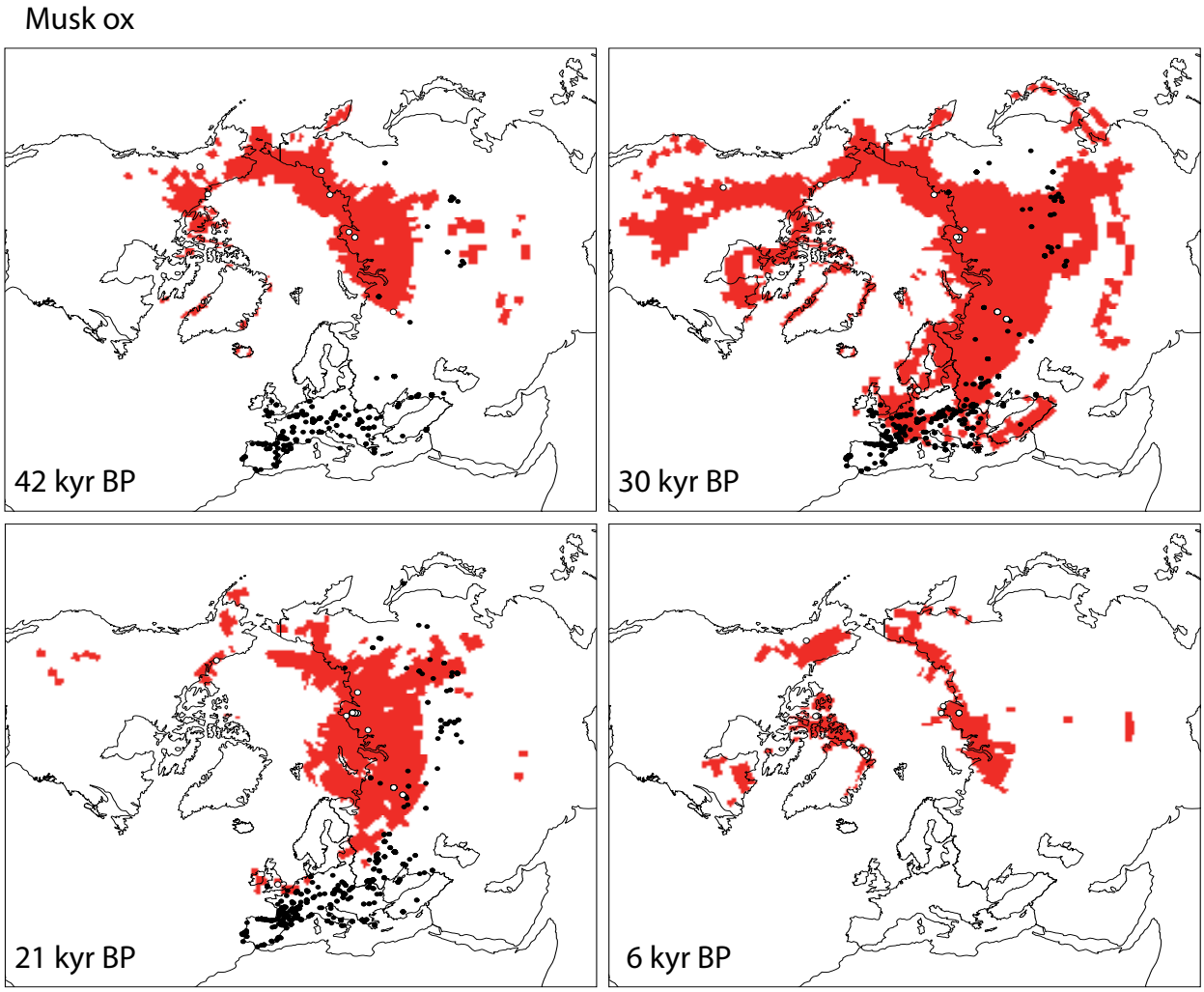
Supplementary Figure S1.3. Continued.



Supplementary Figure S1.3. Continued.



Supplementary Figure S1.3. Continued.



Supplementary Figure S1.3. Continued.

1.4 Species distribution modelling

1.4.1 Model algorithms

Species distribution model (SDM) projections are sensitive to the different statistical techniques used to describe and project species' potential ranges. Mahalanobis Distance (MD)[52], a strict presence-only method, was used to model the 42, 30, 21 and 6 kyr BP distributions of each species. MD is a simple envelope technique that provides a strict presence-only measure of environmental distance, which is calculated in relation to an optimum climatic point, defined as the centroid for all occurrence points in the total climatic space. The distance between this "optimum" and the observed climatic values for each species presence is inverse to the suitability of the climate at that site. MD produces an ellipsoidal envelope around the climatic optimum space by taking into account the covariance among climatic variables.

While a variety of methods are available for modelling species distributions, certain characteristics of the fossil record help to narrow the range of algorithms from which to select. First, the fossil record provides information about the presence of species, but not about their absence, and so presence-absence algorithms must be discarded. Other well-known, more complex algorithms such as GARP[53] or Maxent[54] could be used because they generate pseudo-absences against which to test the models. However, different procedures for calculating pseudo-absences can yield significantly different modelled species ranges[55,56]. In addition, bias in the fossil record is not simply a result of sampling effort, as with extant species, but of unevenly distributed geomorphological conditions affecting the fossilisation and persistence of remains through time, making selection of the study area on which to calculate pseudo-absences far more challenging for fossil data than for extant species for which the extent of the distribution is often known to some degree[55]. After these methods are discarded, the remaining suitable algorithms are restricted to simple presence-only methods based on environmental distances to a climatic optimum, which have been shown to handle bias in the fossil record better than more complex algorithms[57]. Specifically, MD has been shown to perform better than other presence-only methods in a recent comparative study[58]. It has been successfully used for palaeobiology studies[3,57,59] and is specifically recommended[57] for modelling potential species distributions using the fossil record.

1.4.2 Model implementation and performance

Species' distributions used to estimate range size (see below) were modelled using only locality and climate data from the same time periods (42, 30, 21 and 6 kyr BP). A minimum of five dated fossil localities per species/time period were used to build the models[60] (Supplementary Table S1.1). Modelling was implemented using the openModeller cross-platform modelling interface[61]. Each species/time period was modelled with both methods using all available locality data to build the model using both GENESIS2 and HadCM3.

To assess the impact of using a limited subset of the known fossil record on model performance for each species, we performed ten independent model runs in which a different randomly selected 75% of the data were used to build the model and 25% were used to test it each time. This evaluation does not assess the accuracy of model predictions—independent evaluation data (e.g., genetic data or additional fossils) would be required for this purpose. Rather, it provides a measure of internal consistency among repeated runs. Model performance was assessed using the Area Under the (Receiver Operating Characteristic) Curve (AUC; see [62] for a review on the advantages and disadvantages of using the AUC as a performance measure). Scores >0.75 are typically considered adequate for species distribution modelling[63]. Random sub-set model runs were performed using GENESIS2. Species distribution models yielded consistently high AUC values under testing, with only two species/time period yielding a coefficient of variation in AUC greater than 5% (bison and musk ox, 42 kyr BP; Supplementary Table S1.2). This suggests that model performance was relatively robust.

Supplementary Table S1.1. Total number of samples used to build the SDM for each species/ time period. Some localities had multiple dated samples, and the number of unique localities is given in parentheses.

	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP
Woolly rhinoceros	35 (29)	34 (28)	27 (24)	n/a
Woolly mammoth	78 (50)	112 (77)	97 (78)	n/a
Wild horse	35 (21)	42 (29)	53 (24)	32 (26)
Reindeer	16 (10)	33 (23)	35 (12)	46 (18)
Bison	15 (8)	16 (9)	10 (7)	15 (7)
Musk ox	8 (7)	25 (10)	49 (10)	16 (12)

Supplementary Table S1.2. AUC scores for ten model runs using a unique 75% and 25% of the data for building and testing the model, respectively (GENESIS2). Mean AUC, standard deviation and coefficient of variation (CV) are indicated.

	Woolly rhinoceros				**Woolly mammoth**			
	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP
run01	0.97	0.95	0.97	n/a	0.98	0.97	0.95	n/a
run02	0.96	0.97	0.97	n/a	0.98	0.98	0.96	n/a
run03	0.97	0.95	0.98	n/a	0.97	0.95	0.97	n/a
run04	0.94	0.97	0.98	n/a	0.95	0.96	0.96	n/a
run05	0.96	0.94	0.97	n/a	0.94	0.96	0.97	n/a
run06	0.96	0.96	0.98	n/a	0.9	0.96	0.96	n/a
run07	0.95	0.97	0.96	n/a	0.97	0.95	0.98	n/a
run08	0.97	0.98	0.98	n/a	0.92	0.97	0.95	n/a
run09	0.93	0.97	0.98	n/a	0.94	0.97	0.94	n/a
run10	0.96	0.97	0.99	n/a	0.98	0.96	0.95	n/a
Mean	0.96	0.96	0.98	n/a	0.95	0.96	0.96	n/a
StDev	0.01	0.01	0.01	n/a	0.03	0.01	0.01	n/a
CV	1.40%	1.30%	0.86%	n/a	2.93%	0.99%	1.25%	n/a

	Wild horse				**Reindeer**			
	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP
run01	0.98	0.97	0.98	0.97	0.98	0.98	0.96	0.98
run02	0.99	0.98	0.94	0.99	0.96	0.97	0.98	0.98
run03	0.99	0.98	0.96	0.99	0.95	0.98	0.94	0.98

run04	0.98	0.98	0.98	0.99	0.99	0.98	0.98	1
run05	0.99	0.96	0.99	0.98	0.95	0.91	0.98	1
run06	0.98	0.97	0.96	0.96	0.96	0.97	0.97	0.99
run07	0.99	0.96	0.95	0.98	0.99	0.98	0.97	0.99
run08	0.97	0.97	0.97	0.96	0.99	0.97	0.99	0.98
run09	0.99	0.97	0.99	0.99	0.94	0.98	0.96	0.98
run10	0.99	0.97	0.96	0.99	0.98	0.97	0.97	0.99
Mean	0.99	0.97	0.97	0.98	0.97	0.97	0.97	0.99
StDev	0.01	0.01	0.02	0.01	0.02	0.02	0.01	0.01
CV	0.72%	0.76%	1.74%	1.27%	1.97%	2.20%	1.46%	0.83%

	Bison				**Musk ox**			
	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP
run01	0.96	0.98	0.99	0.96	0.99	0.98	1	1
run02	1	0.99	0.94	0.91	0.97	0.98	0.99	1
run03	0.94	1	0.93	0.96	0.82	0.98	0.99	0.99
run04	0.82	0.99	0.87	0.95	0.98	0.99	0.99	0.93
run05	0.94	1	0.98	0.91	0.94	0.98	0.96	0.99
run06	0.95	1	0.94	0.96	0.83	0.98	0.95	0.97
run07	1	0.99	0.99	0.96	0.99	0.98	1	0.99
run08	1	0.98	0.98	0.94	0.96	1	0.94	0.98
run09	0.83	1	1	0.96	0.96	0.99	0.95	0.99
run10	0.93	0.89	0.98	0.99	0.97	0.99	0.99	0.99
Mean	0.94	0.98	0.96	0.95	0.94	0.99	0.98	0.98
StDev	0.06	0.03	0.04	0.02	0.06	0.01	0.02	0.02
CV	**6.92%**	3.39%	4.17%	2.58%	**6.69%**	0.72%	2.38%	2.09%

1.4.3 Measuring range size

To calculate modelled range size for each species/time period, the continuous suitability values mapped by openModeller were converted into deciles and the upper decile (suitability >0.9), the area of most suitable climate conditions, was used to map modelled range size (following a similar approach to [3]). Assuming that only the areas with the highest suitability constituted the potential range is a conservative approach which should prevent overestimation of range sizes. Because all species were not present throughout the Holarctic for all periods, the Holarctic was divided into 3 regions: Europe, Asia and North America. Modelled ranges were cropped to exclude non-land areas and to match those regions for which fossil localities were used to generate the models (e.g., if no fossils were available from North America for a given period, as for the woolly rhinoceros, climatically suitable range from North America was excluded from the range size estimate). This ensured that aDNA and species distribution modelling methods were testing hypotheses relating to

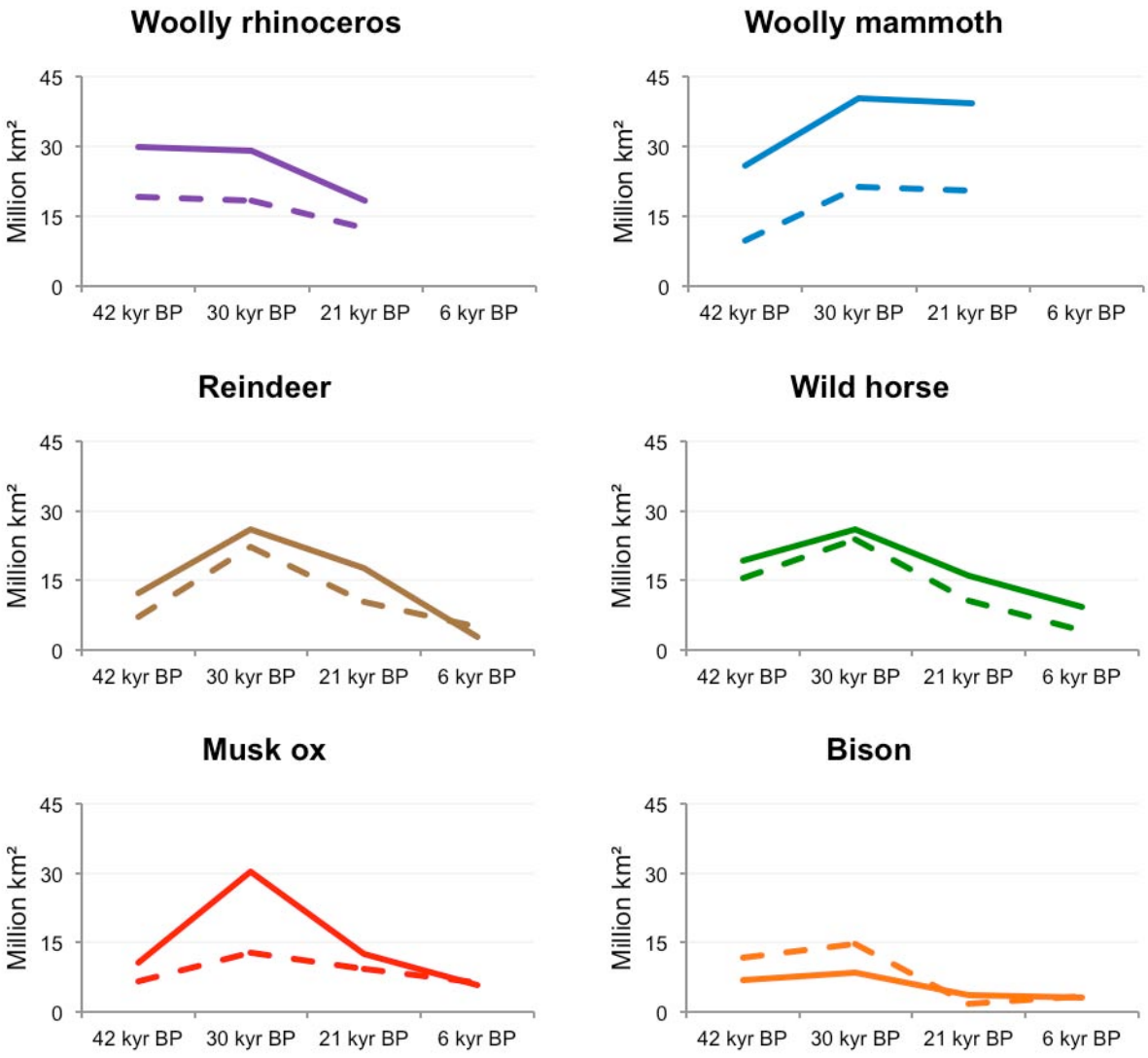
the geographic space actually occupied by each species at a given time as indicated by the fossil record. Cropping was conducted in R[64] using the package sp[65]. Species ranges were then measured (in square kilometres) using IDRISI Taiga (Clark Labs, Worcester, MA, USA).

Distributions for the six megafauna herbivores, reconstructed using SDMs for the periods 42, 30, 21 and 6 kyr BP, contracted in size from 30 kyr BP to the present for all species, although the severity of contraction varies substantially among taxa (Supplementary Table S1.3, Supplementary Fig. S1.3). While the absolute area of species ranges modelled using GENESIS2 and HadCM3 differ, as expected, trends of range expansion and contraction through time are consistent between AOGCMs (Supplementary Figure S1.4). Estimates of range size based on HaDCM3 also shows a significant correlation with estimated effective population size (Section S4; Supplementary Figure S4.3).

Supplementary Table S1.3. Area of potential species' range (climatic suitability ≥ 0.9.), rounded to the nearest 50,000 km^2.

Potential range modelled with GENESIS2				
	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP
Woolly rhinoceros	34,900,000	32,600,000	19,350,000	n/a
Woolly mammoth	26,050,000	42,900,000	39,500,000	n/a
Wild horse	19,400,000	26,250,000	16,000,000	9,200,000
Reindeer	12,250,000	26,250,000	17,800,000	2,750,000
Bison	6,800,000	8,700,000	3,700,000	2,950,000
Musk ox	12,300,000	27,750,000	22,550,000	9,250,000

Potential range modelled with HadCM3				
	42 kyr BP	30 kyr BP	21 kyr BP	6 kyr BP
Woolly rhinoceros	19,400,000	18,600,000	12,650,000	n/a
Woolly mammoth	12,150,000	25,900,000	22,600,000	n/a
Wild horse	15,450,000	24,000,000	10,650,000	4,100,000
Reindeer	6,950,000	22,350,000	10,250,000	4,900,000
Bison	11,800,000	14,900,000	1,850,000	3,350,000
Musk ox	6,650,000	12,900,000	9,200,000	6,150,000



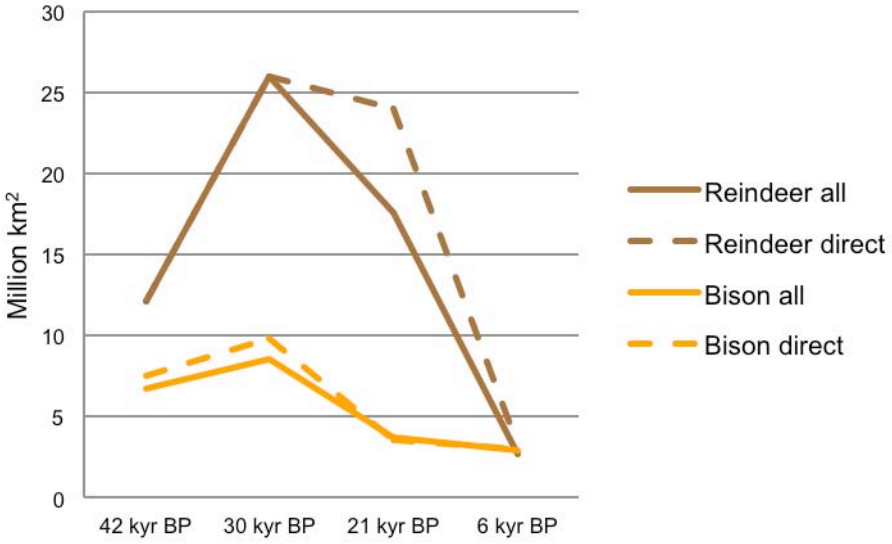
Supplementary Figure S1.4. Estimated potential range size in km^2 for all 6 species at 42, 30, 21 and 6 kyr BP, modelled using the GENESIS2 (solid line) and HadCM3 (dotted line) AOGCMs. While the absolute area of species ranges differ, temporal trends in range size are consistent between AOGCMs.

1.4.4 Sensitivity of range size estimates to fossil record uncertainty

Species Distribution Models are sensitive to the initial conditions used to calibrate the models. When modelling the past distributions of extinct and extant species, initial conditions include the historical climatic data and the distribution of the fossil record. Above we address the effect of using different AOGCMs for estimating range size, but the fossil record is an incomplete and often biased representation of the past distribution of species which can also bias our results Therefore, to incorporate this uncertainty in the estimation of potential range size into the correlation between effective population size and geographic range size (Supplementary Information section S4) we performed ten additional independent model runs using GENESIS2, in which a different randomly selected 90% set of the localities was used to build the model each time. Modelled ranges from the 90% random sub-set runs were cropped and measured as described in S1.4.2. We found a positive correlation between changes in the size of available habitat and genetic diversity for the four species for which we have range estimates spanning all four time-points (although the correlation was not statistically significant for reindeer: $p = 0.101$; Supplementary Information section S4).

1.4.5 Sensitivity of range sizes to radiocarbon dating error

Radiocarbon dates associated with indirectly-dated fossils are considered less reliable than those for which the specimen of interest itself is dated[48] although in some cases (e.g., reindeer, 21 kyr BP, Spanish localities) indirectly-dated fossils can represent important extensions of a species' geographic distribution for which directly-dated fossils are unavailable. Furthermore, for reindeer, ten indirectly-dated specimens from North America were included in the analysis, as the published DNA sequences from the samples[66] were included in the genetic analysis. To examine whether the incorporation of 16 indirectly-dated fossils (ten reindeer, six bison) is likely to have influenced the detected trends in range size through time, we re-ran the models for these species excluding the indirectly-dated specimens (from Supplementary Tables S6.1 and S6.4) using GENESIS2. While the absolute area of species ranges differ, as in the random sub-set model runs, consistent temporal trends of range expansion and contraction through time are detected with these 16 specimens included and excluded (Supplementary Figure S1.5). Incorporating these new measurements into the correlation analysis did not affect the strength of the correlation between range size and effective population size, which was still significant (Supplementary Table S4.).

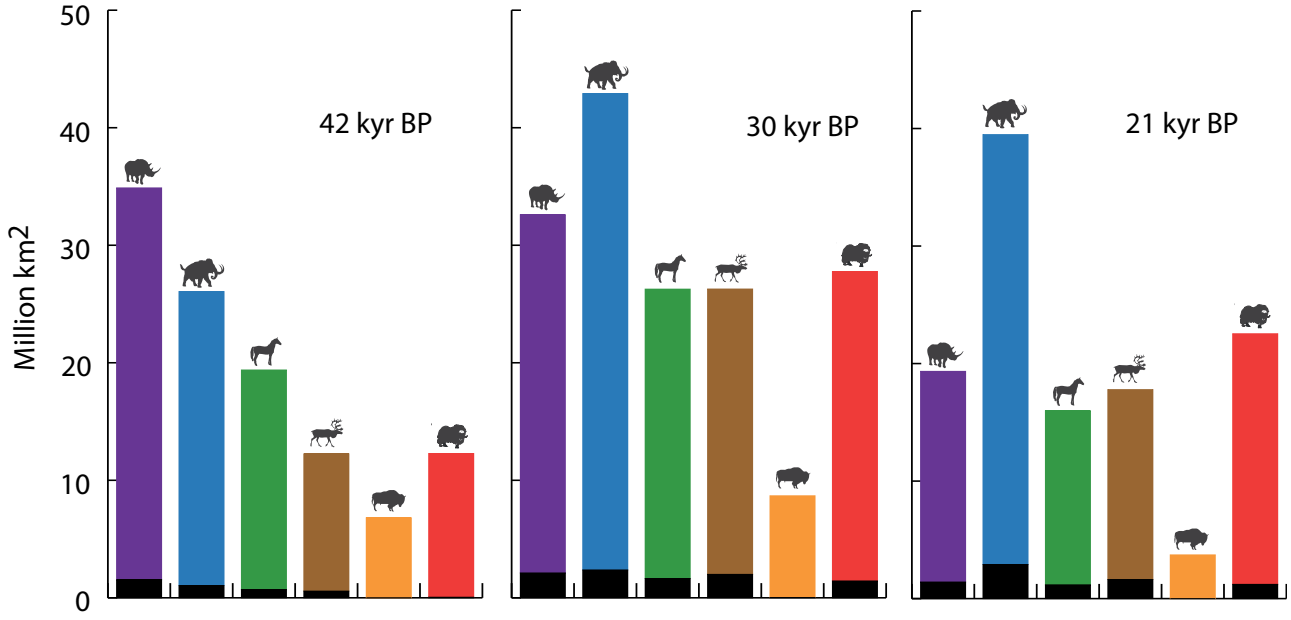


Supplementary Figure S1.5. Estimated potential range size in km^2 for reindeer and bison at 42, 30, 21 and 6 kyr BP, modelled using all (solid line) and only directly-dated (dotted line) fossil localities (GENESIS2). While the absolute area of species ranges differ, temporal trends in range size are consistent between fossil datasets used.

1.5 Human presence within modelled ranges

To calculate the density of Palaeolithic human fossil sites (bones, artefacts and charcoal) within the modelled geographic range for each species/time period, human fossil localities from the calendar time intervals 45–39, 33–27 and 24–18 kyr BP were overlaid on species' modelled ranges (GENESIS2) for 42, 30 and 21 kyr BP, respectively. Human radiocarbon data were derived from the INQUA Palaeolithic Radiocarbon Database v. 11 for Europe[67] and from [68] for Siberia. Details on these radiocarbon determinations and their selection can be found in the associated citations and in Supplementary Information section S5. The most recent period (6 kyr BP) was excluded because localities were only compiled for >9 kyr BP in some data sets.

Palaeolithic human localities (Supplementary Fig. S1.3) were mapped on top of SDM results in IDRISI Taiga to identify grid cells within each species' measured range in which humans were present. Grid cells were then converted into area (km^2) to calculate the extent of each species' range occupied by humans (Supplementary Fig. S1.6).



Supplementary Figure S1.6. Overlap (km^2) between megafauna and humans at 42, 30 and 21 kyr BP. Palaeolithic human localities were mapped on top of species ranges (GENESIS2), and grid cells in which humans were present were converted into area to calculate the extent of each species' range occupied by humans. Column height indicates estimates of megafauna range size; the black portion represents area of human/megafauna overlap.

1.6 Discussion

The key goal for this portion of our study was to relate changes in the megafauna species distributions to estimates of effective population size from the Bayesian skyride models (see Supplementary Information section S4). We used SDMs to estimate the potential range size for each species at four periods for which we have palaeoclimatic data, using a subset of the fossil record. The fossil list was not intended to be an exhaustive survey of all known locations for each species; rather, we targeted the fossil record for data within the same regions for which genetic data were sampled, so that SDMs and Bayesian skyride models could be explicitly compared. Modelled distributions are therefore unlikely to capture the full known distributions for some species (e.g., bison).

We have taken a conservative approach to modelling the distributions of each species; by using only contemporaneous data to build each model, rather than projecting a species' modelled climatic niche from one period to the next, we are able to avoid two potential pitfalls for SDMs: species-

climate equilibrium and climatic niche stability through time[14]. Species-climate equilibrium is the assumption that a species is in equilibrium with its climate, or that a species will exist in all places in which climatic conditions are favourable for its long-term survival[69]. However, many factors other than climate shape a species' ecological niche; these include barriers to dispersal, interactions with other species, and historical contingency[70], and any of these factors may result in a species' distribution being out of equilibrium with its climatic niche. For example, moist conditions on the Bering land bridge have been implicated as a barrier to dispersal to steppe-tundra species such as the woolly rhinoceros, which was adapted to drier conditions[71], even though suitable habitat likely existed on both sides of the strait. Likewise, niche stability through time is the assumption that a species maintains the same climatic niche, with no niche evolution (e.g., behavioural or physiological adaptation) taking place between periods of interest. While this assumption is likely to be true for some species, resulting in either extinction or tracking of suitable conditions under periods of climate change, it will not hold true for all species. Although our fossil data are limited for some species for certain periods, we maintain that building a discrete SDM for each time period, rather than projecting the distribution from those periods for which we have more data, is more relevant for comparison with the genetic data because it represents a species' realised distribution for a given time, rather than a potentially incomplete measure of the species' climatic niche.

SECTION S2: Megafauna ancient DNA extraction, amplification and sequencing

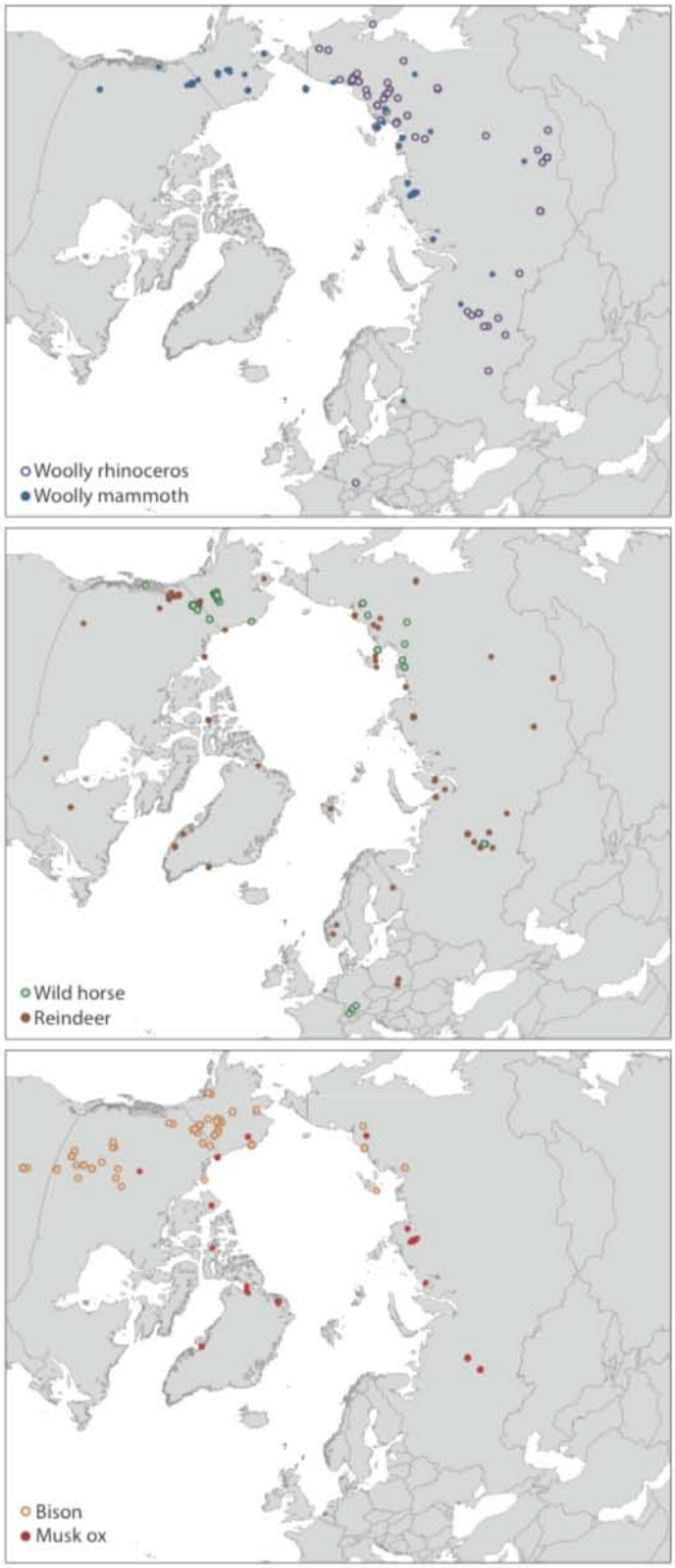
2.1 Megafauna samples

Ancient mitochondrial DNA control region (mtDNA CR) data sets were generated for woolly rhinoceros (*Coelodonta antiquitatis*), wild horse (*Equus ferus*; the fossil species *E. lambei* has been determined to be genetically indistinguishable from *E. ferus*, based upon [72], hence the latter name takes precedence), and reindeer (known as caribou in North America, *Rangifer tarandus*). Sub-fossil samples of bone, tooth and horn were collected across northern Eurasia and North America, including the Canadian Arctic Archipelago and Greenland (Supplementary Fig. S2.1, Supplementary Tables S6.2, S6.3, S6.4). Woolly rhinoceros was only sampled in Eurasia, as the species has never been found in the New World. Sequences data sets from woolly mammoth, bison and musk ox were downloaded from GenBank (Supplementary Information section S3).

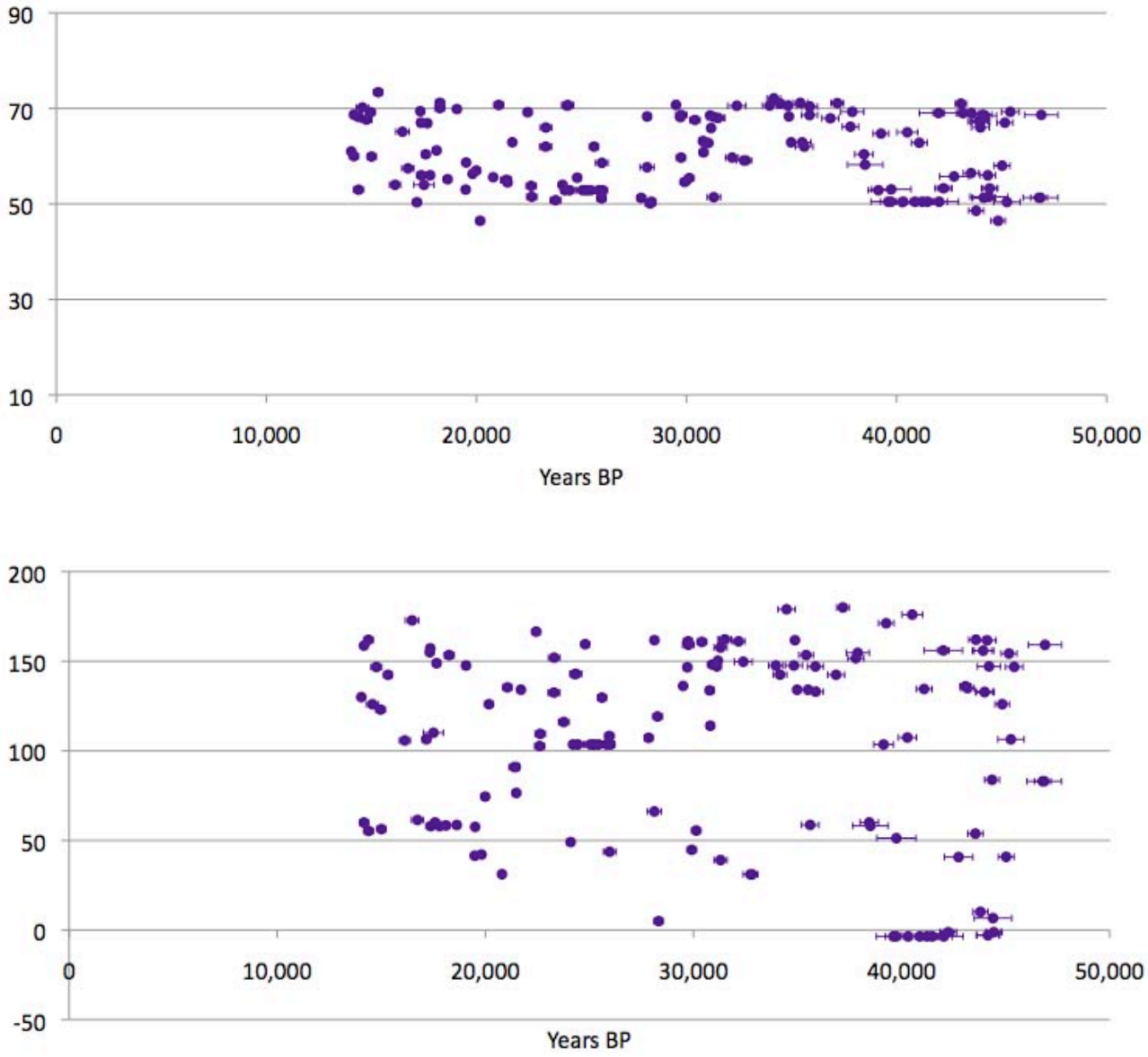
2.2 Accelerator Mass Spectrometry dating

A total of 353 Accelerator Mass Spectrometry (AMS) radiocarbon dates were obtained for woolly rhinoceros (n = 136), wild horse (n = 72) and reindeer (n = 145) from the commercial facilities offered by the Oxford Radiocarbon Accelerator Unit, UK (AMS ID: OxA), AMS 14C Dating Centre, Institut for Fysik og Astronomi, Aarhus University (AMS ID: AAR), Lawrence Livermore National Laboratory's Center for Accelerator Mass Spectrometry (AMS ID: CAMS), and NSF Arizona AMS Facility, Physics Department, University of Arizona (AMS ID: AA) (Supplementary Tables S6.2, S6.3, S6.4).

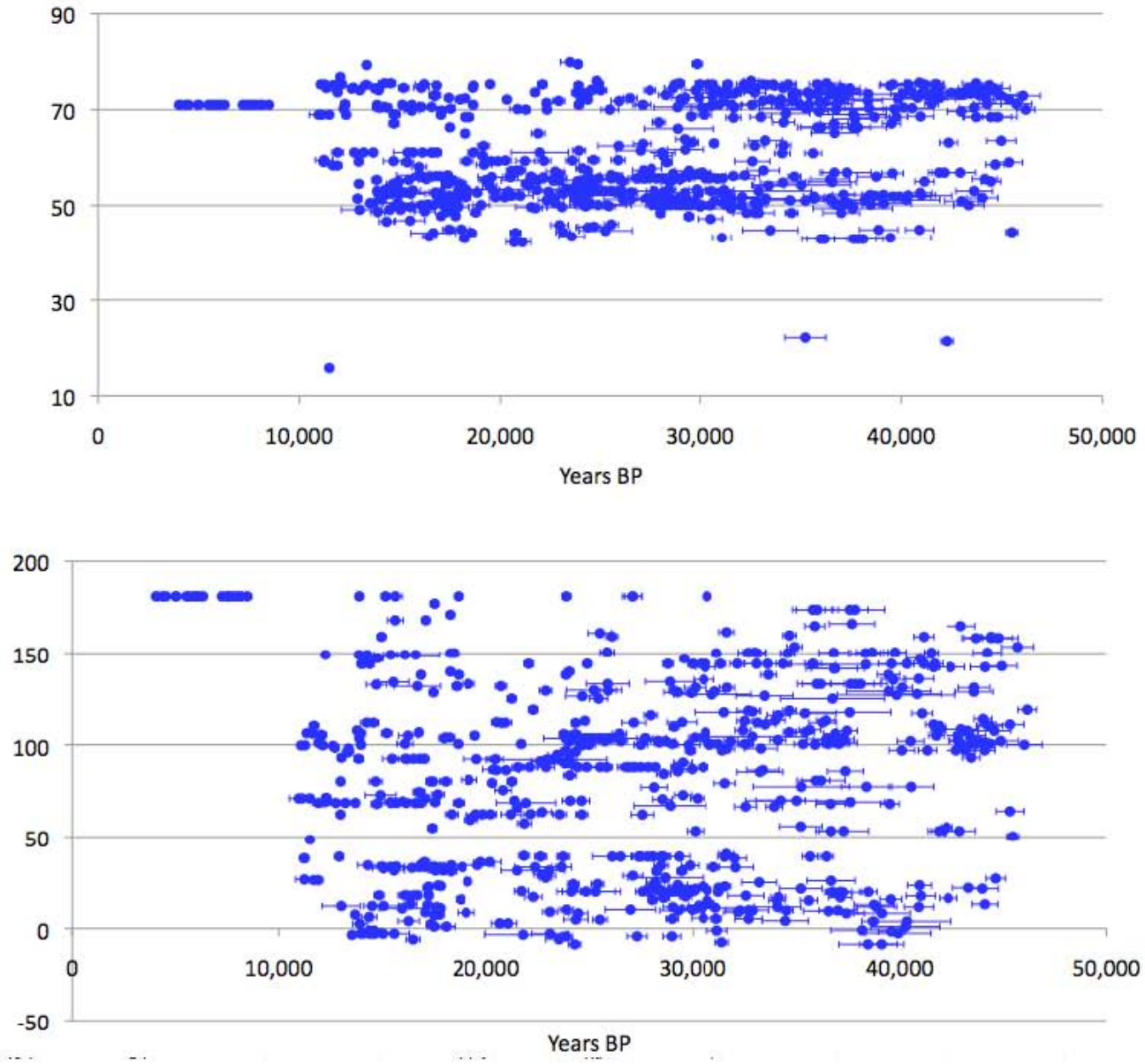
All radiocarbon dates, including those already published, were calibrated using the IntCal09 calibration curve[47] and the OxCal 4.1 online calibration resource (https://c14.arch.ox.ac.uk). Samples with infinite radiocarbon dates and radiocarbon dates past the IntCal09 calibration curve (*c.* >43,000 ^{14}C years before present, depending on the error of the date) were not included in the statistical analyses. Samples where the standard error of the calibrated date fell outside the calibration curve were omitted. Of the 353 new radiocarbon dates generated for this study, 16% were omitted from further analysis as their dates lay beyond the calibration curve or had infinite dates. All samples are discussed as kyr BP throughout the text, where kyr BP is defined as calendar thousand years before the present.



Supplementary Figure S2.1. Polar view of the Holarctic, with Eurasia to the right, indicating the DNA sample localities of the six megafauna species. New data sets were generated for woolly rhinoceros, horse and reindeer; further information in Supplementary Tables S6.2, S6.3, S6.4.



Supplementary Fig. 2.2. The woolly rhinoceros fossil record, including the 136 new radiocarbon dates generated for this study, plotted against latitude (top) and longitude (bottom) coordinates . Data are presented in Supplementary Tables S6.1 and S6.2. The species was present throughout Siberia right up until its disappearance from the fossil record. X-axis in calendar years BP, 1-sigma errors of the calibrated dates are included.



Supplementary Fig. 2.3. The woolly mammoth fossil record plotted against latitude (top) and longitude coordinates (bottom). The species was present throughout Siberia right up until its disappearance from the fossil record. X-axis in calendar years BP, 1-sigma errors of the calibrated dates are included.

2.3 Sequence generation

Stringent ancient DNA protocols were followed to avoid contamination from modern DNA and to assure reliability of results. All DNA extractions and PCR set-ups were performed in a dedicated ancient DNA facility isolated from multi-copy PCR work. PCR amplification, cloning and sequencing were performed at a separate DNA facility.

Ancient DNA sequences were obtained using the extraction procedures reported in [73].

DNA was PCR amplified using overlapping fragments ranging in length from 80–560 bp, depending on the condition of the specimen, and the species being sequenced. Primers were designed to span the entire HVR-1 of the mitochondrial control region. PCR amplifications were performed in 25 µl volumes, using 1xPCR buffer, 2 mM of $MgSO_4$, 2.0 mg/ml Bovine Serum Albumin (BSA), 0.4 µM of each primer, 1 µM of dNTPs, and 5U of High Fidelity Platinum Taq (Invitrogen, Carlsbad, CA). Cycling conditions were: 94˚C for 2 min, 50–70 cycles of 94˚C for 30 sec, 43–63˚C for 30 sec, and 68˚C for 45 sec, followed by 72˚C for 7 min. We included blank extraction controls and blank PCR controls in each reaction. Primer sequences and PCR-annealing temperatures are listed in Supplementary Table S2.1.

PCR products were subsequently purified with either the Invitek PCRapace PCR Purification Kit (Invitek, Berlin, Germany) or the QIAquick PCR purification kit (Qiagen, Valencia, CA), according to manufacturers' instructions. At least two independent PCRs were carried out for each fragment, and the products were either direct sequenced or cloned using TOPO TA cloning kit for sequencing (Invitrogen), with a minimum of six clones sequenced for each fragment. The overlapping of the PCR fragments resulted in a high degree of sequence replication. The sequences were obtained through the commercial service offered by Macrogen (Macrogen, Seoul, South Korea) and by in-house sequencing at Department of Biology, University of Copenhagen, Denmark. DNA sequences were subsequently edited by eye and aligned using Se-Al version 2.0A11 (A. Rambaut, University of Edinburgh). To investigate and account contamination and for errors caused by damage or sequencing, 79% of all consensus sequences were replicated.

The data sets resulted in 274 new mtDNA sequences, including 55 woolly rhinoceros, 115 wild horse and 104 reindeer. Sequences have been submitted to GenBank with the accession numbers JN570760-JN571033, corresponding sample numbers are listed in Supplementary Tables S6.2, S6.3 and S6.4.

Supplementary Table S2.1. Primer sequences and annealing temperatures for woolly rhinoceros, horse and reindeer.

Woolly rhinoceros	Primer name	Primer sequence (5'-3')	Annealing temp. (°C)
	1AF/WR15422F	CCCTAACTTCACCATCAACACCC	55
	1aF/WR15500F2	CACTCCCTTCTTAAACCASAAG	55
	1BF/WR15562F	TATACCAGGTATGTATATCG	55
	1AR/WR15600R	CATGCTTATATGCATGGGGC	55
	1CF2/WR15714F2	TTGATTRATATTGCATAGTAC	55
	1BR3/WR15727R3	GACTYRAATGGGGTATGTACG	55
	1cR/WR15792R	CGCGGCTTGGTGATTAAGCGC	55
	1CR/WR15812R	GAGAGGGTTGATGATTTCCC	55
	1bzF(1AR)	GCCCCATGCATATAAGCATG	55
	1bzR(1CF)	GTACTATGCAATATYAATCAAC	55
	2AF/WR16614F	GCCAAACCCCAAAAACAAG	55
	2aF/WR16635F	GACTAGGTATATAATTACACGC	55
	2BF/WR16704F2	CCCTTCTTTTGATACCAACATGC	55
	2AR3/WR16724R3	CTAGAGGGGTAYGAGTCTAYGTG	55
	2bR/WR16813R	GCTACATTAACAGGTGTATTTG	55

Horse	Primer pairs	Primer sequence (5'-3')	Annealing temp. (°C)
	L1	GCCATCAACTCCCAAAGCT	
	H1	ACATGCTTATTATTCATGG	56
	L2	CCCACCTGACATGCAATAT	
	H2	TGTTGACTGGAAATGATTTG	56
	L3	TCGTGCATACCCCATCCAA	
	H3	CCTGAAGTAGGAACCAGATG	56
	L4	CCATGAATAATAAGCATGT	
	H3	CCTGAAGTAGGAACCAGATG	56
	EQ168-187F	CGTGCATTAAATTGTTTGCC	
	EQCR2bii	CATGGGAGGTGATATGCGTG	56
	EQCR3ai	CGTGCATACCCCATCCAAGTC	
	EQCR3bi	GAACCAGATGCCAGGTATAG	56
	EQCR1F	TCCTCGCTCCGGGCCCAT	
	EQCR136R	TGTGAGCATGGGCTGATTAGTC	60
	EQCR51F	CTGGCATCTGGTTCTTTCTTCAG	
	EQCR210R	CTTTGACGGCCATAGCTGAGT	60
	EQCR163F	ACTGTGGTTTCATGCATTTG	
	EQCR296R	TTGCTGATGCGGAGTAATAA	56
	EQCRend184F	ATCTTGCCAAACCCCAAAAACAAG	
	EQCRend342R	TCTAGGGGGATGCCTGTCTATGG	63
	EQ4F	CATCAACACCCAAAGCTGAA	
	EQ4R	CGAYGTACATAGGCCATTCAT	56
	EQ5F	CATACCCACCTGACATrCAA	
	EQ5R	GACTTGGATGGGGTATGCAC	56

Reindeer	Primer pairs	Primer sequence (5'-3')	Annealing temp. (°C)
	CP1_F	GTCAACATGCGTATCCCG	51
	CP2_R	RTGAGATGGCCCTGAAGAAA	
	Rtp2_F	TCTCCCTAAGACTCAAGGAAG	
	Rtp2_R	GGCTATTGAGTGCAGAACTG	48

Rtp3_F	TCCACAAAATTCAAGAGCCTT	
Rtp3_Rshort	TAGCCGTACAGGACCATA	50
98F	AAGTTCTAATTAAACTATTCCCTG	43
231R	ATATAATATGGCTATTGAGTGC	
67F	TATAGCYCCACTATCAACACCC	47
228R	ATATAAYATGGCTATTGAGTGC	
38F	CCAATCTCCCTAAGACTCAAGG	49
219R	CTGTATTAAATTHTTRAAGGTTTTTRGA	
172F	AAAAACCTTYAADAATTTAATACAGT	46
350R	TGGGRYATRTARTTTAATGTACTATTAT	
186F	CCTTCARGAATTTAATACAGTTCTGC	52
373R	CARGTACTTGCTTATAAGCATGGGG	
272F	GGTCCTGTACGRYTATAGTAC	40
476R	CCCCTAGATCACGAGCT	
442_F	GYCAACATGCGTATCCCG	50
603_R	GCCCTGAAGAAAGAACC	