

The Genomic Legacy of Aurochs hybridization in ancient and modern Iberian Cattle

Reviewed Preprint

v2 • January 7, 2025

Revised by authors

Reviewed Preprint

v1 • January 10, 2024

Torsten Günther , Jacob Chislausky, M Ángeles Galindo-Pellicena, Eneko Iriarte, Oscar Cortes Gardyn, Paulina G Eusebi, Rebeca García-González, Irene Urena, Marta Moreno, Alfonso Alday, Manuel Rojo, Amalia Pérez, Cristina Tejedor Rodríguez, Iñigo García Martínez de Lagrán, Juan Luis Arsuaga, José-Miguel Carretero, Anders Götherström, Colin Smith , Cristina Valdiosera 

Human Evolution, Department of Organismal Biology Uppsala University, Uppsala, Sweden • Centro Mixto UCM-ISCIII de Evolución y Comportamiento Humanos, Madrid, Spain • Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain • Universidad Complutense Madrid, Madrid, Spain • Centre for Palaeogenetics, Stockholm, Sweden • Instituto de Historia – CSIC, Madrid, Spain • Área de Prehistoria, University of the Basque Country, Leioa, Spain • Department of Prehistory and Archaeology, Valladolid University, Valladolid, Spain • Departamento de Prehistoria y Arqueología, UNED, Madrid, Spain • Unidad Asociada de I+D+i al CSIC Vidrio y Materiales del Patrimonio Cultural (VIMPAC), Burgos, Spain • Dept. of Archaeology and History, La Trobe University, Melbourne, Australia

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

eLife Assessment

Using genomic data from ancient and modern samples, this **important** study investigates the genomic history of cattle in Iberia, focusing on the admixture between domestic cattle and their wild ancestors, aurochs. The authors present **convincing** evidence for interbreeding between domestic cattle and wild aurochs since the Neolithic period, although the evidence of sex-biased introgression is weak. The authors also show that the aurochs ancestry in cattle stabilized at ~20% since ~4000 years ago and continues into modern breeds; however, the aurochs ancestry is not heightened in a modern breed of Spanish fighting bulls that are bred for aggressiveness. The work will be of interest to evolutionary biologists and quantitative geneticists who seek to understand the genomic history and genetic basis of trait variation of domesticated animals.

<https://doi.org/10.7554/eLife.93076.2.sa3>

Abstract

Cattle have been a valuable economic resource and cultural icon since prehistory. From the initial expansion of domestic cattle into Europe during the Neolithic period, taurine cattle (*Bos taurus*) and their wild ancestor, the aurochs (*B. primigenius*), had overlapping ranges, leading to ample opportunities for mating (whether intended by farmers or not). We performed a bioarchaeological analysis of 24 *Bos* remains from Iberia dating from the Mesolithic to the Roman period. The archaeogenomic dataset allows us to investigate the

extent of domestic-wild hybridization over time, providing insight into the species' behavior and human hunting and management practices by aligning changes with cultural and genomic transitions in the archaeological record. Our results show frequent hybridization during the Neolithic and Chalcolithic, likely reflecting a mix of hunting and herding or relatively unmanaged herds, with mostly male aurochs and female domestic cattle involved in hybridization. This is supported by isotopic evidence consistent with ecological niche sharing, with only a few domestic cattle possibly being managed. The proportion of aurochs ancestry remains relatively constant from about 4000 years ago, probably due to herd management and selection against first generation hybrids, coinciding with other cultural transitions. The constant level of wild ancestry (~20%) continues into modern western European breeds including the Spanish Lidia breed which is bred for its aggressiveness and fighting ability, but does not display elevated levels of aurochs ancestry. This study takes a genomic glance at the impact of human actions and wild introgression in the establishment of cattle as one of the most important domestic species today.

Introduction

Domestication of livestock and crops has been the dominant and most enduring innovation of the transition from a hunter-gathering lifestyle to farming societies. It represents the direct exploitation of genetic diversity of wild plants and animals for human benefit. Ancient DNA (aDNA) has proved crucial to understanding the domestication process and the interaction between domesticated species and their wild relatives both within domestication centers and throughout the regions that the domestics expanded into (1–14). The origins of the European domestic taurine, *Bos taurus*, are located in the Fertile Crescent (15, 16) and unlike dogs, pigs and goats, where the wild forms are still extant, the wild cow (the aurochs) went extinct in 1627. Aurochs, *B. primigenius*, was present throughout much of Eurasia and Africa before the expansion of domestic cattle from the Levant that accompanied the first farmers during the Neolithisation of Europe. Upon arrival, these early incoming domesticates inevitably coexisted with their wild counterparts in great parts of Europe facilitating gene flow in both directions. In general, taxa within the genus *Bos* can hybridize and produce fertile offspring (17) which may have facilitated and contributed to domestication, local adaptation and even speciation (5, 18–20). Mitochondrial DNA studies have previously indicated gene flow between domestic cattle and aurochs outside their domestication center (21–25) and more recently, genomic studies have shown the presence of European aurochs ancestry in modern taurine cattle breeds (26–28). Although cattle have represented a significant economic resource and a prominent cultural icon for millennia, and have been studied through aDNA for more than a decade (5, 21, 22, 26, 28–30), our understanding of the interaction of early cattle herds and wild aurochs is still limited due to a lack of time-series genomic data. This gap of knowledge includes European aurochs' genetic contribution to modern domestic breeds and human management of these animals in the past.

Aurochs have been widely exploited by humans since the European Palaeolithic and archaeological evidence indicates that the species survived in Europe until historical times. Iberia could have served as a glacial refugium for aurochs (28), and the most recent evidence for aurochs is found at a Roman site in the Basque Country (31). Domestic cattle were introduced into Iberia with the Mediterranean Neolithic expansion and reached the northern coast of the peninsula around 7000 years cal BP (32). Consequently, aurochs and domestic cattle have coexisted in Iberia for about five millennia. Since then, cattle have played an important role in Iberian societies as a source of food and labour, as well as cultural events such as bullfighting. Currently, there are more than 50 bovine breeds officially recognized in the Iberian Peninsula including the Lidia breed, a primitive, isolated population selected for centuries to develop agonistic-aggressive responses with the exclusive purpose of taking part in such socio-cultural events (33). Recently, it has been reported that Lidia breed individuals have the largest brain

size among a comprehensive data set of European domestic cattle breeds and are the most similar to wild aurochs (34 [↗](#)). The combination of aggressiveness and larger brain size in the Lidia breed may suggest a higher proportion of aurochs ancestry compared to other cattle breeds.

Here, we present the genomes and stable isotope data of Iberian Bovine specimens ranging from the Mesolithic into Roman times from four archeological sites (**Fig. 1A** [↗](#)). We explore the extent of interbreeding between wild aurochs and domestic cattle over time and the correlation of genetic ancestry with metric identification and ecological niches. Finally, we compare the results to genomic data obtained from modern Iberian cattle breeds to estimate the genetic contribution of the now-extinct aurochs to the Iberian farming economy.

Results

Exploratory analysis

We successfully sequenced 24 Bovine specimens excavated at four prehistoric sites in Iberia (**Fig. 1A** [↗](#)). Nine of these individuals were inferred to or suspected to represent aurochs based on morphology or chronology. Direct radiocarbon dates and contextual dating placed the individuals between the Mesolithic (oldest sample moo001, 8641-8450 cal BP) and the Roman Age (youngest sample moo010p, 2260-2150 cal BP). It should be noted that while all post-Mesolithic samples were found at archaeological sites with evidence for herding of other domestic fauna such as ovicaprids (Supplementary Information), we do not know whether these bovids were herded or hunted. Based on the number of reads mapping to the X chromosome, 13 individuals were identified as female and 11 as male, for the samples with sufficient amounts of reads for this analysis. Sequencing coverage of the nuclear cattle genome was low to medium, reaching up to 4.7x with a mean of 0.38x (Dataset S1). The sequence data for non-UDG-treated libraries showed damage patterns characteristic of ancient DNA (Supplementary Information, Figure S1). Based on reads mapping to the mitochondrial genome, we were able to estimate contamination for 11 samples with most samples showing low levels (<2%) of contamination. One individual (moo013a) showed a high contamination of 51.7% [33.5, 69.9] and was excluded from further analysis. One Mesolithic individual (moo040) showed 8.3% [3.1, 13.5] contamination, which was included in the initial exploratory analysis but not used for the analysis of hybridization between wild and domestic as this analysis only focused on post-Mesolithic individuals.

Nine individuals were assigned to the mitochondrial P1 haplogroup, one to haplogroup T1, one to haplogroup T4, and 12 to haplogroup T3 (Dataset S1). P haplogroups are dominant among and thought to be endemic to European aurochs (35 [↗](#)), but are occurring at low frequencies in modern European cattle breeds (24 [↗](#)). The prevalence of the T3 haplogroup in our samples is expected; this haplogroup is dominant among modern European *B. taurus* and is the most common haplogroup in ancient Western European domestic cattle. T3 was found in directly dated Neolithic samples from different sites providing direct evidence for the arrival of domestic cattle in northern Iberia during the Neolithic (Dataset S1). The specimen assigned to T1 (moo009a) is notable since this individual was previously used to argue for Bronze Age contact between Iberia and Africa, where the T1 haplogroup is thought to have originated (29 [↗](#), 36 [↗](#)). T4 is usually considered to be restricted to Asian breeds with rare finds in Europe, restricted to the Balkans (24 [↗](#)). The presence of T4 in Chalcolithic Iberia suggests that this haplogroup must have been distributed across Western Europe at low frequencies in prehistory. Furthermore, the fact that some specimens that were morphologically identified as aurochs carry domestic T haplogroups implies some level of interbreeding between the two groups.

As mitochondrial genomes only reflect the maternal line of ancestry, they are not informative about the exact extent of interbreeding in our dataset. To avoid being constrained by the variation in modern domestic breeds as with common approaches such as projected PCA (Supplementary

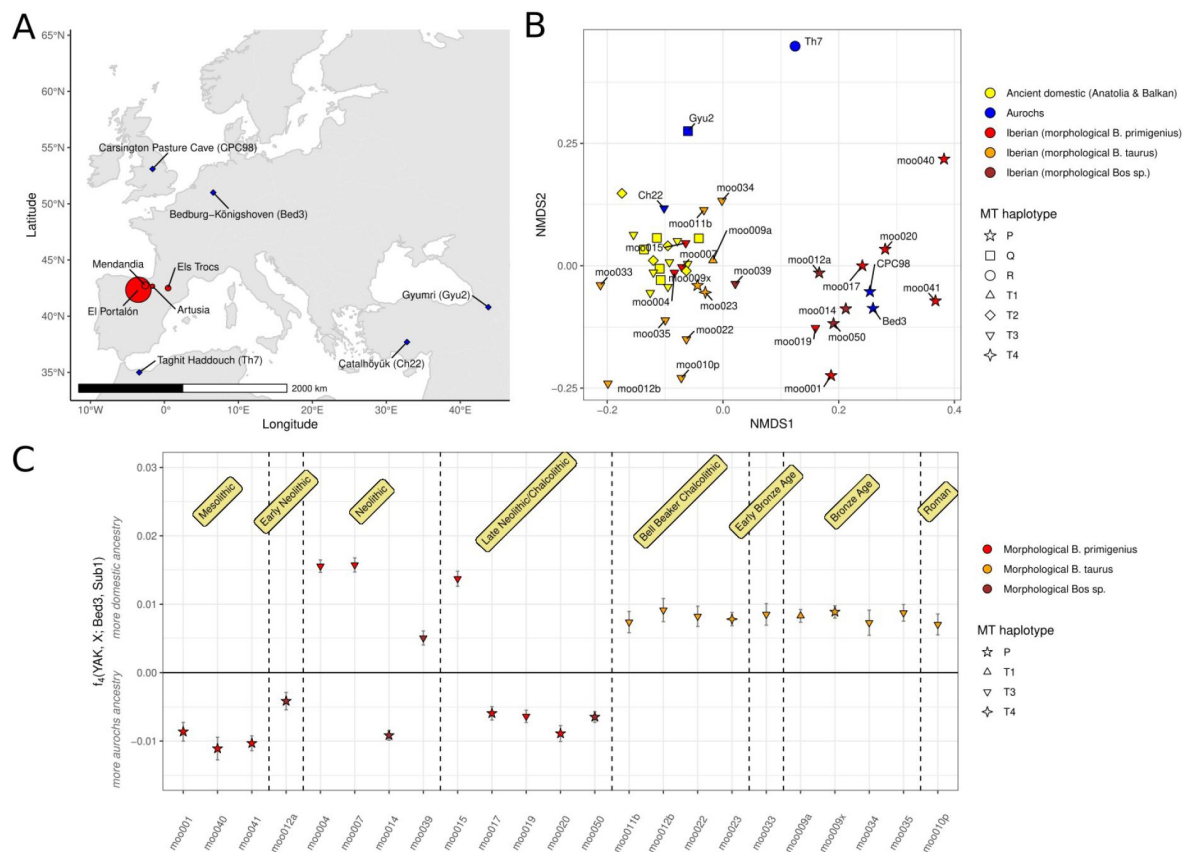


Figure 1

(A) Map of Europe showing the Iberian sampling sites (red circles, size proportional to sample sizes) and the sites for published aurochs genomes used in the analysis (blue diamonds). (B) NMDS ordination of nuclear data of Iberian samples considered *B. taurus* samples (orange), Iberian samples that were morphologically considered *B. primigenius* (red), other Iberian *Bos* samples (brown), ancient domestic cattle from the Balkans and Anatolia (yellow), and aurochs (blue). Data point shape corresponds to mitochondrial haplogroups. (C) f_4 statistic measuring allele sharing of the Iberian samples with European aurochs (Bed3) or Anatolian Neolithic cattle (Sub1). The time periods displayed are contextual.

Information, Figure S4), we performed non-metric multi-dimensional scaling (NMDS) ordinations on a matrix of pairwise outgroup f_3 statistics to explore the genomic ancestry of the sequenced individuals. For reference, we included early cattle genomes from Anatolia and the Balkans as well as aurochs excavated from Morocco (Th7), Armenia (Gyu2), Anatolia (Ch22), Germany (Bed3) and Britain (CPC98) (5, 26, 30), and calculated pairwise outgroup f_3 statistics. The NMDS ordination outcome (Fig. 1B) seems to represent a separation between domestic autosomal ancestry (to the left) and European aurochs ancestry (to the right). In contrast, aurochs from other regions (Th7 and Gyu2) seem genetically distinct. Many early domestic samples from Iberia fall close to early cattle from the Balkans and Anatolia as well as the Anatolian aurochs (Ch22). Notably, at least two of the Iberian samples in this cluster (moo004, moo007) were morphologically identified as aurochs. Eight of the nine Iberian samples with haplogroup P fall to the right in the plot, together with the aurochs from Germany (Bed3) and Britain (CPC98). Additionally, one individual carrying a domestic T3 mitochondrial genome (moo019) appears closer to the aurochs samples than the domestics. Out of nine samples that were presumed aurochs based on their morphological features, only six would be considered aurochs based on this analysis. This highlights a substantial overlap between measurements or criteria that are used to distinguish wild and domestic *Bos* based on morphometrics.

This analysis suggests that one can use other European aurochs such as the German Bed3 (Bedburg-Königshoven, 11802-11326 CalBP) (30) or the British CPC98 (Carsington Pasture Cave, 6874-6602 CalBP) (26) as a reference for Western European aurochs as they seem similar to our three low-coverage Mesolithic Iberian samples. This is also supported by a recent parallel study concluding that all Western European aurochs form a clade, possibly even originating from an Iberian glacial refugium (28). Using Sub1 (Suberde Höyük, 8171-7974 CalBP) (5), a Neolithic domestic Anatolian individual, and the higher coverage aurochs Bed3 as references, we can perform f_4 statistics to measure which Iberian individuals share more alleles with one or the other (Fig. 1C). Despite the relatively low coverage of some samples, the f_4 statistics are highly correlated with the first axis of the NMDS ($R^2=0.84$, $p=8.3 \times 10^{-10}$) implying that they detect the same pattern. Non-overlapping confidence intervals also confirm that the high genetic differentiation between Western European aurochs and domestic cattle allows confident assignment even with low coverage data. The three Mesolithic individuals as well as an additional six, up until the Late Neolithic/Chalcolithic, share most of their alleles with aurochs. Three individuals from the Neolithic and Late Neolithic/Chalcolithic share most of their alleles with domestic Anatolian cattle while two individuals (moo012a and moo039) are more intermediate, suggesting that there could have been some level of hybridization. More recent samples from the Bell Beaker period onwards all appear to have similar amounts of allele sharing with mostly domestic ancestry but some level of aurochs introgression.

Quantifying the extent of introgression

While f_4 statistics measure allele sharing it does not directly quantify the amount of introgression in the different specimens, hence, we employed three different frameworks to estimate ancestry proportions: f_4 ratio (37), qpAdm (38, 39) and Struct-f4 (40) to model each Iberian individual from European aurochs (Bed3) and/or Anatolian Neolithic cattle (Sub1) as sources (Table 1). While the f_4 ratio provides a straightforward-to-interpret estimate of aurochs ancestry under a simple two source model, we also include qpAdm due to the potential of rejecting models and hinting at additional ancestries. We also include Struct-f4 for better samples ($>0.1X$) as it is more flexible than qpAdm not requiring a strict separation between sources and outgroup populations. While quantitative estimates of European aurochs ancestry for the 20 post-Mesolithic individuals are somewhat correlated between f_4 ratio and qpAdm (Spearman's correlation coefficient $\rho=0.57$, $p=0.01$), they differ for certain individuals. This highlights differences between the methods, their assumptions about the relationships of sources and outgroups, and their sensitivity to low coverage data. For most parts of this study, we decide to present the f_4 ratio results but it is important to highlight that our interpretations are based on the general pattern and not on the ancestry estimates for single individuals.

Sample ID	Site	Period	Date calBP	f ₄ ratio	qpAdm	Struct-f4 (K=5)
moo012a	El Portalón	Early Neolithic	<i>contextual</i>	0.908 [0.743, 1.072]	0.928 [0.77, 1.085]	-
moo004	Els Trocs	Neolithic	7152-6890	0.103 [0.025, 0.182]	0.06 [0.026, 0.098] ¹	0.012 ^a
moo007	Els Trocs	Neolithic	7151-6890	-0.089 [-0.204, 0.026]	0.073 [0.033, 0.113] ¹	0.039 ^a
moo014	El Portalón	Neolithic	6491-6403	0.861 [0.791, 0.931]	0.942 [0.905, 0.979] ²	0.942 ^a
moo039	Mendandia	Neolithic	7426-7280	0.435 [0.334, 0.536]	0.866 [0.76, 0.971]	-
moo015	El Portalón	LNCA	5041-4842	-0.055 [-0.182, 0.072]	0.444 [0.32, 0.565]	-
moo017	El Portalón	LNCA	5567-5326	0.756 [0.646, 0.866]	0.811 [0.769, 0.853] ¹	-
moo019	El Portalón	LNCA	5556-5325	0.723 [0.628, 0.818]	0.828 [0.786, 0.87] ¹	0.654
moo020	El Portalón	LNCA	<i>contextual</i>	0.775 [0.652, 0.897]	0.944 [0.80, 1.085]	-
moo050	El Portalón	LNCA	5468-5320	0.788 [0.709, 0.866]	0.809 [0.773, 0.845] ¹	0.663 ^b
moo011b	El Portalón	Bell Beaker	<i>contextual</i>	0.109 [-0.097, 0.316]	0.667 [0.47, 0.865]	-
moo012b	El Portalón	Bell Beaker	4421-4291	0.706 [0.528, 0.883]	0.326 [0.258, 0.394] ¹	-
moo022	El Portalón	Bell Beaker	<i>contextual</i>	0.169 [0.006, 0.332]	0.729 [0.55, 0.910]	-
moo023	El Portalón	Bell Beaker	4153-3976	0.179 [0.073, 0.285]	0.641 [0.54, 0.738]	0.21 ^a
moo033	El Portalón	Early Bronze Age	<i>contextual</i>	0.241 [0.050, 0.432]	0.987 [0.79, 1.182]	-
moo009a	El Portalón	Bronze Age	3884-3635	0.399 [0.297, 0.502]	0.284 [0.244, 0.324] ¹	0.152 ^a
moo009x	El Portalón	Bronze Age	3829-3513	0.198 [0.115, 0.280]	0.259 [0.203, 0.315] ³	0.176 ^a
moo034	El Portalón	Bronze Age	3811-3492	0.547 [0.300, 0.794]	0.112 [-0.061, 0.285] ⁴	-
moo035	El Portalón	Bronze Age	<i>contextual</i>	0.307 [0.170, 0.445]	0.821 [0.68, 0.963]	-
moo010p	El Portalón	Roman	2334-2156	0.397 [0.205, 0.589]	0.281 [0.217, 0.345] ¹	-

1 To produce a fitting and feasible model ($p > 0.01$) a minor contribution ($\leq 5\%$) of indicine ancestry is required to fit the data.

2 Model does not fit with Bed3 as European aurochs source but fits well ($p = 0.49$) when using CPC98.

3 To produce a fitting and feasible model ($p > 0.01$) a contribution of 33.7% Caucasus aurochs ancestry (Gyu2) is required to fit the data. This additional source is not well resolved as the standard error is large (21.2%).

4 To produce a fitting and feasible model ($p > 0.01$) a contribution of 85.9% Caucasus aurochs ancestry (Gyu2) is required to fit the data. This additional source is not well resolved as the standard error is large (74.6%).

a Struct-f4 also assigns a small proportion of Caucasus aurochs ancestry (Gyu2) to this individual ($< 5\%$).

b Struct-f4 also assigns a proportion of Caucasus aurochs ancestry (Gyu2) to this individual (14.0%).

Table 1

European aurochs ancestry proportions in post-Mesolithic Iberian *Bos* samples.

Square brackets are showing Block-jackknife estimates of the 95% confidence interval. f₄ ratio and qpAdm are using Bed3 as source of European aurochs ancestry unless noted otherwise. Footnotes are added when deviations from the two source model were needed. Struct-f4 was run in semi-supervised mode to estimate ancestry in the Iberian samples with K=5 as the different ancestries separated at this point. Only individuals with at least 0.1x coverage were included in this analysis to ensure convergence. LNCA=Late Neolithic/Chalcolithic

Most of the 20 post-Mesolithic individuals show indications of both domestic and European aurochs ancestries (**Table 1**). Only three individuals (f_4 ratio) or one individual (qpAdm) do not show significant proportions of aurochs ancestry while only one individual (f_4 ratio) or three individuals show not significant proportions of domestic ancestry. Furthermore, qpAdm and Struct-f4 suggest low proportions of additional, eastern ancestries represented either by indicine cattle or the Caucasus aurochs Gyu2 in these analyses. While these ancestries are not well resolved and usually have high standard errors, they suggest that multiple western Asian populations contributed to the European early domestic gene pool. Notably, most Neolithic and pre-Bell Beaker Chalcolithic individuals show either predominantly domestic or aurochs ancestry while many Bell Beaker and Bronze Age individuals show more intermediate values of aurochs ancestry. In fact, from the Bronze Age onwards, most estimates overlap with the approximately 25% aurochs ancestry in modern Iberian cattle (**Fig. 2**; Supplementary Information, Table S1).

A limitation of this analysis is the availability of genomes that can be used as representatives of the source populations. We used German and British aurochs to represent western European aurochs ancestry and a single Anatolian Neolithic to represent the original domestic cattle that was introduced into Europe. Our Mesolithic Iberian aurochs contained too little endogenous DNA to be used as a proxy aurochs reference and all Neolithic and Chalcolithic samples estimated with predominantly aurochs ancestry (including the 2.7x genome of moo014) already carry low (but significant) levels of domestic ancestry. However, the fact that all of these aurochs samples carried P mitochondria strongly suggests that western European aurochs can be considered monophyletic. Furthermore, a recent parallel study also concluded that all Western European aurochs form a clade (28). The Anatolian Sub1 might also not be depleted of any European aurochs ancestry and could not fully represent the original European Neolithic gene pool as also indicated by qpAdm and Struct-f4 identifying small proportions of other Asian ancestries in some Iberian individuals. While these caveats should affect our quantitative estimates of European aurochs ancestry, they should not drive the qualitative pattern as our tests would still detect any excess European aurochs ancestry that was not present in Neolithic Anatolia.

An important question that remains unexamined is the exact process that led to the hybridization since this could provide insight into human management practices or, more generally speaking, mating patterns between wild and domestic individuals. The fact that some individuals with predominantly aurochs ancestry carry T haplogroups (moo019) and that some individuals with predominantly domestic ancestry carry P haplogroups (moo009x) implies that females contributed in both directions. To assess whether the admixture process was sex-biased, we compared aurochs ancestry patterns on the X chromosome and autosomes (**Fig. 3**). Since females carry two X chromosomes and males only have one, we can assume that an excess of a certain ancestry on the X chromosome indicates more females from that particular source population. While the estimates are noisy due to the low coverage data and even less sites available for the X chromosome, it is striking that all but one individual with mostly domestic autosomal ancestry (>50%) show even lower point estimates of aurochs ancestry on the X chromosome. This pattern even extends into the modern Iberian individuals. Male-biased aurochs introgression has been suggested based on mitochondrial haplotypes before (5). In the absence of aurochs Y chromosomal data, however, it is difficult to assess sex-biased processes from uniparental data alone. The comparison of X chromosomes and autosomes should theoretically have more power to detect such processes as they are less sensitive to genetic drift due to their recombining nature (42) but estimation of ancestry proportions on the X chromosome can be affected by different biases (43–45). Overall, our results are consistent with previous observations that the contribution of wild ancestry into domestic cattle was mostly through aurochs bulls.

Figure 2

(A) Estimates of aurochs ancestry (estimated using the f_4 ratio with Bed3 as European aurochs source) in directly dated post-Mesolithic Iberian samples over time. Error bars indicate the 95% confidence interval. Modern Iberian whole-genome sequenced Lidia individuals are added around date 0 with some horizontal jitter. Approximate boundaries for the main sampling periods are indicated by dashed vertical lines.

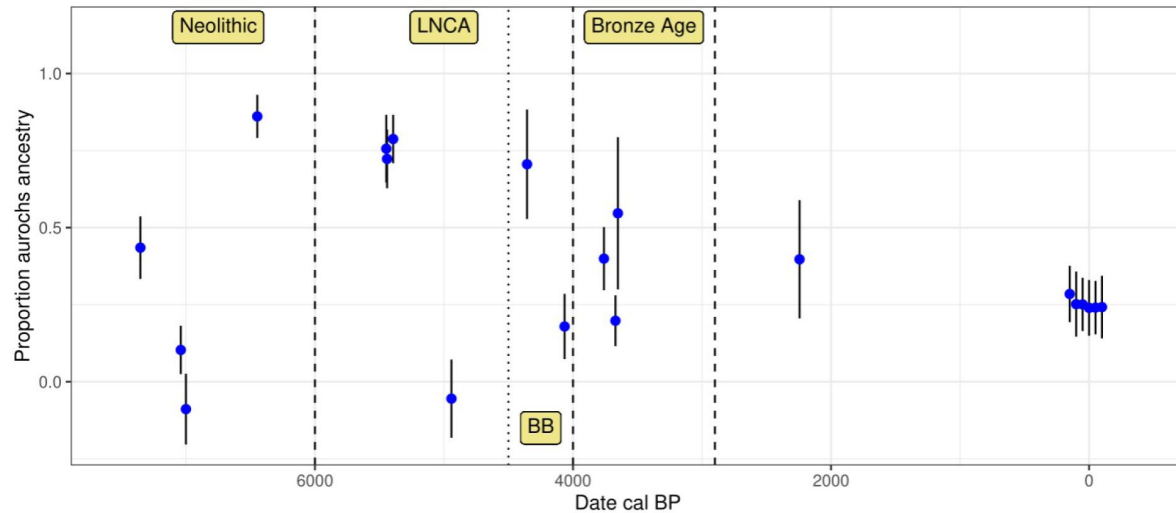
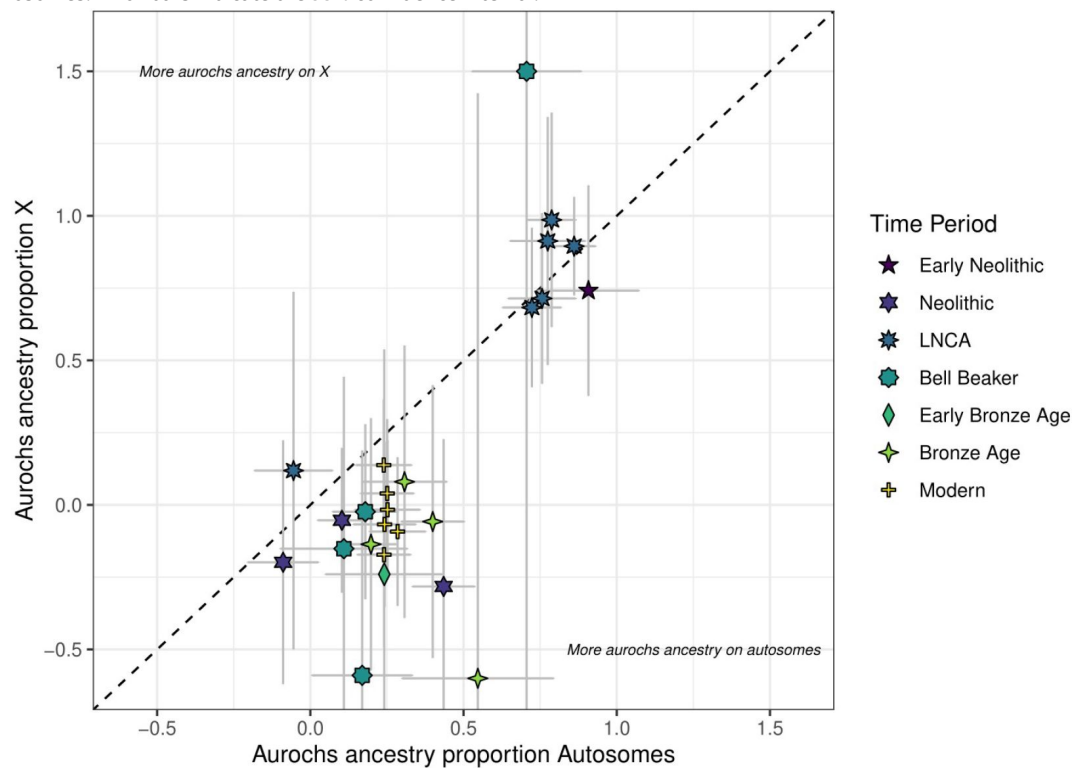


Figure 3

Comparison of f_4 ratio estimated aurochs ancestry of post-Mesolithic Iberian samples on the autosomes versus X chromosomes. Error bars indicate the 95% confidence interval.



Aurochs ancestry in modern breeds and the Spanish Lidia cattle breed

We estimated aurochs ancestry in a set of Western European cattle breeds (27) as we performed for the prehistoric samples. Previous studies have used D statistics for pairwise comparisons between breeds (26, 27, 46). Such D statistics, however, are sensitive to biases including gene flow from populations not included in the analysis (47). Furthermore, qpAdm provides the possibility to reject scenarios not fitting the data. Our point estimates for the aurochs ancestry range between 20% and 30% across all breeds (Fig. 4) and do not show an increase in aurochs ancestry in Iberian breeds (46). This result differs from the previous studies which suggested geographic differences in western and central Europe and we believe this could be due to ancestry from other, non-European groups in some commercial breeds (Supplementary Information). Importantly, not all tested breeds did fit the simple two-source model Anatolian Neolithic domestic + European aurochs, likely representing low levels of contributions from other groups, e.g. indicine cattle (27). The presence of indicine ancestry can be confirmed in a qpAdm analysis using three sources resulting in fitting models for all breeds (Supplementary Information, Table S4).

Cattle have played an important role in Iberian culture during the last centuries as they have been part of numerous traditional popular events including bullfighting. The Lidia breed, a heterogeneous group of Iberian cattle that is mainly bred for aggressive behavior, has commonly been used for such popular festivities (33). Even though Lidia cattle has only been actively bred for agonistic behavior for about 200 years, some people attribute their aggressiveness and appearance as an indication of high levels of aurochs ancestry (34). We additionally use medium coverage genomes of six Lidia individuals (41) to estimate their proportion of aurochs ancestry. Lidia cattle in the (27) data set had a point estimate of 25% (95% confidence interval: [18.5, 31.5]) aurochs ancestry and estimates in the individual genomes ranged from 17.6% [10.9, 24.3] to 23.5% [17.0, 30.0] (Supplementary Information, Figure S8) – all overlapping with the observed range for other western European breeds. Despite some variation between individuals, which might be attributable to noise due to low coverage sequencing data in the reference populations, we do not observe a systematically elevated level of aurochs ancestry compared to other modern breeds or ancient samples since the Bronze Age. While these results reject the idea that the specifics of Lidia cattle can be attributed to a substantially increased genome-wide aurochs ancestry, it does not rule out the possibility that the roots of their aggressiveness and appearance are indeed due to aurochs variants at key loci responsible for those traits. An in-depth investigation of such questions would require a larger dataset of aurochs genomes as well as a more comprehensive Lidia sampling due to their fragmentation in highly distant genetic lineages (33).

Stable Isotope Analysis

In addition to their ancestry, we studied the ecology of the bovids through stable isotope analysis of bone collagen. Lynch et al. (48) suggested that stable isotope data could be used to infer niche separation between the species in Britain, with domestic cattle in more open settings, while aurochs (about 1‰ more depleted in $\delta^{13}\text{C}$) were habitually in more forested areas, or wet ground. This is most likely facilitated through human management of the domestic cattle, separating them from their wild counterparts. In contrast, Noe-Nygaard et al. (49) failed to observe such an effect in samples from Denmark and northern Germany.

Considering our dataset and other data published on Iberian cattle (categorised on morphology/date) (50–59) we observe that the nitrogen isotope means are statistically different only when our data are compared using morphological characteristics, not genetic distinctions (see Dataset S1 and Supplementary Information). This difference is mostly due to

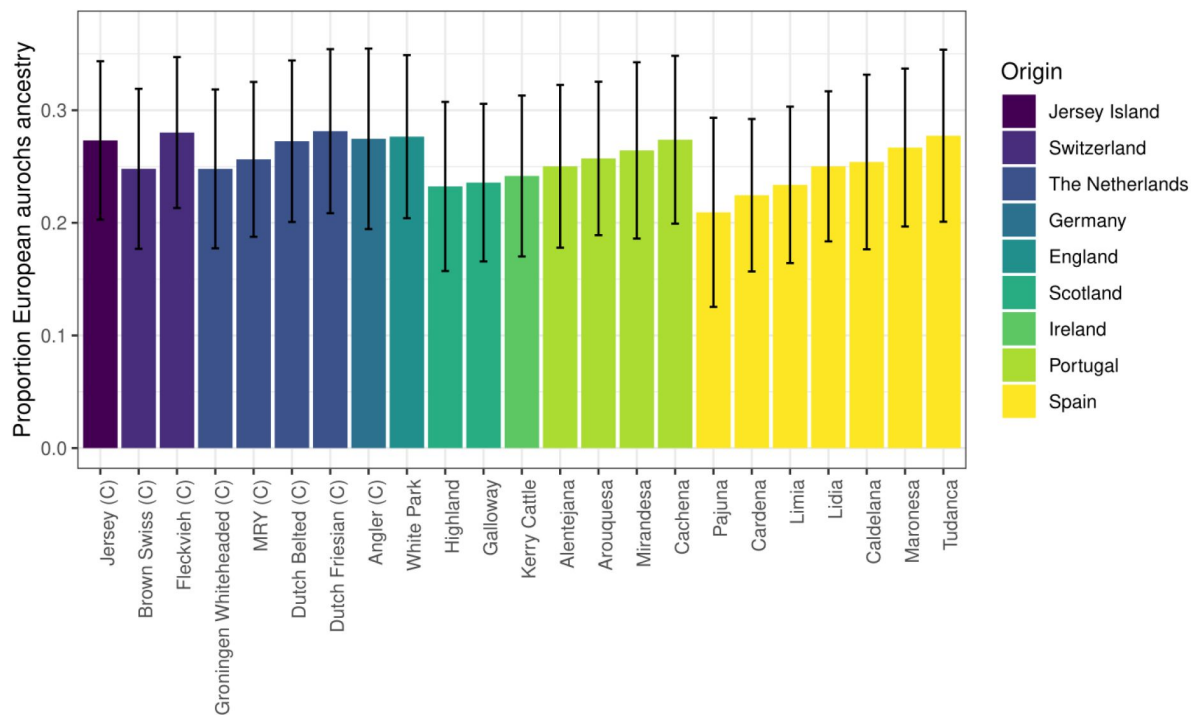


Figure 4

qpAdm estimates of Aurochs ancestry in modern western European cattle breeds from the (27) dataset. Commercial breeds are marked with a "C". The figure is only showing breeds with feasible and non-rejected two source models, all results are shown in Supplementary Information, Table S1. Error bars are showing block-jackknife estimates of the 95% confidence interval.

some domestic cattle with $\delta^{15}\text{N}$ values greater than 6.5‰. This could be explained by some taurine cattle having exclusive habitual access to high nitrogen isotope ratio resources. For example, human management such as corralling on manured ground, or feeding with manured crops, would produce this effect. Nevertheless, there is generally a large amount of overlap in the isotope values for the two groups suggesting that wild and domesticated groups often did not occupy different niches in Iberia.

Discussion

We generated and analyzed biomolecular data from *B. primigenius* and *B. taurus* spanning more than 9000 years in the same region. Cattle are important livestock in the Iberian Peninsula today, and our results illustrate the interaction between domestic cattle and their wild relatives in the past. The two groups show signs of frequent hybridization starting soon after the arrival of cattle to the peninsula, as evident in our oldest directly dated Neolithic individual (moo039, 7426-7280 CalBP) where signals of carrying both ancestries are clear. Throughout the Neolithic, we observed large variations in the wild versus domestic ancestry per individual, but this pattern later stabilized (to 20-30% aurochs ancestry) from the Chalcolithic/Bronze Age onwards. As we do not know whether the sequenced individuals were hunted or herded, this could reflect a transition from hunting and herding to predominantly herding and it is possible that systematic herd management led to the nearly constant levels of aurochs ancestry over the last 4000 years. This period also coincides with several other societal changes; including the Bell Beaker complex and the introduction of human ancestry from the Pontic steppe into the Iberian Peninsula (60–62). Around this time, humans also started processing a significantly higher amount of dairy products connected with the “secondary product revolution” (63, 64). Aurochs were probably present in Iberia until Roman times (31) leaving possibilities for interbreeding but we cannot exclude that various factors such as hunting or changing vegetation had led to a substantial decline in the wild aurochs population around the early Bronze Age. A previous study on cattle morphology from the site of El Portalón described a decrease in size from the Neolithic to the Chalcolithic and a further significant size decrease from the Chalcolithic to the Bronze Age (65) and associated this change in size to the aridification of the area at this time (66). Indeed, this climatic change could also be related to a reduction of the aurochs population contributing to the stabilization of the levels of ancestry in domestic cattle from the Bronze Age to the present. Nonetheless, our stable isotope results suggest that wild and domesticated groups often did not occupy substantially different niches on the Iberian Peninsula. Material excavated from Denmark suggested that aurochs changed their niches over time (49) demonstrating some flexibility depending on local vegetation and the possibility of aurochs adapting to changing environments.

The reduced level of aurochs ancestry on the X chromosome (compared to the autosomes) in admixed individuals suggests that it was mostly aurochs males who contributed wild ancestry to domestic herds, a process that had been suggested based on the distribution of mitochondrial haplotypes before (5). A recent parallel study based using ancient genomes also detected male-biased aurochs introgression using similar methods as our study (28). Consequently, the offspring of wild bulls and domestic cows could be born into and integrated within managed herds. It is unclear how much of this process was intentional but the possibility of a wild bull inseminating a domestic cow without becoming part of the herd suggests that some level of incidental interbreeding was possible. For Neolithic Turkey, it has been suggested that allowing insemination of domesticated females by wild bulls was intentional, maybe even ritual (67). Modern breeders are still mostly exchanging bulls or sperm to improve their stock which manifests in a lower between-breed differentiation on the X chromosome (46).

The lack of correlation between genomic, stable isotope and morphological data highlights the difficulties of identifying and defining aurochs to the exclusion of domestic cattle. All of these data measure different aspects of an individual: their ancestry, ecology or appearance, respectively.

While they can give some indication, none of them are a direct measurement of how these cattle were recognised by prehistoric humans or whether they were herded or hunted. It remains unclear whether our ancestry inferences had any correlation to how prehistoric herds were managed and how much intentional breeding is behind the observed pattern of hybridization. It is even possible that all hybrids identified in this study were part of domestic herds.

Even though wild aurochs populations went extinct, European aurochs ancestry survived into modern cattle with a relatively uniform distribution across western European breeds. Isolated Iberian Lidia, bred for their aggressiveness, appears to be no exception to this pattern. This rejects the notion that an overall increased proportion of aurochs ancestry causes the distinctiveness of certain breeds, but considering the functional relevance of archaic introgression into modern humans (68), it is possible that aurochs variants at functional loci may have a substantial influence on the characteristics of modern cattle breeds. Our low coverage sequencing data did not allow us to investigate this but future bioarchaeological studies combining different types of data will have the possibility to clarify the role of the extinct aurochs ancestry in modern domestic cattle.

Conclusions

Using a bioarchaeological approach we have demonstrated that since cattle arrived in Iberia there has been hybridization with the local aurochs population, and that mainly aurochs bulls contributed to the gene pool still found in domestic herds today. Admixture proportions vary for the first few millennia but stabilize during the Bronze Age at approximately 20-30% of wild ancestry in the individuals found at the Iberian archaeological sites, a level that is still observed in modern Iberian breeds, including the more aggressive Lidia breed. This development could be the result of an initial mix of hunting and herding together with a generally loose management of herds, becoming more controlled over time in combination with a reduced importance of hunting wild aurochs.

The amount of hybridisation observed in the ancient cattle makes it difficult to genetically define what a domestic or wild *Bos* is, bringing into doubt the validity of such categorisations. Our interpretation is made more difficult by the overlap in morphological and metric data, creating further difficulties in species determination (especially in hybrids) and niche sharing as revealed by stable isotopes. To some extent, our interpretation is moot, as the salient matter is, how did prehistoric humans interact with cattle? What was their sense of wild and domestic and hybridisation? While we have recognised individual hybrids, to what extent these were part of domestic herds or intentionally bred and managed is uncertain.

Another source of uncertainty in our determinations is the limited knowledge about the genetic diversity in European aurochs. Further regional (and temporally longitudinal) aurochs genomes would aid future genomic studies defining the genetic variation in the European aurochs population.

Materials and methods

Data generation

We attempted DNA extractions of 50 archaeological remains from which we successfully extracted DNA from 24 individuals identified as domestic cattle and aurochs excavated from four prehistoric sites in Iberia: El Portalón de Cueva Mayor (n=18), Artusia (n=1), Els Trocs (n=2) and Mendandia (n=3). Teeth and bones were UV irradiated (6 J/cm² at 254 nm) and the first millimeter of bone/tooth surface abraded using a Dremel™ tool. DNA was extracted in a dedicated ancient DNA facility using a silica-based DNA extraction protocol (69). For each sample, 100-200mg of bone or tooth powder were incubated for 24 h at 37°C, using the MinElute column Zymo extender

assembly replaced by the High Pure Extender Assembly (Roche High Pure Viral Nucleic Acid Large Vol. Kit) and performed twice for each sample. DNA extracts were subjected to UDG treatment for the removal of deaminated cytosines and were further converted into blunt-end double stranded Illumina multiplex sequencing libraries (70). Between seven and fifteen qPCR cycles were performed to amplify the DNA libraries using indexed primers (70). These were subsequently pooled at equimolar concentrations and shotgun sequenced on Illumina HiSeq and Novaseq sequencing platforms.

Radiocarbon dates

Eight Bone and teeth were directly radiocarbon dated (AMS) at Waikato University in New Zealand and two teeth at Beta Analytics in the United States. Radiocarbon dates were calibrated using the OXcal 4.4 program (71) and the IntCal20 calibration curve (72). Three samples from the site of Mendandia were conventionally radiocarbon dated at Groningen (Netherlands) radiocarbon laboratory and calibrated as above.

Stable Isotopes Analysis

Many of the samples analysed here were radiocarbon dated and stable isotope data (via IRMS) were generated in this process, to augment this data we also produced stable isotope data for some additional samples in this dataset, where they were available. The additional samples underwent bone collagen or tooth dentine collagen extraction at the Laboratorio de Evolución Humana (Universidad de Burgos) following the protocol of (73). In brief, this is a cold acid demineralization, followed by Milli Q water rinsing, gelatinization at pH3 (24 hrs at 70°C), Eze filtering and lyophilization. Collagen yields (as % mass of starting bone) were recorded. Stable isotope values ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$) and %C, %N were measured in duplicate at the Universitat Autònoma de Barcelona, unless only one sample was successful in the analysis. Collagen samples (approx. 0.4 mg) were analysed using a Flash IRMS elemental analyser (EA) coupled to a Delta V Advantage isotope ratio mass spectrometer (IRMS), both from Thermo Scientific (Bremen, Germany) at the Institute of Environmental Science and Technology of the Universitat Autònoma de Barcelona (ICTA-UAB). International laboratory standard IAEA-600 was used, with measurements made relative to Vienna PeeDee Belemnite (V-PDB) for $\delta^{13}\text{C}$, and air N_2 (AIR) for $\delta^{15}\text{N}$. The average analytical error was $<0.2\%$ (1σ) as determined from the duplicate analyses of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$. In house standards used was dog hair collected and homogenized for interlaboratory comparisons.

Data processing

HiSeq X10 reads have been trimmed and merged using AdapterRemoval (74) while adapters for NovaSeq 6000 reads have been trimmed with cutadapt (75) and merging was performed with FLASH (76) requiring a minimum overlap of 11bp. Single-end reads of at least 35bp length were then mapped to the cattle reference genomes UMD3.1 (77) and Btau5 (78) using bwa (79) with the non-default parameters: -l 16500, -n 0.01, and -o 2. Different sequencing runs per sample were merged with samtools (80) and consensus sequences were called for duplicate sequences with identical start and end coordinates (81). Finally, reads with more than 10% mismatches to the reference genome were removed. Biological sex was assigned to the samples mapped to the Btau_5 reference genome (as UMD3.1 does not contain a Y chromosome assembly) using the R_x method (82) modified for 29 autosomes.

Mitochondrial contamination was estimated following the approach used by Green et al. (83) for hominins. We first identified nearly private mutations (less than 5% frequency in the 278 diverse mitogenomes used by MitoToolPy and dometree (84), obtained from Dryad <https://doi.org/10.5061/dryad.cc5kn>) in each individual and then used the proportion of non-consensus alleles at these sites to estimate contamination. We restricted this analysis to sites with at least 10x coverage, a minimum base quality of 30. Furthermore, transition sites with a C or G in the

consensus mitogenome were excluded to avoid over-estimation due to post-mortem damage. Standard errors were estimated assuming a binomial distribution around the point estimate. Code used for this step can be found at https://github.com/GuntherLab/mt_contam_domestic_green

For comparative purposes, we also processed published data from (5, 18, 19, 26) using the same bioinformatic pipeline. Furthermore, we downloaded sequence data for six Spanish Lidia cattle (41), a single modern water buffalo (*Bubalus bubalis*, Jaffrabadi-0845) (85) and a single zebu cattle individual (Sha_3b) (86) and processed them with our ancient DNA mapping pipeline. To obtain a pseudohaploid Yak (*Bos grunniens*) sequence, we followed the approach by (27) splitting the Yak reference genome (87) contigs into 100bp fragments and mapping them to the UMD3.1 reference genome.

Data Analysis

Mitochondrial consensus sequences were called using ANGSD (88) and the options -doFasta 2 -doCounts 1 -minQ 30 -minMapQ 30. Mitochondrial haplogroups were then assigned to the whole mitogenome sequences using the Python script MitoToolPy (84).

For population genomic analysis, we used a panel of SNPs derived from Run6 of the 1000 genomes project (89, 90). We obtained a list of SNPs from (91) and reduced the panel to biallelic SNPs of at least 10% minor allele frequency in the joint European *B. taurus*/Asian *B. indicus* dataset. Prior to genotype calling, all ancient BAM files were modified such that Ts in the first 5 bases of each fragment and As at the last 5 base pairs of each fragment have a base quality of 2. This approach allows to include more sites than excluding all transitions which are potentially affected by post-mortem damage. It produces highly correlated f_4 statistics (Supplementary Information, Figure S2) and f_4 ratios (Supplementary Information, Figure S3) but much lower standard errors in f_4 ratios due to the larger total number of sites (Supplementary Information, Figure S3). To generate pseudohaploid representations of each individual, we randomly draw a single read with mapping and base quality of at least 30 at each SNP position. If the allele carried by the ancient individual was not one of the two known alleles, we removed the site from the panel. Using this approach, ~9.1 million autosomal and 248K X chromosomal SNPs were genotyped in the ancient samples. To compare the ancient samples to a diverse set of modern cattle, we used the panel of modern European breeds presented by (27) which were genotyped at ~770,000 SNPs. The ancient samples were genotyped the same way as for the 1000 Bulls project SNP panel.

To conduct an ordination of the nuclear data, sequences of 43 ancient Eurasian cattle and two aurochs were obtained from (26) and (5). Outgroup f_3 statistics were calculated for all pairs of our Iberian *Bos* samples, using a Yak (*Bos grunniens*) genome as an outgroup, and a distance matrix for all samples was calculated as $1-f_3$. All f -statistics were calculated in R version 4.1.2 (92) package ‘admixtools2’ (93). The distance matrix was used to compute scores for non-metric multi-dimensional scaling (NMDS) ordinations using the metaMDS function in the ‘vegan’ R package and 10000 random starts (94).

European aurochs introgression α into Iberian individual X was estimated using f_4 ratios calculated with POPSTATS (95) and the equation

$$\alpha = 1 - \frac{f_4(\text{Gyu 2, YAK}; X, \text{EuropeanAurochs})}{f_4(\text{Gyu 2, YAK}; i_1, \text{EuropeanAurochs})}$$

Both Bed3 and CPC98 were separately tested as aurochs source and Bed3 was chosen for the results presented in the article due to lower confidence intervals. POPSTATS was run with the non-default options -ratio, -testpop and -not23 to allow for more autosomes than humans have. We also used admixtools2 (93) and qpAdm (38, 39) to model the ancestry proportions in the samples. Bed3 was used as a source for European aurochs ancestry (due to lower standard errors

in the f_4 ratios) while the domestic Anatolian Neolithic Sub1 was used as a source for domesticated cattle ancestry. As “right” populations, we used Gyu2, *B. indicus*, Yak and *Bison bonasus bonasus* PLANTA. qpAdm was run with auto_only=FALSE, maxmiss=0.5 and allsnps=TRUE. When the two source model did not fit ($p < 0.01$) or produced infeasible admixture proportions outside [0, 1], we used rotate_models and qpAdm_multi to find alternative models adding CPC98 as an additional possible source or “right” population. qpAdm was also used for the modern western European breed panel from (27) adding Bes2 (5) to the “right” populations and excluding breeds from Italy and the Balkan from the targets as non-taurine ancestry (27) in them would lead to a rejection of the models. Finally, we also used Struct-f4 (40) to estimate ancestry proportions. First, input files were generated with the provided helper scripts and f_4 statistics were calculated in blocks of 5Mbp. Struct-f4 was then run in semi-supervised mode to estimate ancestries in Iberian individuals with at least 0.1x coverage. This cutoff was chosen as lower coverage samples prevented conversion. CPC98, YAK, Ch22, Gyu2, Bed3, Sub1 and Sha_3b were used as additional individuals to provide a framework of different possible ancestries.

Data availability

Raw sequence data and aligned reads for the new ancient individuals are available through the European Nucleotide Archive under accession number PRJEB63140. All metric and isotope data are available in Dataset S1.

Acknowledgements

We are incredibly grateful to the excavation teams at the four archaeological sites. Sequencing was performed at The National Genomics Infrastructure (NGI) Stockholm. The computations and data handling were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) at Uppmax, partially funded by the Swedish Research Council through grant agreements no. 2022-06725 and no. 2018-05973. This project was supported by grants from the Royal Physiographic Society of Lund (Nilsson-Ehle Endowments) to T.G. and C.V., Vetenskapsrådet (2017-05267) to T.G., Ramón y Cajal (RYC2018-025223-I) to C.V and a Beatriz Galindo Fellowship (BGS220-461AA-69201) to C.S.

References

1. Frantz L. A. F. *et al.* (2019) **Ancient pigs reveal a near-complete genomic turnover following their introduction to Europe** *PNAS* **116**:17231–17238
2. Frantz L. A. F., Bradley D. G., Larson G., Orlando L. (2020) **Animal domestication in the era of ancient genomics** *Nat Rev Genet* :1–12
3. Botigué L. R. *et al.* (2017) **Ancient European dog genomes reveal continuity since the Early Neolithic** *Nat Commun* **8**
4. Irving-Pease E., Ryan H., Jamieson A., Dimopoulos E. A., Larson G., Frantz L., Lindqvist C., Rajora O. P. (2018) **Paleogenomics of Animal Domestication** *Paleogenomics* Springer
5. Verdugo M. P. *et al.* (2019) **Ancient cattle genomics, origins, and rapid turnover in the Fertile Crescent** *Science* **365**:173–176
6. Daly K. G. *et al.* (2018) **Ancient goat genomes reveal mosaic domestication in the Fertile Crescent** *Science* **361**:85–88
7. Daly K. G. *et al.* (2021) **Herded and hunted goat genomes from the dawn of domestication in the Zagros Mountains** *Proceedings of the National Academy of Sciences* **118**
8. Fages A. *et al.* (2019) **Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series** *Cell* **177**:1419–1435
9. Librado P. *et al.* (2021) **The origins and spread of domestic horses from the Western Eurasian steppes** *Nature* **598**:634–640
10. Yurtman E. *et al.* (2021) **Archaeogenetic analysis of Neolithic sheep from Anatolia suggests a complex demographic history since domestication** *Commun Biol* **4**:1–11
11. Bergström A. *et al.* (2020) **Origins and genetic legacy of prehistoric dogs** *Science* **370**:557–564
12. Ginja C. *et al.* (2023) **Iron age genomic data from Althiburos – Tunisia renew the debate on the origins of African taurine cattle** *iScience* **26**
13. Larsson M. N. A. *et al.* (2024) **Ancient Sheep Genomes Reveal Four Millennia of North European Short-Tailed Sheep in the Baltic Sea Region** *Genome Biology and Evolution* **16**
14. Kaptan D. *et al.* (2024) **The Population History of Domestic Sheep Revealed by Paleogenomes** *Molecular Biology and Evolution* **41**
15. Peters J., Helmer D., Von den Driesch A., Segui M. Sana (1999) **Early Animal Husbandry in the Northern Levant** *Paléorient* **25**:27–48
16. Helmer D., Gourichon L., Monchot H., Peters J., Segui M., Vigne J.-D., Peters J., Helmer D. (2005) **Identifying early domestic cattle from Pre-Pottery Neolithic sites on the Middle Euphrates using sexual dimorphism** *The First Steps of Animal Domestication: New Archaeozoological Approaches* Oxbow Books :86–96

17. Zhang K., Lenstra J. A., Zhang S., Liu W., Liu J. (2020) **Evolution and domestication of the Bovini species** *Animal Genetics* **51**:637–657
18. Palacio P., Berthoud V., Guérin C., Lambourdière J., Maksud F., Philippe M., Plaire D., Stafford T., Marsolier-Kergoat M.-C., Elalouf J.-M. (2017) **Genome data on the extinct Bison schoetensacki establish it as a sister species of the extant European bison (Bison bonasus)** *BMC Evol Biol* **17**
19. Węcek K. *et al.* (2017) **Complex Admixture Preceded and Followed the Extinction of Wisent in the Wild** *Molecular Biology and Evolution* **34**:598–612
20. Ward J. A., McHugo G. P., Dover M. J., Hall T. J., Ng'ang'a S. I., Sonstegard T. S., Bradley D. G., Frantz L. A. F., Salter-Townshend M., MacHugh D. E. (2022) **Genome-wide local ancestry and evidence for mitonuclear coadaptation in African hybrid cattle populations** *iScience* **25**
21. Beja-Pereira A. *et al.* (2006) **The origin of European cattle: Evidence from modern and ancient DNA** *Proceedings of the National Academy of Sciences* **103**:8113–8118
22. Achilli A. *et al.* (2008) **Mitochondrial genomes of extinct aurochs survive in domestic cattle** *Current Biology* **18**:R157–R158
23. Schibler J., Elsner J., Schlumbaum A. (2014) **Incorporation of aurochs into a cattle herd in Neolithic Europe: single event or breeding?** *Sci Rep* **4**
24. Cubric-Curik V. *et al.* (2022) **Large-scale mitogenome sequencing reveals consecutive expansions of domestic taurine cattle and supports sporadic aurochs introgression** *Evolutionary Applications* **15**:663–678
25. Bro-Jørgensen M. H., Carøe C., Vieira F. G., Nestor S., Hallström A., Gregersen K. M., Etting V., Gilbert M. T. P., Sinding M.-H. S. (2018) **Ancient DNA analysis of Scandinavian medieval drinking horns and the horn of the last aurochs bull** *Journal of Archaeological Science* **99**:47–54
26. Park S. D. *et al.* (2015) **Genome sequencing of the extinct Eurasian wild aurochs, Bos primigenius, illuminates the phylogeography and evolution of cattle** *Genome biology* **16**
27. Upadhyay M. *et al.* (2016) **Genetic origin, admixture and population history of aurochs (Bos primigenius) and primitive European cattle** *Heredity*
28. Rossi C. *et al.* (2024) **The genomic natural history of the aurochs** *Nature* :1–6
29. Anderung C., Bouwman A., Persson P., Carretero J. M., Ortega A. I., Elburg R., Smith C., Arsuaga J. L., Ellegren H., Götherström A. (2005) **Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle** *Proceedings of the National Academy of Sciences of the United States of America* **102**:8431–8435
30. Erven J. A. M., Scheu A., Verdugo M. P., Cassidy L., Chen N., Gehlen B., Street M., Madsen O., Mullin V. E. (2024) **A High-Coverage Mesolithic Aurochs Genome and Effective Leveraging of Ancient Cattle Genomes Using Whole Genome Imputation** *Molecular Biology and Evolution* **41**
31. Altuna J., Mariezkurrena K. (2002) **El problema de la domesticación de bovinos en el País Vasco y resto de la Región Cantábrica** *Congresos de Estudios Vascos* **15**:123–128

32. Cubas M., Altuna J., Álvarez-Fernández E., Armendariz A., Fano M. Á., López-Dóriga I. L., Mariezkurrena K., Tapia J., Teira L. C., Arias P. (2016) **Re-evaluating the Neolithic: The Impact and the Consolidation of Farming Practices in the Cantabrian Region (Northern Spain)** / *World Prehist* **29**:79–116
33. Caíón J., Tupac-Yupanqui I., García-Atance M. A., Cortés O., García D., Fernández J., Dunner S. (2008) **Genetic variation within the Lidia bovine breed** *Animal Genetics* **39**:439–445
34. Balcarcel A. M., Veitschegger K., Clauss M., Sánchez-Villagra M. R. (2021) **Intensive human contact correlates with smaller brains: differential brain size reduction in cattle types** *Proceedings of the Royal Society B: Biological Sciences* **288**
35. Sinding M.-H. S., Gilbert M. T. P. (2016) **The Draft Genome of Extinct European Aurochs and its Implications for De-Extinction** *Open Quaternary* **2**
36. Colominas L. *et al.* (2015) **Detecting the T1 cattle haplogroup in the Iberian Peninsula from Neolithic to medieval times: new clues to continuous cattle migration through time** *Journal of Archaeological Science* **59**:110–117
37. Patterson N., Moorjani P., Luo Y., Mallick S., Rohland N., Zhan Y., Genschoreck T., Webster T., Reich D. (2012) **Ancient admixture in human history** *Genetics* **192**:1065–1093
38. Haak W. *et al.* (2015) **Massive migration from the steppe was a source for Indo-European languages in Europe** *Nature* <https://doi.org/10.1038/nature14317>
39. Harney E., Patterson N., Reich D., Wakeley J. (2021) **Assessing the performance of qpAdm: a statistical tool for studying population admixture** *Genetics* **217**
40. Librado P., Orlando L. (2022) **Struct-f4: a Rcpp package for ancestry profile and population structure inference from f4-statistics** *Bioinformatics* **38**:2070–2071
41. Eusebi P. G., Cortés Ó., Contreras E., Caíón J., Dunner S., Sevane N. (2022) **A novel missense variant in endothelin-2 (EDN2) causes a growth and respiratory lethal syndrome in bovine** *Animal Genetics* **53**:583–591
42. Keinan A., Mullikin J. C., Patterson N., Reich D. (2009) **Accelerated genetic drift on chromosome X during the human dispersal out of Africa** *Nat Genet* **41**:66–70
43. Lazaridis I., Reich D. (2017) **Failure to replicate a genetic signal for sex bias in the steppe migration into central Europe** *Proceedings of the National Academy of Sciences* **114**:E3873–E3874
44. Goldberg A., Günther T., Rosenberg N. A., Jakobsson M. (2017) **Reply to Lazaridis and Reich: Robust model-based inference of male-biased admixture during Bronze Age migration from the Pontic-Caspian Steppe** *Proceedings of the National Academy of Sciences* **114**:E3875–E3877
45. Pfennig A., Lachance J. (2023) **Challenges of accurately estimating sex-biased admixture from X chromosomal and autosomal ancestry proportions** *The American Journal of Human Genetics* **110**:359–367
46. da Fonseca R. R., Ureía I., Afonso S., Pires A. E., Jørsboe E., Chikhi L., Ginja C. (2019) **Consequences of breed formation on patterns of genomic diversity and differentiation: the case of highly diverse peripheral Iberian cattle** *BMC Genomics* **20**

47. Rogers A. R., Bohlender R. J. (2015) **Bias in estimators of archaic admixture** *Theoretical population biology* **100**:63–78
48. Lynch A. H., Hamilton J., Hedges R. E. M. (2008) **Where the wild things are: aurochs and cattle in England** *Antiquity* **82**:1025–1039
49. Noe-Nygaard N., Price T. D., Hede S. U. (2005) **Diet of aurochs and early cattle in southern Scandinavia: evidence from 15N and 13C stable isotopes** *Journal of Archaeological Science* **32**:855–871
50. Fernández-Crespo T., Schulting R. J. (2017) **Living different lives: Early social differentiation identified through linking mortuary and isotopic variability in Late Neolithic/Early Chalcolithic north-central Spain** *PLOS One* **12**
51. Fontanals-Coll M., Subirà M. E., Bonilla M. D.-Z., Duboscq S., Gibaja J. F. (2015) **Investigating palaeodietary and social differences between two differentiated sectors of a Neolithic community, La Bòbila Madurell-Can Gambús (north-east Iberian Peninsula)** *Journal of Archaeological Science: Reports* **3**:160–170
52. GarciaGuix E., Richards MichaelP., Subir M. E. (2006) **Palaeodiets of Humans and Fauna at the Spanish Mesolithic Site of El Collado** *Current Anthropology* **47**:549–556
53. Jones J. R., Maeso C. V., Ballester E. C., Martín L. V., Arceo M. E. D., Marín-Arroyo A. B. (2019) **Investigating prehistoric diet and lifeways of early farmers in central northern Spain (3000–1500 CAL BC) using stable isotope techniques** *Archaeol Anthropol Sci* **11**:3979–3994
54. Navarrete V., Colonese A. C., Tornero C., Antolín F., Von Tersch M., Subirà M. Eulàlia, Comes P., Rosell-Melé A., Saia M. (2017) **Feeding Management Strategies among the Early Neolithic Pigs in the NE of the Iberian Peninsula** *International Journal of Osteoarchaeology* **27**:839–852
55. Salazar-García D. C. (2009) **Estudio de la dieta en la población neolítica de Costamar: Resultados preliminares de análisis de isótopos estables de Carbono y Nitrógeno** *Torre la Sal (Ribera de Cabanes, Castellón): evolución del paisaje antrópico desde la Prehistoria hasta el Medioevo, 2009 Servei d'Investigacions Arqueològiques i Prehistòriques* :411–418
56. Salazar-García D. (2011) **Aproximación a la dieta de la población calcolítica de La Vital a través del análisis de isótopos estables del carbono y del nitrógeno sobre restos óseos** *La Vital (Gandia, Valencia). Vida y Muerte En La Desembocadura Del Serpis Durante El III y El I Milenio a.C* :139–143
57. Salazar-García D. C., Aura J. E., Olària C. R., Talamo S., Morales J. V., Richards M. P. (2014) **Isotope evidence for the use of marine resources in the Eastern Iberian Mesolithic** *Journal of Archaeological Science* **42**:231–240
58. Sarasketa-Gartzia I., Villalba-Mouco V., le Roux P., Arrizabalaga Á., Salazar-García D. C. (2018) **Late Neolithic-Chalcolithic socio-economical dynamics in Northern Iberia. A multi-isotope study on diet and provenance from Santimamié and Pico Ramos archaeological sites (Basque Country, Spain)** *Quaternary International* **481**:14–27
59. Villalba-Mouco V., Utrilla P., Laborda R., Lorenzo J. I., Martínez-Labarga C., Salazar-García D. C. (2018) **Reconstruction of human subsistence and husbandry strategies from the Iberian Early Neolithic: A stable isotope approach** *American Journal of Physical Anthropology* **167**:257–271

60. Olalde I. *et al.* (2018) **The Beaker phenomenon and the genomic transformation of northwest Europe** *Nature* **555**:190–196
61. Valdiosera C. *et al.* (2018) **Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia** *PNAS*
62. Olalde I. *et al.* (2019) **The genomic history of the Iberian Peninsula over the past 8000 years** *Science* **363**:1230–1234
63. Francés-Negro M., Iriarte E., Galindo-Pellicena M. A., Gerbault P., Carrancho A., Pérez-Romero A., Arsuaga J. L., Carretero J. M., Roffet-Salque M. (2021) **Neolithic to Bronze Age economy and animal management revealed using analyses lipid residues of pottery vessels and faunal remains at El Portalón de Cueva Mayor (Sierra de Atapuerca, Spain)** *Journal of Archaeological Science* **131**
64. Evershed R. P. *et al.* (2022) **Dairying, diseases and the evolution of lactase persistence in Europe** *Nature* **608**:336–345
65. Galindo-Pellicena M. A., Arsuaga J. L., Pérez-Romero A., Iriarte E., De Gaspar I., Carretero J. M. (2020) **Metrical analysis of bovine bone remains from the Neolithic to the Bronze Age at the El Portalón site (Atapuerca, Burgos) in the Iberian context** *Quaternary International* **566**:211–223
66. Martínez-Pillado V. *et al.* (2014) **Upper Pleistocene and Holocene palaeoenvironmental records in Cueva Mayor karst (Atapuerca, Spain) from different proxies: speleothem crystal fabrics, palynology and archaeology** *International Journal of Speleology* **43**
67. Peters J., Arbuckle B. S., Pöllath N. (2012) **Subsistence and beyond: animals in Neolithic Anatolia** *The Neolithic in Turkey* **6**:1–65
68. Reilly P. F., Tjahjadi A., Miller S. L., Akey J. M., Tucci S. (2022) **The contribution of Neanderthal introgression to modern human traits** *Current Biology* **32**:R970–R983
69. Dabney J. *et al.* (2013) **Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments** *Proc. Natl. Acad. Sci. U.S.A* **110**:15758–15763
70. Meyer M., Kircher M. (2010) **Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing** *Cold Spring Harbor Protocols* **2010**
71. Ramsey C. B., Others (2009) **Bayesian analysis of radiocarbon dates** *Radiocarbon* **51**:337–360
72. Reimer P. J., Austin W. E., Bard E., Bayliss A., Blackwell P. G., Ramsey C. B., Butzin M., Cheng H., Edwards R. L., Friedrich M. (2020) **The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP)** *Radiocarbon* **62**:725–757
73. Richards M. P., Hedges R. E. M. (1999) **Stable Isotope Evidence for Similarities in the Types of Marine Foods Used by Late Mesolithic Humans at Sites Along the Atlantic Coast of Europe** *Journal of Archaeological Science* **26**:717–722
74. Schubert M., Lindgreen S., Orlando L. (2016) **AdapterRemoval v2: rapid adapter trimming, identification, and read merging** *BMC Research Notes* **9**

75. Martin M. (2011) **Cutadapt removes adapter sequences from high-throughput sequencing reads** *EMBnet.journal* **17**:10–12
76. Magoč T., Salzberg S. L. (2011) **FLASH: fast length adjustment of short reads to improve genome assemblies** *Bioinformatics* **27**:2957–2963
77. Zimin A. V. *et al.* (2009) **A whole-genome assembly of the domestic cow, *Bos taurus*** *Genome Biology* **10**
78. Baylor College of Medicine, Human Genome Sequencing Center (2015) **Btau_5.0.1** *NCBI Genome Assembly*
79. Li H., Durbin R. (2009) **Fast and accurate short read alignment with Burrows-Wheeler transform** *Bioinformatics* **25**:1754–1760
80. Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., 1000 Genome Project Data Processing Subgroup (2009) **The Sequence Alignment/Map format and SAMtools** *Bioinformatics* **25**:2078–2079
81. Kircher M. (2012) **Analysis of high-throughput ancient DNA sequencing data** *Methods in Molecular Biology (Clifton, N.J)* Springer :197–228
82. Mitnik A., Wang C.-C., Svoboda J., Krause J. (2016) **A Molecular Approach to the Sexing of the Triple Burial at the Upper Paleolithic Site of Dolní Věstonice** *PLOS One* **11**
83. Green R. E. *et al.* (2008) **A Complete Neandertal Mitochondrial Genome Sequence Determined by High-Throughput Sequencing** *Cell* **134**:416–426
84. Peng M.-S., Fan L., Shi N.-N., Ning T., Yao Y.-G., Murphy R. W., Wang W.-Z., Zhang Y.-P. (2015) **DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals** *Molecular Ecology Resources* **15**:1238–1242
85. Dutta P. *et al.* (2020) **Whole genome analysis of water buffalo and global cattle breeds highlights convergent signatures of domestication** *Nat Commun* **11**
86. Chen N. *et al.* (2018) **Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia** *Nat Commun* **9**
87. Hu Q., Ma T., Wang K., Xu T., Liu J., Qiu Q. (2012) **The Yak genome database: an integrative database for studying yak biology and high-altitude adaption** *BMC Genomics* **13**
88. Korneliussen T. S., Albrechtsen A., Nielsen R. (2014) **ANGSD: Analysis of Next Generation Sequencing Data** *BMC Bioinformatics* **15**
89. Daetwyler H. D. *et al.* (2014) **Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle** *Nat Genet* **46**:858–865
90. Hayes B. J., Daetwyler H. D. (2019) **1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes** *Annual Review of Animal Biosciences* **7**:89–102

91. Naval-Sánchez M., Porto-Neto L. R., Cardoso D. F., Hayes B. J., Daetwyler H. D., Kijas J., Reverter A. (2020) **Selection signatures in tropical cattle are enriched for promoter and coding regions and reveal missense mutations in the damage response gene HELB** *Genetics Selection Evolution* **52**
92. R Core Team (2016) **R: A Language and Environment for Statistical Computing** Vienna, Austria: R Foundation for Statistical Computing
93. Maier R., Flegontov P., Flegontova O., Isildak U., Changmai P., Reich D. (2023) **On the limits of fitting complex models of population history to f-statistics** *eLife* **12**
94. Oksanen J. *et al.* (2020) **vegan: Community Ecology Package**
95. Skoglund P., Mallick S., Bortolini M. C., Chennagiri N., Hünemeier T., Petzl-Erler M. L., Salzano F. M., Patterson N., Reich D. (2015) **Genetic evidence for two founding populations of the Americas** *Nature* **525**:104–108

Author information

Torsten Günther

Human Evolution, Department of Organismal Biology Uppsala University, Uppsala, Sweden
ORCID iD: [0000-0001-9460-390X](https://orcid.org/0000-0001-9460-390X)

For correspondence: torsten.gunther@ebc.uu.se

Jacob Chisauksy

Human Evolution, Department of Organismal Biology Uppsala University, Uppsala, Sweden

M Ángeles Galindo-Pellicena

Centro Mixto UCM-ISCIH de Evolución y Comportamiento Humanos, Madrid, Spain

Eneko Iriarte

Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain
ORCID iD: [0000-0001-8365-5616](https://orcid.org/0000-0001-8365-5616)

Oscar Cortes Gardyn

Universidad Complutense Madrid, Madrid, Spain
ORCID iD: [0000-0001-7685-3980](https://orcid.org/0000-0001-7685-3980)

Paulina G Eusebi

Universidad Complutense Madrid, Madrid, Spain
ORCID iD: [0000-0002-0682-038X](https://orcid.org/0000-0002-0682-038X)

Rebeca García-González

Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain

Irene Urena

Centre for Palaeogenetics, Stockholm, Sweden

Marta Moreno

Instituto de Historia – CSIC, Madrid, Spain

Alfonso Alday

Área de Prehistoria, University of the Basque Country, Leioa, Spain

ORCID iD: [0000-0002-0024-3462](https://orcid.org/0000-0002-0024-3462)

Manuel Rojo

Department of Prehistory and Archaeology, Valladolid University, Valladolid, Spain

ORCID iD: [0000-0002-9317-8654](https://orcid.org/0000-0002-9317-8654)

Amalia Pérez

Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain

Cristina Tejedor Rodríguez

Department of Prehistory and Archaeology, Valladolid University, Valladolid, Spain

Iñigo García Martínez de Lagrán

Departamento de Prehistoria y Arqueología, UNED, Madrid, Spain

Juan Luis Arsuaga

Universidad Complutense Madrid, Madrid, Spain

ORCID iD: [0000-0001-5361-2295](https://orcid.org/0000-0001-5361-2295)

José-Miguel Carretero

Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain, Unidad Asociada de I+D+i al CSIC Vidrio y Materiales del Patrimonio Cultural (VIMPAC), Burgos, Spain

Anders Götherström

Centre for Palaeogenetics, Stockholm, Sweden

Colin Smith

Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain, Dept. of Archaeology and History, La Trobe University, Melbourne, Australia

ORCID iD: [0000-0001-5901-8780](https://orcid.org/0000-0001-5901-8780)

For correspondence: cismith@ubu.es

Cristina Valdiosera

Laboratorio de Evolución Humana, Universidad de Burgos, Burgos, Spain, Dept. of Archaeology and History, La Trobe University, Melbourne, Australia

ORCID iD: [0000-0003-4948-2226](https://orcid.org/0000-0003-4948-2226)

For correspondence: cevaldiosera@ubu.es

Editors

Reviewing Editor

Ziyue Gao

University of Pennsylvania, Philadelphia, United States of America

Senior Editor

George Perry

Pennsylvania State University, University Park, United States of America

Reviewer #2 (Public review):

Summary:

In this paper, the authors investigated the admixture history of domestic cattle since they were introduced into Iberia, by studying genomic data from 24 ancient samples dated to ~2000-8000 years ago and comparing them to modern breeds. They aimed to (1) characterize genomic variation of skeletal remains and concordance (or discordance) with morphological features; (2) test for hybridization between wild aurochs and domestic cattle; (3) test for correlation between genetic ancestry and stable isotope levels (which are indicative of ecological niche); and (4) test for previously hypothesized higher aurochs ancestry in a modern breed of fighting bulls.

Strengths:

Overall, this study collects valuable new data and tests several important hypotheses regarding the evolutionary history and genomic variation of domestic cattle in Iberia, such as admixture between domestic and wild populations, and correlation between genome-wide aurochs ancestry and aggressiveness.

Weaknesses:

Most conclusions are well supported by the data presented, with the strengths and caveats of each analysis clearly explained. The presence of admixed individuals in prehistorical periods strongly support hybridization between wild and domestic populations, although the evidence for sex-biased introgression and ecological niche sharing is relatively weak. Lastly, the authors presented convincing evidence for relatively constant aurochs ancestry across all modern breeds, including the Lidia breed that has been bred for aggressiveness for centuries.

Major comments:

As the authors pointed out, a major limitation of this study is uncertainty in the "population identity" for most sampled individuals (i.e., whether an individual belonged to the domesticated or wild herd when they were alive). Based on chronology, morphology and genetic data, it is clear the Mesolithic samples from the Artusia and Mendandia sites are bona fide aurochs, but the "population identities" of individuals from the other two sites are much less certain. Indeed, archeological and morphological evidence from El Portalon supports the presence of both domestic animals and wild aurochs, which is echoed by the inter-individual heterogeneity in genetic ancestry. Despite the strong evidence of hybridization, it is unclear whether these admixed individuals were raised in the domestic population or lived in the wild population and hunted, limiting the authors' ability to draw conclusions regarding the direction of gene flow.

In general, detecting sex-bias admixture is an inherently challenging problem, especially given limited data. The differential ancestry proportions (estimated by f_4 ratios) on autosomes and X chromosome are indicative of sex-biased hybridization and consistent with previous mtDNA results and other non-genetic data. However, as shown in Fig 3, the confidence intervals of X and autosomal estimates overlap for all but a couple of individuals, despite the overall trend of the point estimates. Moreover, even if there is significant difference, it only suggests existence of sex-bias but does not speak to the extent (unless further quantitative argument is made). Statements such as "it was mostly aurochs males who contributed wild ancestry to domestic herds" is too strong and may be interpreted as extreme bias. The authors did a good job noting the caveats of this analysis and down-toned

the statement in the main text, but claims regarding sex-bias hybridization that use the phrase "mostly" in the abstract and discussion need to be further weakened.

The stable isotope analysis is very under-powered, due to issues of categorization of wild vs domestic Bos, as discussed by the authors. Although the considerable overlap in stable isotope values between domestic and wild groups is consistent with shared ecological niche, but the absence of evidence (ie significant difference between groups) is not evidence of absence. Two alternative, non-mutually exclusive scenarios are (1) prevalent errors in classification of wild vs domestic individuals; (2) different ecological niches share similar isotope profiles. Thus, the claim "suggesting that wild and domesticated groups often did not occupy different niches in Iberia" is still too strong.

<https://doi.org/10.7554/eLife.93076.2.sa2>

Reviewer #3 (Public review):

Summary:

Günther and colleagues leverage ancient DNA data to track the genomic history of one of the most important farm animals (cattle) in Iberia, a region showing peculiarities both in terms of cultural practices as well as a climatic refugium during the LGM, the latter of which could have allowed the survival of endemic lineages. They document interesting trends of hybridisation with wild aurochs over the last 8-9 millennia, including a stabilisation of auroch ancestry ~4000 years ago, at ~20%, a time coincidental with the arrival of domestic horses from the Pontic steppe. Modern breeds such as the iconic Lidia used in bullfighting or bull running retain a comparable level of auroch ancestry.

Strengths:

The generation of ancient DNA data has been proven crucial to unravel the domestication history of traditional livestock, and this is challenging due to the environmental conditions of the Iberian peninsula, less favourable to DNA preservation. The authors leverage samples unearthed from key archaeological sites in Spain, including the karstic system of Atapuerca. Their results provide fresher insights into past management practices and permit characterisation of significant shifts in hybridization with wild aurochs.

Comments on revisions:

The authors have satisfactorily addressed my previous concerns. Last questions:

- How many MCMC iterations were run for Structf4? Can they show the likelihood of the last 10% of MCMC iterations? The results seem way too different for K = 4 vs. K = 5, but only for moo014 and moo019.
- I guess the authors also lack an "a" superindex in Table 1 for moo019.
- That Gyu2-related ancestry appears systematically for K=5 suggests that the Caucasus-related ancestry was already present in the pool that led to domesticates. Is it not important to discuss the implications of this possibility, for future analyses?
- If monophyletic, why choose between Bed3 and CPC98 if both could be combined as a single population to further reduce qpAdm and f4 confidence intervals?
- Why not combine all auroch Iberian samples as a single population for testing gene flow from this whole group of samples to ancient Iberian cattle? Would be the resulting coverage still too low?

- What is subindex 1 in the denominator of the f_4 ratio (main methods)?

Thanks for your efforts

<https://doi.org/10.7554/eLife.93076.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public Review):

My main concern is the use of the 700K SNP dataset. This set of SNPs suffers from a heavy ascertainment bias, which can be seen in the PCA in the supplementary material where all the aurochs cluster in the center within the variation of cattle. Given the coverage of some of the samples, multiple individuals would have less than 10K SNP covered. The majority of these are unlikely to be informative here given that they would just represent fixed positions between taurine and indicine or SNPs mostly variable in milk cattle breeds. The authors would get a much better resolution (i.e. many more SNPs to work with their very low genome coverage data) using the 1000 bull genome project VCF data set:

<https://www.ebi.ac.uk/ena/browser/view/PRJEB42783> which based on whole genome resequencing data from many cattle. This will certainly help with improving the resolution of qpAdm and f_4 analysis, which have huge confidence intervals in most cases. Right now some individuals have huge confidence intervals ranging from 0 to 80% auroch ancestry...

We thank the reviewer for this suggestion. We repeated our analyses with a SNP panel from Run 6 of the 1000 Bulls project presented in Naval-Sanchez et al 2020. This panel reduced standard errors and narrowed down confidence intervals for the ancient samples. Another consequence is that more single-source qpAdm models can now be rejected highlighting the abundance of hybridization. For our comparison to modern breeds, we still use the 700K dataset as it provides a set of different modern European cattle breeds.

I agree with the authors that qpAdm is likely to give quite a noisy estimate of ancestry here (likely explain part of the issue I mentioned above). Although qpAdm is good for model testing here for ancestry proportion the authors instead could use an explicit f_4 ratio - this would allow them to specify a model which would make the result easier to interpret.

We have added ancestry estimates from f_4 ratios to the manuscript and display them together with qpAdm and Struc- f_4 (as suggested by reviewer #3) in our new Table 1. We decided to keep all three different estimates to illustrate that results are not consistent for all analyses. An additional feature of qpAdm is the possibility that two source models can be rejected and additional ancestries can be identified.

The interpretation of the different levels of allele sharing on X vs autosome being the result of sex-bias admixture is not very convincing. Could these differences simply be due to a low recombination rate on the X chromosome and/or lower effective population size, which would lead to less efficient purifying selection?

Following this comment (and another comment referring to the X chromosome analysis by reviewer #2), we decided to remove sex bias from the title of our study and add more information on the caveats of this analysis. While estimating ancestry on the X chromosome can be difficult, we also add that our patterns are consistent with what has been suggested based on ancient mitochondrial data (Verdugo et al 2019). For Neolithic Anatolia, it has been suggested that the insemination of domestic cows by auroch bulls has been intentional or even ritual (Peters et al 2012). A recent parallel archaeogenomic study also concluded sex-biased introgression from autosomal, X-chromosomal and Y-chromosomal data (Rossi et al 2024). As our results are consistent with these previous studies as well as the lower differentiation of modern breeds on the X chromosome (da Fonseca et al 2019), we still consider the general pattern of our results valid even if the exact extent of sex bias is difficult to assess.

The authors suggest that 2 pop model rejection in some domestic population might be due to indicine ancestry, this seems relatively straightforward to test.

We had already performed this analysis of modeling their ancestry from three sources using qpAdm. The results are shown in Supplementary Table S6 and we now refer to this more explicitly in the text: “The presence of indicine ancestry can be confirmed in a qpAdm analysis using three sources resulting in fitting models for all breeds (Supplementary Table S6).”

The first sentence of the paper is a bit long-winded, also dogs were domesticated before the emergence of farming societies.

We rephrased the first sentence to “Domestication of livestock and crops has been the dominant and most enduring innovation of the transition from a hunter-gathering lifestyle to farming societies.”

It would be good to be specific about the number of genomes and coverage info in the last paragraph of the intro.

This information is included in the first paragraph of the results section and we decided to not duplicate the numbers in the preceding introduction paragraph to retain a flow for the readers.

Reviewer #2 (Public Review):

Summary:

In this paper, the authors investigated the admixture history of domestic cattle since they were introduced into Iberia, by studying genomic data from 24 ancient samples dated to ~2000-8000 years ago and comparing them to modern breeds. They aimed to (1) test for introgression from (local) wild aurochs into domestic cattle; (2) characterize the pattern of admixture (frequency, extent, sex bias, directionality) over time; (3) test for correlation between genetic ancestry and stable isotope levels (which are indicative of ecological niche); and (4) test for the hypothesized higher aurochs ancestry in a modern breed of fighting bulls.

Strengths:

Overall, this study collects valuable new data that are useful for testing interesting hypotheses, such as admixture between domestic and wild populations, and correlation between genome-wide aurochs ancestry and aggressiveness.

Thank you for highlighting the importance of our study and the potential of our dataset.

Weaknesses:

Most conclusions are partially supported by the data presented. The presence of admixed individuals in prehistorical periods supports the hypothesized introgression, although this conclusion needs to be strengthened with an analysis of potential contamination. The frequency, sex-bias, and directionality of admixture remain highly uncertain due to limitations of the data or issues with the analysis. There is considerable overlap in stable isotope values between domestic and wild groups, indicating a shared ecological niche, but variation in classification criteria for domestic vs wild groups and in skeletal elements sampled for measurements significantly weakens this claim. Lastly, the authors presented convincing evidence for relatively constant aurochs ancestry across all modern breeds, including the Lidia breed which has been bred for aggressiveness for centuries. My specific concerns are outlined below.

Contamination is a common concern for all ancient DNA studies. Contamination by modern samples is perhaps unlikely for this specific study of ancient cattle, but there is still the possibility of cross-sample contamination. The authors should estimate and report contamination estimates for each sample (based on coverage of autosomes and sex chromosomes, or heterozygosity of Y or MT DNA). Such contamination estimates are particularly important to support the presence of individuals with admixed ancestry, as a domestic sample contaminated with a wild sample (or vice versa) could appear as an admixed individual.

We thank the reviewer for this suggestion. Due to our low coverage data, we focused on estimating contamination from the mitochondrial data by implementing the approach used by Green et al (2008). We make the code for this step available on Github. While most samples displayed low levels of contamination, we identified one sample (moo013a) with a surprisingly high (~50%) level of contamination which was excluded from further analysis.

A major limitation of this study is uncertainty in the "population identity" for most sampled individuals (i.e., whether an individual belonged to the domesticated or wild herd when they were alive). Based on chronology, morphology, and genetic data, it is clear the Mesolithic samples from the Artusia and Mendandia sites are bona fide aurochs, but the identities of individuals from the other two sites are much less certain. Indeed, archeological and morphological evidence from El Portalon supports the presence of both domestic animals and wild aurochs, which is echoed by the inter-individual heterogeneity in genetic ancestry. Based on results shown in Fig 1C and Fig 2 it seems that individuals moo017, moo020, and possibly moo012a are likely wild aurochs that had been hunted and brought back to the site by humans. Although the presence of individuals (e.g., moo050, moo019) that can only be explained by two-source models strongly supports that interbreeding happened (if cross-contamination is ruled out), it is unclear whether these admixed individuals were raised in the domestic population or lived in the wild population and hunted.

The reviewer is pointing out an important topic, the unknown identity of the studied individuals. We have revised the text making clear that we do not know whether the individuals were hunted or herded. At the same time, their genomic ancestry speaks for itself showing that there was hybridization between wild and domestic and that different individuals carried different degrees of wild ancestry. In the revised version, we have added the unknown identity as well as the fact that our results can be affected by both, changes in human hunting and herding practices over time. Regardless of the exact identity of the

individuals, our results can still be seen as (a) evidence for hybridization and (b) changes in human practices (hunting and/or herding) and their relationship to bovids over time.

Such uncertainty in "population identity" limits the authors' ability to make conclusions regarding the frequency, sex bias, and directionality of gene flow between domestic and wild populations. For instance, the wide range of ancestry estimates in Neolithic and Chalcolithic samples could be interpreted as evidence of (1) frequent recent gene flow or (2) mixed practices of herding and hunting and less frequent gene flow. Similarly, the statement about "bidirection introgression" (on pages 8 and 11) is not directly supported by data. As the genomic, morphological, and isotope data cannot confidently classify an individual as belonging to the domesticated or wild population, it seems impossible to conclude the direction of gene flow (if by "bidirection introgression" the authors mean something other than "bidirectional gene flow", they need to clearly explain this before reaching the conclusion.)

We have removed “bidirectional introgression” from the text and replaced it with the more neutral term “hybridization”. Furthermore, we used the revision to mention at several places in the text that it is not clear whether the sequenced individuals were hunted and herded and that the observed pattern likely reflects changes in both hunting and herding practices.

The f_4 statistics shown in Fig 3B are insufficient to support the claim regarding sex-biased hybridization, as the f_4 statistic values are not directly comparable between the X chromosome and autosomes. Because the effective population size is different for the X chromosome and autosomes (roughly 3:4 for populations with equal numbers of males and females), the expected amount of drift is different, hence the fraction of allele sharing (f_4) is expected to be different. In fact, the observation that moo004 whose autosomal genome can be modeled as 100% domestic ancestry still shows a higher f_4 value for the X chromosome than autosomes hints at this issue. A more robust metric to test for sex-biased admixture is the admixture proportion itself, which can be estimated by $qpAdm$ or f_4 -ratio (see Patterson et al 2012). However, even with this method, criticism has been raised (e.g., Lazaridis and Reich 2017; Pfennig and Lachance, 2023). In general, detecting sex-bias admixture is a tough problem.

In response to this comment and another comment by reviewer #1, we decided to remove sex bias from the title. In the revised version of our study, we have now switched this analysis from f_4 statistics to comparing f_4 ratios between the X chromosome and autosomes (Figure 3). Furthermore, we have added more information on the caveats of this analysis citing the articles mentioned by the reviewer. At the same time, we highlight that our patterns are consistent with what has been suggested based on ancient mitochondrial data (Verdugo et al 2019). Unfortunately, the low coverage data does not allow to call Y chromosomal haplotypes which would also allow an analysis of the paternal lineage. But our results are consistent with additional examples from the literature: For Neolithic Anatolia, it has been suggested that the insemination of domestic cows by auroch bulls has been intentional or even ritual (Peters et al 2012) and there is a lower differentiation of modern breeds on the X chromosome (da Fonseca et al 2019). A recent parallel archaeogenomic study also concluded sex-biased introgression from autosomal, X-chromosomal and Y-chromosomal data (Rossi et al 2024). Similar to the broader hybridization signal, our interpretation does not depend on the estimates for single individuals as we describe the broader pattern. As our results are consistent with previous results based on other types of data, we still consider the general pattern of our results valid even if the exact extent of sex bias is difficult to assess.

In general, the stable isotope analysis seems to be very underpowered, due to the issues of variation in classification criteria and skeletal sampling location discussed by the authors in supplementary material. The authors claimed a significant difference in stable

nitrogen isotope between (inconsistently defined) domestic cattle and wild aurochs, but no figures or statistics are presented to support this claim. Please describe the statistical method used and the corresponding p-values. The authors can consider including a figure to better show the stable isotope results.

In combination with updated tables, we have added a supplementary figure showing the stable isotope results (S9). In light of the reanalysis of the genetic data, we have reassessed the genetic models used to assign species in the stable isotope analysis. We have provided more details of the statistical methods used and the p-values are given in the supplementary materials. There is a significant difference in the nitrogen isotope values when comparing *B. taurus* and *B. primigenius* (identified on morphology) but no other comparisons are significant at the $p = 0.05$ threshold. The reviewer highlights what we have mentioned in the supplementary material regarding the varied skeletal elements used for stable isotope analysis and the difficulty of assigning a species identity (as this depends on what criteria are used; morphological or some kind of genetic threshold of ancestry). Indeed, how to identify the species is at the heart of the paper. Given that identity could be defined in many ways, we have used 3 different genetic models to reflect this and the morphological categories, to help explore different possible scenarios. The reviewer is correct to point out that some of this analysis is not helped by the variety of skeletal elements used, but we have been careful not to over-interpret the results. The only samples that have nitrogen values higher than one standard deviation from the mean are domestic cattle, so it is not unreasonable to suggest that only domestic cattle have high nitrogen isotope values.

Reviewer #3 (Public Review):

Summary:

Günther and colleagues leverage ancient DNA data to track the genomic history of one of the most important farm animals (cattle) in Iberia, a region showing peculiarities both in terms of cultural practices as well as a climatic refugium during the LGM, the latter of which could have allowed the survival of endemic lineages. They document interesting trends of hybridisation with wild aurochs over the last 8-9 millennia, including a stabilisation of auroch ancestry ~4000 years ago, at ~20%, a time coincidental with the arrival of domestic horses from the Pontic steppe. Modern breeds such as the iconic Lidia used in bullfighting or bull running retain a comparable level of auroch ancestry.

Strengths:

The generation of ancient DNA data has been proven crucial to unravel the domestication history of traditional livestock, and this is challenging due to the environmental conditions of the Iberian peninsula, less favourable to DNA preservation. The authors leverage samples unearthed from key archaeological sites in Spain, including the karstic system of Atapuerca. Their results provide fresher insights into past management practices, and permit characterisation of significant shifts in hybridization with wild aurochs.

We thank the reviewer for their positive assessment of our work and for highlighting the strength and potential of the study.

Weaknesses:

- Treatment of post-mortem damage: the base quality of nucleotide transitions was recalibrated down to a quality score of 2, but for 5bp from the read termini only. In some specimens (e.g. moo022), the damage seems to extend further. Why not use dedicated tools (e.g. mapDamage), or check the robustness by conditioning on nucleotide transversions?

We agree that using such a non-standard data preparation approach requires some testing. Since our main analyses are all based on f_4 statistics, we compared f_4 statistics and f_4 ratios of our rescaled base quality data with data only using transversion sites. While estimates are highly correlated, the data set reduced to transversions produces larger confidence intervals in f_4 ratios due to the lower number of sites. Consequently, we decided to use the rescaled data for all analyses displayed in main figures. We also prefer not to perform reference based rescaling as implemented in mapDamage as it might be sensitive to mapping bias (Günther & Nettelblad 2019).

- Their more solid analyses are based on qpAdm, but rely on two single-sample donor populations. As the authors openly discuss, it is unclear whether CPC98 is a good proxy for Iberian aurochs despite possibly forming a monophyletic clade (the number of analysed sites is simply too low to assess this monophyly; Supplementary Table S2). Additionally, it is also unclear whether Sub1 was a fully unadmixed domestic specimen, depleted of auroch ancestry. The authors seem to suggest themselves that sex-biased introgression may have already taken place in Anatolia ("suggesting that sex-biased processes already took place prior to the arrival of cattle to Iberia").

We expanded the discussion on this topic but removed the analysis of whether European aurochs form a clade due to the low number of sites. We do highlight that a recent parallel study on aurochs genomes confirmed that Western European aurochs form a clade, probably even originating from an Iberian glacial refugium (Rossi et al 2024). Even if minor structure in the gene pool of European aurochs might affect our quantitative results, it should not drive the qualitative pattern. The same should be the case for Sub1 as our tests would detect additional European aurochs ancestry that was not present in Sub1. The corresponding paragraph now reads:

"A limitation of this analysis is the availability of genomes that can be used as representatives of the source populations as we used German and British aurochs to represent western European aurochs ancestry and a single Anatolian Neolithic to represent the original domestic cattle that was introduced into Europe. Our Mesolithic Iberian aurochs contained too little endogenous DNA to be used as a proxy aurochs reference and all Neolithic and Chalcolithic samples estimated with predominantly aurochs ancestry (including the 2.7x genome of moo014) already carry low (but significant) levels of domestic ancestry. However, the fact that all of these aurochs samples carried P mitochondria strongly suggests that western European aurochs can be considered monophyletic. Furthermore, a recent parallel study also concluded that Western European aurochs all form a clade (27). The Anatolian Sub1 might also not be depleted of any European aurochs ancestry and could not fully represent the original European Neolithic gene pool as also indicated by qpAdm and Struct-f4 identifying small proportions of other Asian ancestries in some Iberian individuals.

While these caveats should affect our quantitative estimates of European aurochs ancestry, they should not drive the qualitative pattern as our tests would still detect any excess European aurochs ancestry that was not present in Neolithic Anatolia."

Alternatively, I recommend using Struct-f4 as it can model the ancestry of all individuals together based on their f_4 permutations, including outgroups and modern data, and without the need to define pure "right" and "left" populations such as CPC98 and Sub1. It should work with low-coverage data, and allows us to do f_4 -based MDS plots as well as to estimate ancestry proportions (including from ghost populations).

We thank the reviewer for this suggestion. We added Struct-f4 as an analysis but observed that it would not converge in an individual-based analysis due to the low coverage of most of

our samples. We added Struct-f4 results for samples with >0.1X to the new Table 1, the results are similar to the results obtained using f4 ratios and (to a lower degree) the qpAdm results.

- In the admixture graph analyses (supplementary results), the authors use population groups based on a single sample. If these samples are pseudohaploidised (or if coverage is insufficient to estimate heterozygosity - and it is at least for moo004 and moo014), f3 values are biased, implying that the fitted graph may be wrong. The graph shown in Fig S7 is in fact hard to interpret. For example, the auroch Gyu2 from Anatolia but not the auroch CPC98 also from Anatolia received 62% of ancestry from North Africa? The Neolithic samples moo004 and moo014 also show the same shocking disparity. I would consider re-doing this analysis with more than a sample per population group

There seems to be some confusion relating to the sample identity in these figures. CPC98 is British and not Anatolian while Gyu2 is from the Caucasus and not Anatolia which would explain why they are different. Furthermore, moo004 is mostly of domestic ancestry while, moo014 is mostly of European aurochs ancestry according to our other analyses, which should explain why they also behave differently in this analysis. To avoid confusion and since this is a supplementary analysis from which we are not drawing any major conclusions, we decided to remove the graphs and the analysis from the study.

Recommendations for the authors:

Reviewer #2 (Recommendations For The Authors):

Fig 3A: The red regression line is misleading. It seems to show that the average aurochs ancestry fraction has been steadily decreasing since ~8000 years ago, but the "averaging" is not meaningful as not all samples necessarily represent domestic cattle remains and the sample size is rather small. In other words, the samples are just a small, random collection of domestic and wild animals, and the average ancestry is subject to large sampling noise. I would suggest removing the regression line (along with the associated confidence interval) in this figure. It would also be helpful to label the samples with their IDs and morphology in the plot for cross-reference with other figures. Also, it is said in the legend that "Modern Iberian breeds... are added around date 0 with some vertical jitter". Do the authors mean "horizontal jitter" instead?

Thank you for noticing this! We have removed the regression line and corrected the figure legend.

Fig 2 vs Fig 3A: are the error bars the same in these two plots? They seem to be highly similar, if not identical, but the legends read very differently ("95% confidence interval by block-jackknife vs. on standard error"). Please explain.

The figure legends have been corrected.

Fig 3B: What do the error bars in Fig 3B mean? 95% confidence interval or one standard error? Please clarify in the legend.

We have removed this figure and replaced it with a different way of displaying the results (now Figure 3). We ensured that the error bars are displayed consistently across figures.

According to the f4 statistics shown in Fig 1C and Fig 3B, moo012b carries a relatively high amount of domestic ancestry. How is this compatible with the observation in Fig 2 that this individual can be modeled with 100% aurochs (i.e., aurochs as the single source)? Does this simply reflect the low genome coverage?

moo012b is indeed one of the lowest coverage samples in our has at <0.02x sequencing depth. Even in our revised analysis using more sites, there is a discrepancy between the results of f4 statistics and qpAdm (suggesting mostly domestic ancestry) and f4 ratio suggesting mostly aurochs ancestry (Figure 1C and Table 1). We believe that this highlights the sensitivity of different methods to assumptions about the relationships of sources and potential “outgroups” which might not be well resolvable with low coverage data and in the presence of potentially complex admixture. Our general results, however, do not depend on the estimates for single individuals as our interpretations are based on the general pattern.

I don't fully understand the rationale behind the statement "However, at some point, the herding practices must have changed since modern Iberian breeds show approximately 20-25% aurochs ancestry". Can the stable ancestry fraction from 4000 years to the present (relative to the highly variable ancestry before) reflect of discontinuation of hunting rather than changes in herding practices?

We agree that this statement was not justified here, we rephrased the sentence to “In fact, from the Bronze Age onwards, most estimates overlap with the approximately 25% aurochs ancestry in modern Iberian cattle” and generally tried to make the text more nuanced on the issue of herding and hunting practices.

Reviewer #3 (Recommendations For The Authors):

Thanks for this interesting piece of work. The results are clearly presented, and I have no additional concerns other than those reflected in the public report, except perhaps:

(i) trying to use more informative sample names (eg. including the date and location). It may facilitate reading without going back and forth to the table "Sample List".

We have now added a main table listing our post-Mesolithic samples together with their age, site and estimated aurochs ancestry proportions. We hope that his table makes it easier for readers to follow our sample IDs.

(ii) Briefly describe in the main the age of aurochs and Sub1 not generated in this study.

Fixed.

<https://doi.org/10.7554/eLife.93076.2.sa0>