Evidence for early domestic yak, taurine cattle, and their hybrids on the Tibetan Plateau

Chen, Ningbo; Zhang, Zhengwei; Hou, Jiawei; Chen, Jialei; Gao, Xuan; Tang, Li; Wangdue, Shargan; Zhang, Xiaoming; Sinding, Mikkel Holger S.; Liu, Xuexue; Han, Jianlin; Lü, Hongliang; Lei, Chuzhao; Marshall, Fiona; Liu, Xinyi

Published in:
Science Advances

DOI:
10.1126/sciadv.adi6857

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NG

Citation for published version (APA):
Evidence for early domestic yak, taurine cattle, and their hybrids on the Tibetan Plateau

Ningbo Chen†, Zhengwei Zhang‡, Jiawen Hou†, Jialei Chen†, Xuan Gao‡, Li Tang#, Shargan Wangdue§, Xiaoming Zhang¶, Mikkel-Holger S. Sinding*, Xuexue Liu*,†, Jianlin Han*,†, Hongliang Lu*, Chuzhao Lei*, Fiona Marshall*, Xinyi Liu*‡

Domestic yak, cattle, and their hybrids are fundamental to herder survival at high altitudes on the Tibetan Plateau. However, little is known about their history. Bos remains are uncommon in this region, and ancient domestic yak have not been securely identified. To identify Bos taxa and investigate their initial management, we conducted zooarchaeological analyses of 193 Bos specimens and sequenced five nuclear genomes from recently excavated assemblages at Bangga. Morphological data indicated that more cattle than yak were present. Ancient mitochondrial DNA and nuclear genome sequences identified taurine cattle and provided evidence for domestic yak and yak-cattle hybridization ~2500 years ago. Reliance on diverse Bos species and their hybrid has increased cattle adaptation and herder resilience to plateau conditions. Ancient cattle and yak at Bangga were closely related to contemporary livestock, indicating early herder legacies and the continuity of cattle and yak husbandry on the Tibetan Plateau.

INTRODUCTION

People living at high altitudes on the Tibetan Plateau have long relied for survival on domesticated and wild mammals adapted to extreme conditions (1–8). Domestic yak (Bos grunniens), taurine cattle (Bos taurus), and their hybrids are indispensable to the everyday life of people on the Tibetan Plateau today (Fig. 1). Contemporary highland pastoralists rely on the strength and hardiness of domestic yak for transportation across vast mountainous terrain and for supplies of milk, meat, fiber, and dung for fuel (9). Cattle and yak-cattle hybrids complement yak, supporting herders under different conditions. Hybrids are hardy and productive at midelevations (approximately 2500 to 3500 m above sea level, henceforth m a.s.l.), and cattle are especially valued at low elevations. Little is known, however, about the timing and processes of yak domestication, the introduction of cattle across the Tibetan Plateau, and the early hybridization of yak and cattle.

Wild yak (Bos mutus) are sympatric with domestic yak and are especially well adapted to low oxygen, high solar radiation, and extremely cold conditions at altitudes above 4000 m a.s.l. (10). Genetic studies show that extant wild yak are closely related to the ancestors of domestic yak, suggesting local domestication (10, 11). Continuous admixture of wild yak with domestic yak is widespread (11, 12) and has adaptive advantages for yak herders (13–15). Scholars have also argued that hybridization between locally adapted Tibetan yak and introduced cattle facilitated a rapid adaptation of Tibetan cattle to hypoxia (11, 16, 17). This husbandry practice was used to extend the ranges of cattle and yak (16, 17). Hybrids produce more milk and meat and are stronger than cattle at high altitudes. They are better adapted to low elevations than yak and expand possibilities for milk and traction at the midelevations (2500 to 3500 m a.s.l.), where male hybrid “dzo” and female hybrid “dzomo” are commonly used today.

Exploration of the early history of bovine (cattle, yak, and hybrids) husbandry on the high Tibetan Plateau has been hindered by a lack of archaeological data because yak remains are rare, osteological discrimination between wild and domestic yak has not been fully investigated, and early cattle bones have seldom been found outside the northern fringes of the Tibetan Plateau. Arguments based on modern genetic data on yak population expansion and information on pollen and linguistic shifts suggest two possible scenarios for yak domestication (9). These include early selection on wild yak by mid-Holocene hunter-gatherers and/or late yak domestication at approximately 4000 to 3000 calibrated radiocarbon years before the present (cal B.P.), somewhat after the introduction of domestic sheep, goat, wheat, and barley to the region and cattle first appearing on the northeastern margins of the plateau (7, 9, 14, 18–21).

However, evidence that may constrain these interpretations is lacking. The analysis of milk proteins from human dental calculus has documented the role of milk from sheep and goats in early herder diets, but bovine milk has not been securely identified (8). Three sites yielded a few bones (<5 specimens) that could possibly...
be attributed to domestic yak, including Qugong (3480 to 3160 cal B.P.) on the southern Tibetan Plateau (22–25) and two Nuomuhong Culture sites (~3400 to 2450 B.P., Tawendaliha and Xiariyamebu) on the northeastern Tibetan Plateau (5). No quantitative zooarchaeological data are available, but a small, slender yak horn core was found at Qugong (22, 24), while single specimens were attributed to domestic yak at Tawendaliha and Xiariyamebu (5). The lack of established morphological markers for discriminating between domestic and wild yak, or ancient DNA (aDNA) analyses has offered no chance of confidently identifying early domestic yak at these sites. As a result, empirical data critical for the examination of domestication contexts, primary yak husbandry, and their integration into early pastoral and farming practices are needed.

In contrast, abundant fauna and genetic evidence indicate that taurine cattle were introduced to China via routes along the northern fringes of the Tibetan Plateau (26, 27). A spread to the northern edge of the plateau ~5000 B.P. has been suggested on the basis of cattle bones, although specimens from Majiayao Culture sites such as Shichaocun and Xishanping lack direct dates (20). Currently, the introduction of cattle along with sheep, goats, wheat, and barley ~3900 B.P. is better supported (18, 20, 26, 27). Beyond its northern fringe, archaeological records for ancient cattle are less common. Unlike the northern section, the southern portion of the plateau lies close to the ranges of South Asian aurochs (*Bos primigenius*) and early domesticated indicine cattle (*Bos indicus*). Cattle bones were recovered from Gebusailu on the southwestern Tibetan Plateau ~3600 to 3000 B.P., but whether they are taurine or indicine cattle is unknown (28). Among other possible sources, barley in this region could have been initially introduced from South Asia through a trans-Himalayan connection (29), and such a route could also have been followed by domestic indicine cattle. Although present-day Tibetan cattle are of taurine ancestry (16), aDNA analysis is required to verify the species and/or region of origin of cattle identified at sites on the southern Tibetan Plateau.
Bangga (3000 to 2200 B.P.) is one of the earliest agropastoral settlements and the only site on the southern Tibetan Plateau with abundant faunal remains to have been systematically excavated in recent decades (30). This context provides a rare opportunity to explore the history of early yak, cattle, and their hybrids. To identify early Bos taxa on the high-altitude Tibetan Plateau and investigate their initial management, we conducted zooarchaeological and genomic analyses of faunal specimens from Bangga. From a zooarchaeological assemblage of >10,000 mammalian specimens (31), we identified 193 specimens as belonging to genus Bos. Established morphological criteria (see Materials and Methods) (32) were used to identify domestic taurine or indicine cattle [number of identified specimens (NISP) = 32] and domestic and wild yak (NISP = 10). To further scrutinize their species, we sequenced 5 whole genomes from the ancient Bos samples and 11 genomes from extant Tibetan cattle. These data were jointly analyzed with publicly available ancient (n = 74) (16, 33) and present-day (n = 166) cattle (16, 34–40) and yak (n = 86) whole genomes (11, 21, 41, 42) (Fig. 1). Comparative analyses of ancient and modern mitochondrial DNA (mtDNA) and nuclear genome sequences, combined with zooarchaeological and contextual data, provide secure evidence for early domestic yak, their adaptive hybridization, and the integration of taurine cattle and yak into an agropastoral system at elevations of ~3750 m a.s.l. on the southern Tibetan Plateau 3000 to 2200 B.P.

Archaeological site of Bangga
The site of Bangga (29°05′13.66″N, 91°43′15.36″E, 3750 m a.s.l.) lies adjacent to the modern agropastoral village of Bangga near the Yarlung River Valley on the southern Tibetan Plateau (Fig. 1A). Recent excavations at Bangga led by the Tibetan Autonomous Region Cultural Relic and Conservation Institute and the Sichuan University, revealed eight stone enclosures and over 400 other features, including hearths, ash pits, and post holes. Ancient residential and corral structures were also identified (30). These material remains provided evidence for a cultural change in the first millennium BC associated with distinct ceramic and lithic assemblages and a barley-dominant agricultural system revealed by paleoethnobotanical analysis (30, 43). Lying above the Yarlung River Valley (valley bottom at ~3700 m a.s.l.), surrounded by high mountain grasslands (~3800 to 4200 m a.s.l.), the site had accessibility to a wide range of elevations and resources. The fauna was well preserved and dominated by domesticated animals, predominantly sheep (Ovis aries) and goats (Capra hircus), with a bovine (Bos) component and a few horse (Equus cf. caballus) bones (30). Traces of dog (Canis familiaris), woolly hare (Lepus cf. oioios tolus), musk deer (Moschus sp.), and Tibetan gazelle (Procapra pcticaudata) were also found (30). Isotopic data and age profiles indicated provisioning and high juvenile sheep mortality that was likely associated with stresses resulting from penning and extreme conditions (32).

RESULTS
Zooarchaeological analysis and radiocarbon dating of bovine remains from the Bangga site
From the >10,000 mammalian remains, we attributed 193 specimens from the early phase to the genus Bos (table S1). We differentiated the cranial and postcranial elements of cattle and yak using morphological indicators (size and shape characters; see Materials and Methods) (32, 44) and identified 31 (16.1%) taurine or indicine cattle and 13 (6.7%) domestic or wild yak specimens (Fig. 1, B and C, and table S1). Specimens identified as yak included carpals, proximal metacarpal, astragalus, first phalanx, and second phalanx (Fig. 1C and table S1). Measurements of the astragalus and phalanx specimens suggested that the body sizes of the Bangga yak were slightly larger than those of present-day domestic yak (Fig. S1). The lack of wild yak comparative material and well-studied morphological criteria for distinguishing between domestic and wild yak prohibited the clarification of yak domestication status. The presence, consumption, and discard of Bos species in the settlement were demonstrated by cranial and postcranial element representation (Fig. 1C and table S2), modification (e.g., chopping, cutting, and burning; table S2), and the distribution of bones in house structures and corrals (fig. S2). Mortality profiles based on 77 ageable Bos postcranial specimens revealed that most of the animals survived the first 4 years of life (fig. S3 and table S3), which indicated a high adult yak and cattle survivorship at Bangga. Accelerator mass spectrometry (AMS) measurements were conducted on two Bos bone samples used for aDNA analysis (BG-33, humerus, corresponding to cattle and BG-67, metacarpal, corresponding to yak), yielding radiocarbon ages of 2666 to 2342 (cattle) and 2699 to 2350 cal B.P. (yak) (all radiocarbon ages are reported at 95.4% probability range) (fig. S4 and table S4).

Molecular identification of the species of the Bangga bovine specimens
Genomic data from the bovine specimens of Bangga presented a singular opportunity to deepen our understanding of the timing and strategies of early herding of cattle and yak on the Tibetan Plateau and their relationship to current livestock. We selected five well-preserved bovine bones from Bangga for molecular assignment of species (Fig. 1B and table S5), a key step in establishing the presence or absence of domestic yak at this site. In the future, perhaps sufficient specimens will be excavated from the plateau to enable analysis at the population level. We performed aDNA laboratory work in the aDNA-dedicated laboratory of the Molecular Paleontology Research Group at Kunming Institute of Zoology, Chinese Academy of Sciences. Using initial shotgun sequencing to screen data, we first assessed the aDNA authenticity and preservation situation according to several key indices, including the terminal damage rate (TDR), average fragmental length, and mapping rate. The results showed that the average TDRs were ≥10%, except for the BG-15 sample, of which all libraries were treated with USER enzyme, and the average fragmental lengths ranged from 78 to 90 bp for the five samples, indicating their typical aDNA molecular characteristics (figs. S5 and S6 and table S6).

The mitochondrial genome (mitogenome) sequences of the five bovine specimens were first aligned to both domestic yak (GenBank accession no. NC_006380.3) and taurine cattle (V00654.1) mitogenomes. Combining shotgun and captured sequencing data, we obtained sequencing depths for the five bovine mitogenomes ranging from 7.16× to 256.72× (table S6). Next, we assigned the species of the five bovine mitogenomes by placing them within the currently known matrilineal ancestries of 76 present-day yak and cattle (45–49) (table S7). The maximum likelihood (ML) tree revealed that four of the five mitogenomes fell within domestic taurine cattle haplogroup T, while the remaining mitogenome was within the yak clade III of haplogroup A (Fig. 2A and figs. S7 and S8) (48, 49).
To obtain further insight into the whole-genomic ancestry of these five ancient bovine specimens, we mapped the captured sequencing data to the taurine cattle reference genome (ARS-UCD1.2) and obtained sequencing depths of their nuclear genomes ranging from 0.014× to 1.723× (table S6). All four Bangga taurine cattle were female, while the Bangga yak was male (see Materials and Methods, fig. S9, and tables S8 and S9). All these five whole-genome sequencing data were retained for subsequent analyses. We combined whole-genome sequences of 57 present-day cattle and yak for species identification of the five Bangga bovine specimens (16, 21) (table S10 and S11). Principal components analysis (PCA) of autosomal single-nucleotide polymorphisms (SNPs) confirmed these mitogenome assignments (Fig. 2B and fig. S10) by clustering the four Bangga taurine cattle specimens into the present-day East Asian taurine cattle group and the Bangga yak specimen into the extant yak group.

**Evolution of the Bangga taurine cattle**

To investigate the relationship of the Bangga taurine cattle with ancient and present-day East Asian taurine cattle, we first compared their mitogenomes with published data of mtDNA hypervariable fragments from 109 ancient cattle dated between 4000 and 2200 B.P. from sites across North China (20, 26, 27, 50–53) (table S12) and then compared them with 157 present-day cattle mitogenome sequences (table S7). The results showed that all four Bangga taurine cattle were classified into the T3 haplogroups, including two T3_{119} haplogroups that had extremely high frequencies in ancient East Asian taurine cattle and two other T3 haplogroups that were clustered Tibetan cattle (Fig. 3, A and B). Most of the 75 ancient cattle (~4000 to 3000 B.P.) mtDNA sequences (88.23%) were reported to belong to the T3_{119}, T3_{055}, and T4 haplogroups, indicating an ancient Southwest Asian origin for this maternal lineage ~4000 B.P. and its subsequent predominance in East Asia (Fig. 3, A and B, and table S11) (26, 27, 50). After ~3000 B.P., other T3 and T2 ancestries were introduced, leading to a partial replacement of the ancient mtDNA lineages and then a complex matrilineal ancestry in East Asian taurine cattle (Fig. 3A), which was consistent with the previous suggestion of two matrilineal ancestries for taurine cattle in East Asia (16). The shared ancient matrilineal ancestry of two Bangga taurine cattle with Shimao cattle dating to ~4000 B.P., together with the other two Bangga taurine cattle, clustered within extant Tibetan taurine cattle, indicated a matrilineal continuity of Tibetan cattle from the ~2500 years ago to the present.

A projection PCA was used to investigate the genetic affinities of the four Bangga taurine cattle nuclear genomes alongside 240 published genomes of 74 ancient and 155 present-day cattle as well as 11 newly sequenced genomes of cattle from the Tibetan Plateau (tables S10 and S14) (16, 21, 33, 35–37, 42). A clear division was observed among the African, Southwest Asian, and East Asian cattle, while all four Bangga taurine cattle were positioned between the older Shimao cattle in Northwest China (~3900 B.P.) and present-day East Asian taurine cattle, including the South Korean Hanwoo and Tibetan taurine cattle (Fig. 3C and fig. S11). This genetic structure was consistent with the findings of a previous study (33). The estimated ancestral proportions of all four Bangga taurine cattle were similar to those of the Shimao and Near Eastern cattle, confirming their origin from ancient East Asian taurine cattle in Southwest Asia (Fig. 3D and fig. S12). In the identity-by-state (IBS) neighbor-joining (NJ) phylogenetic tree, all four Bangga cattle fell...
within the same branch of the ancient Shimao and present-day taurine cattle from the Tibetan Plateau (fig. S13).

To further examine the genetic relatedness between the ancient Bangga taurine cattle and present-day Tibetan cattle, we calculated the outgroup $f_3$ statistic from whole-genome sequences to measure their shared drift (buffalo; Bangga, East Asian taurine cattle) (fig. S14 and table S15). This analysis was corroborated by the PCA result with the ancient Bangga taurine cattle to be highly similar to the Shimao cattle and present-day Changdu and Yushu cattle from Tibet and Qinghai Provinces, China. Taking the mtDNA and nuclear genome data together, we found that the Shimao cattle at low altitudes were related to the ancient Bangga and present-day East Asian taurine cattle, including those on the Tibetan Plateau. Our results demonstrated that early domestic taurine cattle from West Asia arrived on the southernmost Tibetan Plateau no later than ~2500 years ago.

**Genomic footprint of the ancient Bangga yak among extant wild and domestic yak**

As it was difficult to verify the domestic or wild status of the ancient Bangga yak using mitogenome data alone (34), we mapped its nuclear genome data to the reference yak genome (BosGru3.0) (42) and retrieved genome data at a depth of 0.072× for comparison with the published data of 86 present-day wild and domestic yak on the Tibetan Plateau (table S16). The projection PCA showed a clear division between the wild and domestic yak, where the ancient Bangga yak fell close to present-day domestic yak than to the wild yak (Fig. 4A). However, within the domestic yak, the ancient Bangga yak was still separated from present-day yak group (Fig. 4B), making it difficult to further reveal the relationship between the ancient Bangga yak and present-day domestic yak. The NJ phylogeny based on IBS estimates indicated the same genetic affinity, in which the ancient Bangga yak was clustered with present-day domestic yak (Fig. 4C). The outgroup $f_3$ analysis (wild yak; Bangga...
yak, domestic yak) revealed that the ancient Bangga yak had a relatively close affinity with the domestic yak (table S16). Overall, these results indicated that the yak from Bangga was clustered close to present-day domestic yak but not to present-day wild yak, suggesting its same genetic ancestry as the domestic yak that are now found in the region today. Our analyses also documented the diversification of domestic yak into their current genetic structure ~2500 years ago or by the time of their presence at Bangga.

**Genetic evidence of hybridization between the ancient Bangga yak and taurine cattle**

Because present-day cattle and yak can naturally hybridize to produce fertile dzomo and both ancient taurine cattle and yak were present at Bangga, we applied both the $D$ statistic and the $f_4$ ratio to test for possible gene flow between them in the past. The two $D$ statistics [(buffalo, Bangga yak; Shimao cattle, Bangga cattle) and (buffalo, Bangga cattle; Bangga yak, wild yak)] showed a significant allele sharing between the taurine cattle and yak at Bangga (table S17), whereas the $f_4$ ratios identified a low proportion of common alleles derived from the Bangga yak (10.9 to 12.0%) in the Bangga taurine cattle, both supporting a relatively early hybridization between the ancient taurine cattle and yak (Fig. 4D and table S18). We further investigated potential sources of gene flow using qpAdm and found that the Bangga taurine cattle were best described as having mixed ancestry with domestic yak (12.1 to 19.5%) (Fig. 4E). We also accounted for possible impacts of the ancient Bangga taurine cattle on present-day domestic yak on the Tibetan Plateau using the $D$ statistic (buffalo, Bangga cattle; wild yak, present-day domestic yak) and observed that all present-day domestic yak shared significantly different levels of alleles with the Bangga
taurine cattle (table S19). A similar pattern was also evident when the ancient Shimao cattle were included as a source group (Fig. 4F). We inferred that the early domestic taurine cattle in East Asia and on the Tibetan Plateau, including those from Bangga and Shimao, contributed to a considerable amount of gene flow to with early domestic yak on the Tibetan Plateau. The yak-derived aDNA evidence represented additional evidence of ancient hybridization between domestic yak and cattle and supported the argument for early hybridization based on the ancient Shimao cattle and present-day yak data (16). Our findings suggested that yak-cattle hybridization occurred on the Tibetan Plateau by 2670 to 2360 cal B.P.

**DISCUSSION**

Our aDNA results from the high-altitude agropastoral site of Bangga provided the earliest evidence for domestic yak and yak-cattle hybridization. They also documented the presence of taurine cattle on the southernmost Tibetan Plateau by ~2500 B.P. Combined with zooarchaeological data on the frequencies as well as the timing and location of domestic yak and cattle utilization, our findings provide insights into yak domestication, early Tibetan cattle, and the faunal legacies of early herders.

**Yak domestication**

Genetic simulations based on present-day yak samples have suggested that yak were domesticated as early as c. 7500 years ago (21), while other lines of evidence advocated possible later dates. Until now, however, direct evidence for early domestic yak was lacking. Comparative analyses of the genome data from the ancient Bangga yak with those of present-day wild and domestic yak indicated the presence of domestic yak at high altitudes in the southern Tibetan Plateau, 2670 to 2360 cal B.P. These results validated and amplified yak and cattle morphological results from Bangga, revealing the presence of domestic yak and taurine cattle. They also strengthened the likelihood of domestic yak at Qugong ~3480 to 3160 cal B.P. and Tawendaliha and Xiariyamakebu on the northern Tibetan Plateau ~3300 B.P., as suggested by morphological indicators (5, 24). Nevertheless, no present-day wild yak genome matched either ancient or present-day domestic yak ancestry, leaving the ancestral population(s) of domestic yak an open question.

Selection on domestic yak at Bangga that took place in a setting where ancient herders intensively managed sheep and goats through corolling and provisioning at a high lamb mortality, nevertheless, suggested environmental stress (31). Yak cranial and postcranial remains were present in houses and corrals, consistent with butchery near the settlement. Genomic evidence for hybridization between yak and cattle indicated adaptive introgression, likely associated with intentional selection (e.g., culling, castration, or breeding of individual yak or cattle) to reduce environmentally related mortality and to increase productivity (e.g., milk, meat, and traction). Large-scale aDNA sampling at earlier sites is needed to address the question of whether hunter-gatherers or earlier herders influenced initial domestication processes (9).

**Long-term genetic continuity in the taurine cattle of Tibetan agropastoralists**

Comparative analyses of whole genomic aDNA data from Bangga with the genomes of ancient and present-day Southwest Asian and East Asian cattle showed that early herders at Bangga kept domestic taurine cattle at high altitudes. The Bangga site is close to the Himalayan Mountain corridors and within the ranges of the South Asian bovine species of B. nambadicus and B. indicus, and the possible routes followed by interregional exchanges, including barley. However, the four 2500-year-old taurine cattle at Bangga were genetically similar to the ancient cattle from Shimao on the Loess Plateau (1100 to 1300 m a.s.l., ~3900 B.P.) and present-day Tibetan taurine cattle (26, 27, 50). These findings were consistent with the premises of an eastern dispersal of taurine cattle along the Inner Asian Mountains and the Hexi Corridor and supported the genetic continuity of taurine cattle from the Bronze Age to the modern day but not a parallel movement of indicine cattle over the Himalayas. This observation agreed with recent human genomic evidence that suggested an extensive genetic admixture among the populations on the Tibetan Plateau and in East Asia (55, 56).

**Domestic yak and taurine cattle contributed to an undocumented agropastoral system**

Zooarchaeological and genetic findings indicated that domestic yak and taurine cattle were integrated into an undocumented agropastoral system that allowed early pastoralists to survive challenges posed by extreme environments on the Tibetan Plateau and to occupy the settlement of Bangga by approximately 2500 B.P. Herders at this site relied on large numbers of sheep and goats, culled more cattle than yak or horses, and used domestic yak-cattle hybrids, which have been advantageous for a range of reasons including alitudinal adaption. The Yarlung River Valley is located at approximately 3700 m a.s.l., which is challenging for early cattle but is within the range of domestic yak and at the low end of the range of wild yak. Several modern genetic studies have identified gene flow between cattle and yak on the Tibetan Plateau (11, 16, 17, 50, 57). At low elevations, it is common today to hybridize yak with cattle (e.g., cattle bulls mating yak cows) to combine the environmental hardiness and strength of yak with the productivity (meat and milk) of cattle (13). However, direct evidence has been lacking regarding the earliest management of Tibetan cattle and yak. The unique specimens from Bangga site provided aDNA data for hybridization between taurine cattle and yak. These genetic data were complemented by a substantial faunal assemblage and detailed archaeological information that documented a setting, in which people managed yak and cattle together at ~3700 to 4000 m a.s.l. ~2500 B.P. and thus afforded opportunities for interbreeding and experimentation.

Ancient herders at Bangga integrated yak, cattle, and their hybrids along with sheep, goats, and horses to generate multiple livelihoods (e.g., meat, milk, fiber, dung fuel, traction, and transport) and a degree of redundancy that increased herders’ resilience to weather- and disease-related shocks as a result of different mortality risks, reproductive rates, culling potentials, temperaments, and roles in the agropastoral economy (58). Traces of butchery and burning showed that cattle and yak were eaten in the Bangga settlement. However, mortality profiles revealed that young cattle, yak or hybrids were not subject to culling or natural mortality, unlike small ruminants. Manure identified in the corral (30) indicated its use as fuel and fertilizer (7). This mixed livestock production system at the site of Bangga (44), together with barley and possible millet cultivation, revealed the coalescence of a diverse
form of agropastoralism on the southern Tibetan Plateau around 2500 B.P., many elements of which persist today.

Pastoral legacy
The genetic continuity among domestic yak and taurine cattle on the Tibetan Plateau revealed by our aDNA and present-day genomic data is notable. Although future excavations and aDNA sampling will undoubtedly increase variability, some level of continuity has been established. In many regions of the world, early domesticated animals were replaced, e.g., the complete genetic turnover of European pigs (59) and the replacement of taurine ancestry by indicine cattle in Africa and the Near East (33). On the central Tibetan Plateau, expected die-offs did not result in local extinction or replacement of domestic yak or taurine cattle, indicating their successful long-term adaptation to high altitudes. The lack of faunal turnover and the subsequent 2500 years of selection for survival under local conditions represent the lasting legacy of pastoralism on the Tibetan Plateau.

MATERIALS AND METHODS
Zooarchaeological analysis
Maximally identifiable body parts that can be attributed to the genus Bos (60, 61) were the focus of this analysis. Specimens were tallied to obtain the NISP. Diagnostic osteological traits were compared to modern comparative collections in the Zooarchaeology Laboratory at Sichuan University, China and to published comparative osteological criteria for differentiating taurine cattle and domestic yak (7, 32, 44, 62–66) during the taxonomic identification of the Bos remains. Osteomorphological research on Tibetan Plateau mammalian species was conducted at five institutions (the Smithsonian Institution, the Natural History Museum New York, the Field Museum, the Henan Provincial Institute of Cultural Heritage and Archaeology, and the Shaanxi Provincial Institute of Archaeology [also see (42)]). Elements with robust osteomorphological criteria for distinguishing taurine cattle and yak included mandibular teeth, atlas, scapula, humerus, carpals, metacarpal, tibia, malleolus lateralis, astragalus, and navicular-cuboid. We conducted morphometric measurements following Von den Driesch (32). Body part representation and bone modification (surface marks) were recorded following standard zooarchaeological practices (67–69). We reconstructed Bos mortality profiles based on the analysis of the fusion status of postcranial specimens (70, 71). Because of the lack of systematic data on epiphyseal fusion times of yak postcranial elements, the analysis of the ageing of Bos remains was based on the epiphyseal fusion data of taurine cattle (72, 73). We also correlated information on the spatial context of individual specimens with archaeological features (e.g., houses, manure, and pens).

Osteomorphological landmarks used in this study for differentiating taurine cattle and yak
The descriptions of validated landmarks are as follows: Mandibular teeth: Metaconid and paraconid of the fourth premolar (P4) of yak form an open “V” shape but a “O” shape on the P4 of taurine cattle. Atlas: In dorsal view, a notch is present on the caudal border in taurine cattle but absent in yak; and two foramina are present on each side of the atlas body in yak, but usually only one is present in taurine cattle. Scapula: The coracoid process is notable reduced in yak compared with taurine cattle. Humerus: The proximal border of the greater trochanter exhibits a single notch in yak but two lesser notches in taurine cattle. Lunate: In proximal view, the lateral border of the proximal articular facet is relatively smooth in yak, but it exhibits a step in taurine cattle. Scaphoid: In proximal view, the medial border of the proximal articular facet is relatively smooth in yak, but it has a projection near the anterior end in taurine cattle; in lateral view, the posterior border forms an oblique line in yak, but it is curved in taurine cattle. Metacarpal: In posterior view, the proximal articular facets are farther extended distally and reach the distal articular facets in yak than in taurine cattle; in posterior view, a marked distal-medial tuberosity is present in yak but absent in taurine cattle. Malleolus lateralis: The proximal articular surface at which the spine is located is restricted and forms a two-blade propeller shape in yak, but it is not restricted in taurine cattle; in lateral view, the distal border is more curved in yak than in taurine cattle. Astragalus: In dorsal view, a notch is present at the middle of the proximal rim of the trochlea tali distalis in taurine cattle but absent in yak. Navicular-cuboid: The plantar side of the lateral half exhibits a well-pronounced plantar prominence in taurine cattle, but it is nearly absent in yak; in the abaxial view, a marked distal tuberosity is present in taurine cattle but nearly absent in yak; in the proximal view, the small articular facet joins the articular facet for calcaneum in yak, but it is separated from the articular facet for calcaneum in taurine cattle.

Radiocarbon dating
We chose bone fragments from two animals, which were also used for aDNA extraction and species identification, for AMS radiocarbon dating at the Beta Analytic Testing Laboratory, Miami (fig. S4 and table S4). All chronological data were calibrated by the OxCal (v4.4) and the IntCal20 calibration curve (74, 75). All ages reported are relative to AD 1950 at a 95.4% probability range (referred to as “cal B.P.”).

aDNA extraction, library preparation, genome capture, and sequencing
We collected five bovine bone samples from the site of Banga for molecular identification (Fig. 1B and table S5). The extraction of aDNA, qualification of aDNA preservation, and construction of double-strand aDNA library (DSL) were performed following a previous report (60). In brief, each sample was independently handled together with a blank control, strictly following aDNA laboratory operation standards. Before the aDNA extraction, we conducted ultraviolet irradiation on the samples for ~2 hours and removed the surface layer using a sterile dentistry trill. For each sample, we drilled ~200 mg of bone powder for the aDNA extraction, which was achieved by using the silica adsorption protocol (4), resulting in 200 μl of extraction per sample.

For qualifying the DNA preservation, we performed polymerase chain reactions (PCRs) in a final volume of 50 μl using KAPA uracil + polymerase (KAPA BIOSYSTEMS), which can incorporate nucleotides across uracils and preserve the deamination-induced damage patterns of authentic aDNA (61). For the five samples, we treated remaining endogenous DNA (eDNA) extraction with uracil-DNA glycosylase (UDG) to remove uracil residues but retain abasic sites in the eDNA for the digestion by endonuclease VIII (Endo VIII).
through cutting the 5′ and 3′ sides of the abasic sites. Following this treatment, the DSL was prepared for genome capture using the RNA baits transcribed from cattle genomic DNA as in the reported method (76). Both pre- and postcapture PCRs were carried out using the KAPA HiFi Hot Start Polymerase protocol (KAPA BIO-SYSTEMS). We sequenced the constructed DSL aDNA libraries (including all later genomic captured libraries) using the Illumina HiSeq X Ten (PE-150) platform [Wuxi NextCODE Genomics (Shanghai) Co., Ltd.].

**Mapping, alignment, filtering, and mtDNA assembly from the Bangga bovine aDNA reads**

Adapter sequences of the paired-end reads were identified and removed using AdapterRemoval (v2.2.0) (77). The reads were separately aligned to mitogenome references of the taurine cattle (V00654.1) and domestic yak (NC_006380.3) using Burrows-Wheeler Alignment (BWA) (v0.7.5a-r405) (78). To improve the alignment to circular mitogenomes, a 30-bp sequence from the end was attached to the beginning of the mitogenome sequence. Reads were subsequently screened for PCR duplicates and a mapping quality of <30 using SAMtools (v0.1.19) (79). MtDNA coverage was calculated by using Qualimap (v2.2) (80) (table S6).

We retrieved the mtDNA data of five ancient bovine samples with coverage depths ranging from 7.16× to 256.72×. The average lengths of the bovine molecules ranged from 78 to 90 bp, the TDR of all short gun sequences was >10%, and the maximum mapping rate was 1.83%, all of which indicated typical aDNA characteristics. The final mtDNA BAM files were manually and visually inspected for all mutations against the reference sequences, converted to FASTA format, and subsequently used to assemble consensus mitogenomes with Mapping Iterative Assembler (v1.0).

**Phylogenetic construction with the Bangga bovine mitogenomes**

The consensus mitogenomes of the four Bangga cattle and one Bangga yak were aligned together with available known matrilineal ancestries of 76 mitogenomes of yak and cattle for species identification; B. grunniens (NC_012346.1) and Bubalus bubalis (NC020617) were used as outgroup, using ClustalW (v2) (81). For the Bangga yak, we collected published data of mitogenomes from 55 present-day yak. For the Bangga cattle, we first collected published data of hypervariable mtDNA fragments from 109 ancient cattle dated between 4000 and 2200 B.P. from 13 sites across North China and assigned these data according to lineage-specific sites and the results from the literature (table S12) (20, 26, 27, 50, 53). We also collected 157 cattle mitogenome sequences to compare the Bangga cattle and present-day cattle (16, 50) (table S7). Phylogenetic relationships were inferred using the ML approach within IQ-TREE (82).

**Alignment of the Bangga bovine aDNA reads to taurine cattle and domestic yak genomes**

Dataset 1 contained all five Bangga bovine aDNA genomes and the whole-genome sequences of 57 present-day cattle and yak (16, 21), including 12 yak (4 wild yak and 8 domestic yak) and 45 domestic taurine cattle (9 African cattle, 31 Asian cattle, and 5 European cattle) (tables S10 and S11) at an average sequencing depth of ~11.559×, for species identification of the five Bangga bovine specimens. Dataset 2 included the four Bangga taurine cattle aDNA data and 11 Tibetan cattle genomes sequenced in this study as well as previously published 155 present-day cattle genomes and 74 Near Eastern cattle aDNA dataset for phylogenetic analyses of the four Bangga taurine cattle (16, 33, 35, 36) (tables S10 and S14). Dataset 3 consisted of the single Bangga yak genome and 86 previously published wild and domestic yak genomes (representing 13 regions from the five provinces of Tibet, Qinghai, Sichuan, Gansu, and Yunnan in China) at an average sequencing depth of ~9.95× for phylogenetic analyses of the single Bangga yak (11, 21, 42) (table S11). All procedures involving sample collection and experiments were approved by the Animal Ethical and Welfare Committee, Northwest A&F University (approval no. DK2022065).

ARS-UCD1.2 plus Btaru5.0 Y were used as the taurine cattle references, while BosGru3.0 (GCA_005887515.2) was treated as the domestic yak reference. For present-day cattle and yak genomes, all quality-filtered reads were aligned using BWA (v0.7.5a-r405) (78). Duplicated reads were filtered using Picard (v2.20.3) (83). Reads around indels were realigned with GATK (v3.8) (84). For the aDNA data, all raw reads were processed through a standardized pipeline. Adapter sequences of the paired-end reads were identified and removed using AdapterRemoval (v2.2.0) (77), and read pairs were merged if expected index was observed, based on an overlap of at least 11 bp. Alignments were performed using BWA (v0.7.5a-r405) (78) with the parameters “aln” and “sames” based on the options "seeding disabled -l 1024" and "-r" to define read groups and produce unfiltered SAM files, which were subsequently converted to BAM files using the command samtools view -Sb. The BAM files were sorted using SAMtools (v0.1.19) (79) and merged using Picard (v2.20.3) (83). Duplicates were removed using Picard (v2.20.3) (83), and indel realignment was then performed using GATK (v3.8) (84). Unaligned reads and reads with mapping quality less Q25 were also removed using SAMTools. Coverage statistics calculations was performed using Qualimap (v2.2) (80) (table S6).

**Variant discovery and genotyping**

The commands “Bcf tools mpileup” and “call” of BCFTools (v1.9) (79) were used to call variants from the “high-confidence” sample of present-day cattle and yak genomes based on the options “-q 30 -C 50 -Q 20 -B -a FORMAT/AD, FORMAT/SP, FORMAT/ADF, FORMAT/ADR” and generate vcf files “-O z -m -o.” The indels and any variants within 3 bp of the indels were removed using the BCFTools (v1.9) (79) filtering expressions. Repeat regions as defined by the UCSC Browser and the Repeatmasker files (85) were removed. Tri- and quad-allelic sites were also discarded. Variants with coverage twice the mean coverage were marked as missing (./.). Heterozygous variants present in a single individual or in more than 75% of individuals were excluded.

**Ancient variant genotyping**

For eight ancient samples with >2× mean coverage, the above-defined high-confidence sites were called using the command “mpileup” of BCFTools (v1.9) (79) with the same options implemented for present-day samples but without both recalibration with command “mpileup -B” and filtering for variant sites. After the indels and sites within 3 bp of the indels were removed, tri- and quad-allelic sites were further discarded. Sites were set to missing as described above (> twice the mean coverage). For other samples with depths <2× mean coverage, no minimum
coverage filter was imposed, and while a maximum of four read coverages per site was permitted. Individuals were then pseudo-haploidized by random sampling, obtaining a read at each site, and setting that individual as homozygous for the corresponding allele.

The present-day samples of dataset 1 contained 28,399,925 SNPs. This call set was used to call the Bangga bovine aDNA data. Variant positions with missing data in all Bangga bovine aDNA data and all transitions were removed after removing the loci based on high levels of pairwise linkage disequilibrium (LD) using PLINK (v1.90) (86) with settings “-indipairwise 50 5 0.2,” resulting in 321,537 SNPs. The present-day samples of dataset 2 had 41,893,695 SNPs. After removing the loci based on high level of pairwise LD and all transitions, a further filter of sites with minor allele frequency < 0.05 resulted in 2,503,796 SNPs. This call set was used to call all the Bangga taurine aDNA data, and variant positions with missing data in all bovine aDNA data and all transitions were removed, lastly resulting in 2,338,022 SNPs. The present-day samples of dataset 3 included 19,591,862 SNPs. This call set was then merged with the Bangga yak aDNA data after removing the loci based on high levels of pairwise LD, transitions, and variants with missing data in the Bangga yak aDNA data, resulting in 68,511 SNPs.

Sex determination
To infer the genetic sex of each Bangga bovine sample, we counted the reads mapped to the X chromosome versus autosomes using SAMtools idxstats (79) and used a custom Python script (87). This script assumed that the obtained number of reads should reflect the chromosome copy numbers and chromosome lengths; thus, two binominal models were constructed, one for males and another for females, and a likelihood ratio test was used to distinguish between the two models. To ensure the reliability of identification results based on low-depth sequencing data, we also determined the genetic sex of the samples by computing the number of alignments to the Y chromosome as a fraction of the total number of alignments to both sex chromosomes (88). The sequence was evaluated by calculating the number of alignments to the Y chromosome (nY) as a fraction of the total number of alignments to both sex chromosomes (nX + nY), denoted as RY, where RY = nY/(nX + nY). The 95% confidence interval of RY was computed using a normal approximation. The identification results of the two methods were consistent.

Population genetic structure analysis
We evaluated the population genetic structure of three Dataset among present-day and ancient cattle and yak by performing PCA and constructing phylogenetic trees. We projected our data onto the PCA space using the Smartpca program and the Isoplot option “YES” of EIGENSOFT (89).

For all analyses using ANGSD (90), only autosomal SNPs were included. When triallelic sites (skipTriallelic 1) and transition (–rmTrans 1) were ignored, a SNP filter was applied (–SNP_pval 1e-6). We calculated a distance matrix using single-base IBS in ANGSD. The filters were as follows: –doIBS 1 -bad 0 -remove_bads 1 -doCounts 1 -doCov 1 -makeMatrix -minInd 2/3 sample number -GL 1 -minMapQ 25 -minQ 20 -doGlf 2 -doMajorMinor 1 -doMaf 1 -minMaf 0.05. NGXdata was then run over a range of K values from 2 to 5. The optimal value of K was determined by iteratively running NGXdata as described above 10 times and changing the seed value for each run. The log likelihoods of each run (all iterations of K = 3 to 5) were analyzed to determine the optimal value of K by using evalAdmix (91). A positive correlation of the pairwise residuals indicated a poor model fit and was used to identify the best-fitting value of K.

F statistics and D statistics
For f statistic–based analyses, ADMIXTOOLS v7.0.2 (92) was used to convert PLINK format files to eigenstrat format. D statistics were computed using ANGSD with random read sampling option “-doAbbaba 1.” River buffalo was used to define the ancestral allele. Biallelic sites were limited to transversions (–rmTrans 1), and the analysis was restricted to autosomal SNPs. D statistics results are displayed in Tables S17 and S19. Statistically significant results, defined by |z scores| ≥ 3. Introggression between the Bangga taurine cattle and yak was estimated using the qp4ratio software in ADMIXTOOLS (92) with water buffalo as the outgroup (table S18).

Admixture modeling with qpAdm
We applied qpAdm (v1520) from ADMIXTOOLS v7.0.2 (92) to model the Bangga taurine cattle ancestry. The “allsnps: YES” option was used to calculate every f4 statistic to determine the power of the tests without introducing bias. The Bangga cattle, Bangga yak, and Shimao cattle were considered as target and source populations, while river buffalo, Sub1 (ancient cattle in Anatolia), and Hereford cattle were included in the reference populations.


Acknowledgments: For the assistance of excavation at Bangga and analysis of the faunal assemblage recovered, we thank Z. Chen, L. Zhaxi, and Z. Ciren from the Institute for Conservation and Research of Cultural Relics of Tibetan Autonomous Region; D. Daji, X. Wei, and L. Qunpei from the Museum of Shannan City; X. Chen, H. Xu, J. Song, F. Han, D. Zhao, M. Zhou, X. Zhang, and F. Yang from Sichuan University; and residents at Bangga village. We are grateful to S. Ketelsen, M. Surovy, E. Hoeger, and N. Duncan at the American Museum of Natural History; J. Oosoky and D. Lunde at the National Museum of Natural History at the Smithsonian Institution; and A. Ferguson and J. Phelps Jr. at the Field Museum for facilitating comparative osteomorphological studies of bovine specimens at their facilities. We are grateful to colleagues for advice and the perspectives of Bos geneticists. We also thank the High Performance Computing Center of NWAFU for providing computing resources. Funding: This work was supported by the National Key R&D Programs of China (2021YFC1523600 and SQ2021YFF1000041), the earmarked grant CARS-37, the National Natural Science Foundation of China (31872317, 32372854, and 32102523), fellowships provided by Postdoctoral Science Foundation of China (2021 T140564 and 2020 M683587), the Natural Science Basic Research Program of Shaanxi (2021JQ-137), Foundational Research Programs for Universities of China (2452021153), the National Science Foundation of USA (2017247), the Carlsberg Foundation (CF20-0355), and CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources in Beijing (2023-YWF-ZX-02). Author contributions: Z.Z., F.M., Xinyi Liu, H.L., N.C., and C.L. designed the research. Z.Z. and X.G. conducted zooarchaeological analyses. N.C., J.Hou, and J.C. conducted aDNA analyses. H.L., Z.Z., L.T., S.W., and Xinyi Liu conducted excavation. N.C., Z.Z., H. L., F.M., and Xinyi Liu wrote the paper. C.L., F.M., Xuezue Liu, J.Han, H.L., M.-H.S.S., X.Z., N.C., Z.Z., and Xinyi Liu revised the paper. Competing interests: The authors declare they have no competing interests. Data and materials availability: The generated sequences for ancient and present-day samples are available from the Sequence Read Archive (SRA) under Bioproject accession number PRJNA788155 (www.ncbi.nlm.nih.gov/bioproject/PRJNA788155), PRJNA898088(www.ncbi.nlm.nih.gov/bioproject/PRJNA898088), and PRJNA905718 (www.ncbi.nlm.nih.gov/bioproject/PRJNA905718). All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 16 May 2023
Accepted 7 November 2023
Published 13 December 2023
10.1126/sciadv.adi6857
Evidence for early domestic yak, taurine cattle, and their hybrids on the Tibetan Plateau

Ningbo Chen, Zhengwei Zhang, Jiawen Hou, Jialei Chen, Xuan Gao, Li Tang, Shargan Wangdue, Xiaoming Zhang, Mikkel-Holger S. Sinding, Xuexue Liu, Jianlin Han, Hongliang Lü, Chuzhao Lei, Fiona Marshall, and Xinyi Liu

Sci. Adv. 9 (50), eadi6857. DOI: 10.1126/sciadv.adi6857

View the article online
https://www.science.org/doi/10.1126/sciadv.adi6857
Permissions
https://www.science.org/help/reprints-and-permissions