Most Beefalo cattle have no detectable bison genetic ancestry

Beth Shapiro[1,2], Jonas Oppenheimer[3], Michael P Heaton[4], Kristen L Kuhn[4], Richard E Green[3], Harvey D Blackburn[5], Timothy PL Smith[4]

1. Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA 95064, USA
2. Howard Hughes Medical Institute, University of California Santa Cruz, Santa Cruz, CA 95064, USA
3. Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA 95064, USA
4. USDA, ARS, U.S. Meat Animal Research Center, Clay Center, NE, USA
5. USDA, ARS, National Animal Germplasm Program, Fort Collins, CO, USA

Corresponding author: Beth Shapiro (bashapir@ucsc.edu)

Abstract

Hybridization is common among lineages in the genus *Bos*, often mediated through human management for the selection of adaptive or desirable traits. A recent example is the American Beefalo cattle breed, which was developed in the 1970s and defined as a hybrid between American bison (*Bison bison*) and cattle (*Bos taurus*). The American Beefalo Association typically require ⅜ bison ancestry to qualify as Beefalo. Here, we sought to characterize admixed ancestry among Beefalo as a component of a larger project to understand the role of hybridization in shaping present-day diversity in bison and cattle. We generated genomic data from 50 historical and present-day Beefalo and bison hybrids, including several important founding animals, as well as from 10 bison originating from commercial herds that represent potential sources of bison ancestry in Beefalo. We found that most Beefalo did not contain detectable bison ancestry. No individual Beefalo within our data set satisfies the ancestry requirements specified by the American Beefalo Association (ABA), although several Beefalo had smaller proportions of bison ancestry (2-18%). Some beefalo had detectable indicine cattle ancestry (2-38%), suggesting that hybridization of taurine and zebu cattle may contribute to morphological similarity between some Beefalo and bison. Overall, ancestry profiles of Beefalo and bison hybrid genomes are consistent with repeated backcrossing to either parental species rather than the breeding between hybrids themselves, implying significant barriers to gene flow between bison and cattle. Our results call into question the ⅜ bison ancestry targeted by the breed association and

demonstrate the value of genomic information in examining claims of interspecies gene flow among *Bos* species.

Introduction

Gene flow has been common among lineages within the genus *Bos*, including among cattle (*Bos taurus*), bison (*Bison bison*), yak (*Bos mutus*), and gaur (*Bos gaurus*) (Wu et al. 2018; Zhang et al. 2020). Much of this gene flow was facilitated by human livestock management and breeding. For example, hybridization led to gene flow between taurine (*Bos taurus*) and indicine (zebu; *Bos indicus*) cattle breeds in the Near East during the Bronze Age (Verdugo et al. 2019), and later in Africa (Kim et al. 2020) and North America (McTavish et al. 2013). Human-mediated gene flow also occurred between more deeply diverged *Bos* lineages, including Chinese cattle and banteng (Chen et al. 2018) and yak and Tibetan cattle (Wu et al. 2018). In each case, gene flow is believed to have conferred genetic benefits that led to easier management of hybrid herds or enabled adaptation to local environmental conditions (Zhang et al. 2020).

While admixture is common among *Bos* lineages, genomic incompatibilities often induce reproductive challenges to the hybrids. These incompatibilities are probably due to the considerable divergences across the *Bos* clade, with the oldest split between cattle and the other lineages dating to around three million years ago (Wang et al. 2018). Except for hybrids of taurine and indicine cattle, hybrids of all other *Bos* species follow Haldane's rule, in which male offspring of F1 crosses are sterile (Zhang et al. 2020). While taurine-indicine cattle hybrids are fertile, hybrid African cattle have uneven proportions of indicine nuclear and mitochondrial ancestry (Kwon et al. 2022; Ward et al. 2022), which may be due to mitonuclear incompatibilities. Analysis of the dynamics of speciation and recurrent gene flow among *Bos* would improve understanding of the evolutionary process of species divergence.

Beefalo are a purported example of interspecies *Bos* hybridization. The American Beefalo Association (ABA) defines Beefalo as a stable hybrid cross with ⅝ cattle and ⅜ bison ancestry ("American Beefalo Association"). The Beefalo breed was established in the early 1970s by Mr. D.C. "Bud" Basolo, a cattle rancher from California, through complex interbreeding of bison and cattle (Miller 1982). The establishment of this breed followed a long history of attempts to crossbreed bison and cattle for commercial production (Boyd 1908; Goodnight 1914). As early 20th century breeders had been unable to create a viable hybrid population due to reproductive difficulties (Deakin et al. 1942), the establishment of the beefalo breed was both surprising and celebrated (Miller 1982; "Business" 1973).

Basolo did not reveal the pedigree for his original Beefalo, but subsequent analyses had allowed better understanding of the breed origins. (Paim et al. 2020) illustrated two pedigree approaches that could create a 3/8 – 5/8 combination, both of which require first generation hybrid males to be fertile. However, analyses of Beefalo karyotypes showed that all Beefalo possess taurine Y chromosomes (Lenoir and Lichtenberger 1978), such that the initial cross would have to have been a bison cow and a taurine bull. Intriguingly, Stormont et al. used blood typing to show that foundational Beefalo almost completely lack bison-specific markers (Stormont et al. 1986), although the authors believed that bison ancestry became more prevalent in subsequent generations. Beefalo ancestry has not, however, been evaluated with genomic information, despite that the breed was granted a special roll stamp for voluntary federal inspection by the United States Department of Agriculture (USDA, 1984, 9 CFR part 352) and the ongoing sale and purchase of pedigreed animals.

Here, we present results of the first genome-wide investigation of the ancestry composition of the Beefalo breed. We generate whole-genome shotgun (WGS) data from 47 Beefalo, three bison hybrids reported to carry majority bison ancestry, and ten commercial bison that represent a source of bison ancestry in Beefalo, and co-analyze this data along with publicly available genomes from a range of bovids. These samples were donated to the USDA Agricultural Research Service (ARS) National Animal Germplasm Program (NAGP) collection and include ancestors from the early establishment of the breed, such as Joe's Pride, a foundational Beefalo bull that Mr. Basolo sold to a Canadian breeder's firm for $2.5 million in 1975 ("Most Expensive Cattle"). We find that most Beefalo, including key founding individuals, contain no detectable bison ancestry, while bison hybrids have approximately their reported bison ancestry. Indicine cattle ancestry is common among Beefalo, suggesting that breeding with indicine cattle, which share a number of phenotypic similarities with bison in commercially-relevant traits (Zeder 2006), was an early strategy in the establishment of the Beefalo breed. Finally, we find that some Beefalo contain small amounts of bison ancestry, ranging from 2-18%. While this is less bison ancestry than the targeted 37.5%, these genomes prove that bison and cattle can hybridize in some circumstances. However, we always infer minor parental ancestry to be in a heterozygous state, providing evidence only of backcrossing to either parent, rather than of breeding between hybrids themselves. This is inconsistent with a scenario in which a stable bison-cattle hybrid Beefalo population had been established.

Results

Sequencing Beefalo genomes

We generated genomic data from 47 registered Beefalo and three bison hybrids using preserved semen samples. Beefalo semen samples were obtained from the USDA, ARS, NAGP (Supplementary Table 1). Most of the NAGP Beefalo samples were collected during the 1970s and 80s, around the time when the breed was founded, and donated to NAGP around 2007. We retrieved bison ancestry composition for each sample from metadata accompanying NAGP samples. These 50 Beefalo and bison hybrids had reported bison ancestry ranging from 75% in a bison hybrid to 12.5% in several Beefalo, with the bison ancestry of most Beefalo reported as the breed-standard 37.5%. We sequenced each sample to 2.7x median genomic coverage (range: 1.1-39x; Supplementary Table 1). Seven prominent Beefalo were selected for higher coverage sequencing (>30x), including animals from the Basolo foundation herds and others with different proportions of reported bison ancestry.

We generated high coverage (>17.5x) genomes of ten bison from commercial herds (Supplementary Table 1). We added these to a larger data set of previously published bison genomes (Yang et al. 2020; Stroupe et al. 2022; Wu et al. 2018; Shirazi et al. 2022) and genomes from several breeds of zebu and taurine cattle (Heaton et al. 2016); Supplementary Table 1), as well as the genome of Buzz, an F1 Yellowstone bison-Simmental cattle hybrid (Oppenheimer et al. 2021). Our final data set included 41 bison, 26 cattle, and 51 Beefalo and bison hybrid genomes. For some analyses we also incorporated published genomes from outgroups, including yak (Qiu et al. 2012, 2015), gaur (Heaton et al. 2016), banteng (Heaton et al. 2016), and water buffalo (Sun et al. 2020).

Estimating bison ancestry in Beefalo

Most Beefalo samples cluster with taurine cattle in principal component analysis (PCA) conducted on bison and both cattle subspecies, indicating they do not contain appreciable bison ancestry (Fig. 1A). This includes several foundational individuals, such as Joe's Pride (NAGP5887). The first principal component (PC1) separates bison from cattle, with the F1 hybrid Buzz falling halfway along this axis, while PC2 distinguishes taurine and indicine cattle. While most Beefalo samples cluster with taurine cattle, bison hybrids (>50% bison ancestry) fall in an intermediate position between bison and cattle (Fig. 1A). These hybrids are closer to bison than is Buzz, confirming their majority bison ancestry. The few Beefalo that are slightly shifted toward bison in the PCA are unlikely to contain the breed-standard 37.5% bison ancestry (Fig 1B, C). Beefalo formed a cline between taurine and indicine cattle on PC2, suggesting indicine ancestry is a widespread and variable ancestry component found in Beefalo.

One animal (NAGP9109) fell entirely within the indicine cluster and is likely of 100% indicine origin. The reported pedigree in the NAGP for this animal lists its composition

as ½ Brahman, ¼ Charolais, ⅛ bison, 1/16th Hereford, and 1/16th Shorthorn, but the American Brahman Breeders Association records this animal (#309519) as purebred Brahman. We believe NAGP9109 was erroneously labeled as Beefalo by the contributors. This result highlights that PCA analyses of even low coverage genomes can uncover inconsistencies between genomic ancestry and animal metadata.

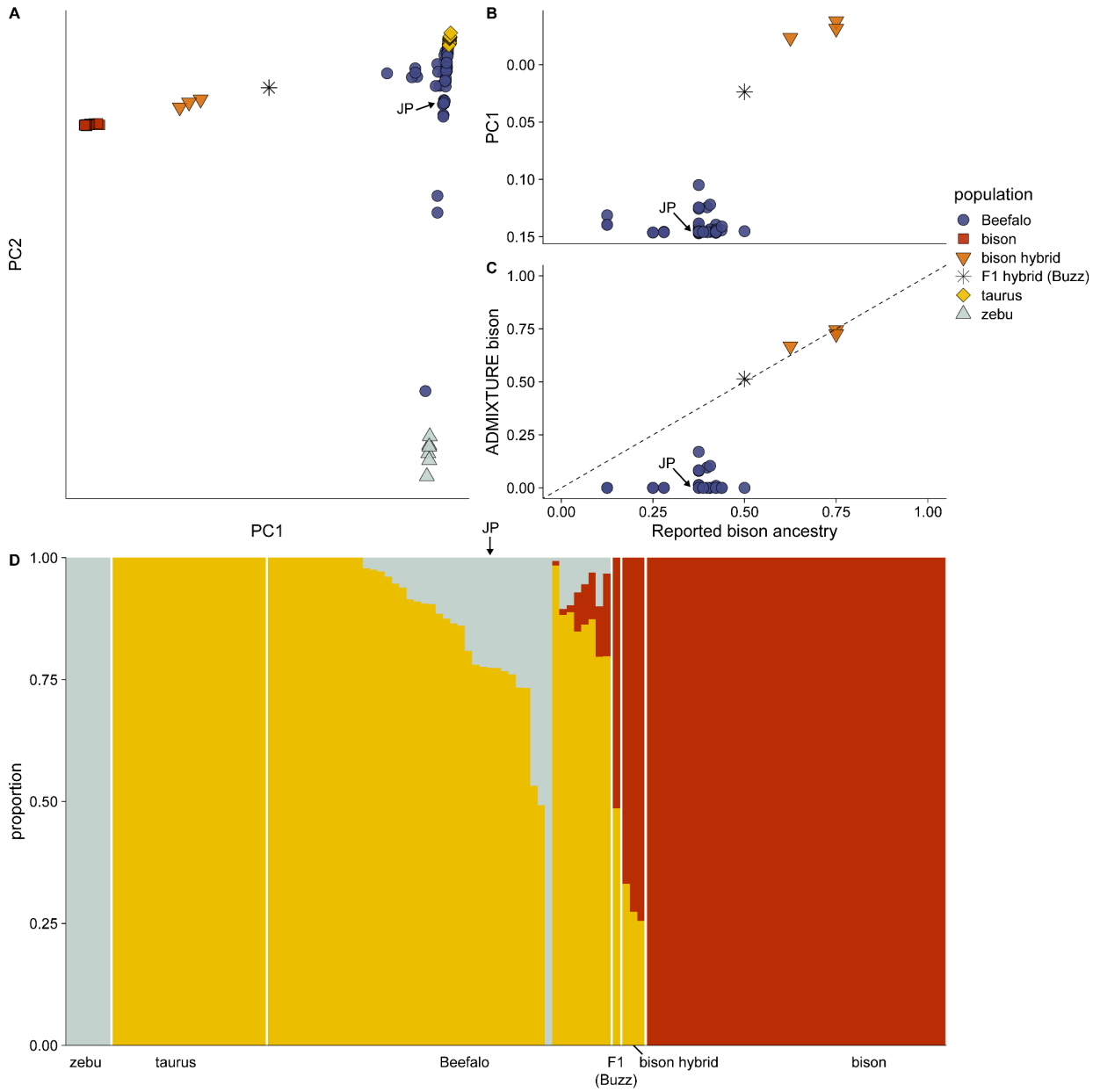


Figure 1. A) *Principal component analysis with Beefalo (blue), bison-cattle hybrids (orange), and a bison-cattle F1 hybrid (black star) projected onto axes computed from bison (red) and taurine (yellow) and indicine (gray) cattle. PC1 separates bison and cattle while PC2 separates the two cattle groups.* ***B)*** *Comparison of Beefalo reported bison ancestry and position on PC1. Placement of bison hybrids on PC1 is correlated*

with reported bison ancestry, though Beefalo do not follow this trend. ***C)*** *depicts the same as in* ***B)*** *though compares inferred bison proportions from ADMIXTURE to reported bison ancestry. Reported and inferred bison proportions match well for bison hybrids, but all Beefalo were inferred to have less bison ancestry than reported.* ***D)*** *Autosomal ADMIXTURE analysis of bison, taurine and indicine cattle, bison hybrids, and Beefalo. The position of Joe's Pride (JP) is depicted with an arrow in* ***A-D****.*

An unsupervised ADMIXTURE analysis supports the conclusion that most Beefalo have no bison ancestry. This analysis estimates the ancestry profiles present across all bison, cattle, bison hybrid, and Beefalo individuals examined for a given value of *K* source populations. At *K* = 3, bison, taurine cattle, and indicine cattle are divided into homogenous groups, with Beefalo assigned variable levels of these three ancestry types, mirroring the major patterns seen in the PCA (Fig. 1D). As in the PCA, bison-cattle hybrid backcrosses have bison ancestry estimates consistent with their pedigrees, while most of the sampled Beefalo, including Basolo founding individuals, possess no bison ancestry. Eight of the 47 registered Beefalo tested are, however, inferred to have a small amount (<18%) of bison ancestry, although less than was indicated by pedigrees and well below the breed standard defined by the ABA (Fig. 1C). These eight individuals are those shifted toward bison in PCA (Fig. 1A). The unsupervised ADMIXTURE analysis correctly assigned the F1 bison-Simmental genome equal amounts of taurine and bison ancestry, demonstrating that this analysis is sensitive to the presence of large fractions of bison ancestry. Supervised ADMIXTURE clustering, which models Beefalo ancestry as coming specifically from bison, taurine, and indicine source panels, provides similar estimates to those obtained from the unsupervised ADMIXTURE analysis (Supplementary Fig. S1).

Tests for excess allele sharing between Beefalo and bison relative to taurine cattle confirmed the results from PCA and ADMIXTURE, underscoring that most Beefalo have no appreciable bison ancestry, while those that do have bison ancestry have far less than that reported. *D*-statistics of the form *D(taurus, Beefalo; bison, water buffalo)*, which test whether Beefalo share more alleles with bison than taurine cattle, are consistent with 0 for most individual Beefalo, although the same eight Beefalo identified in PCA and ADMIXTURE as having bison ancestry also have an excess of bison alleles, confirming their bison ancestry (Fig. 2A). f4-ratios estimate that these Beefalo derive between 2-18% of their genomes from bison (Fig. 2B), consistent with estimates from ADMIXTURE and less than their reported bison ancestry, which ranged from 37.5-50%. *D*-statistics also confirm that many individual Beefalo have significantly more indicine alleles than is found in taurine breeds, demonstrating that indicine ancestry is widespread across Beefalo (Fig. 2C). Excess affinity with indicine cattle is observed in many Beefalo individuals with no evidence of bison ancestry, suggesting that allele

sharing with bison, which are deeply diverged from both taurine and indicine cattle, is detectable even in the presence of indicine ancestry, and vice versa.

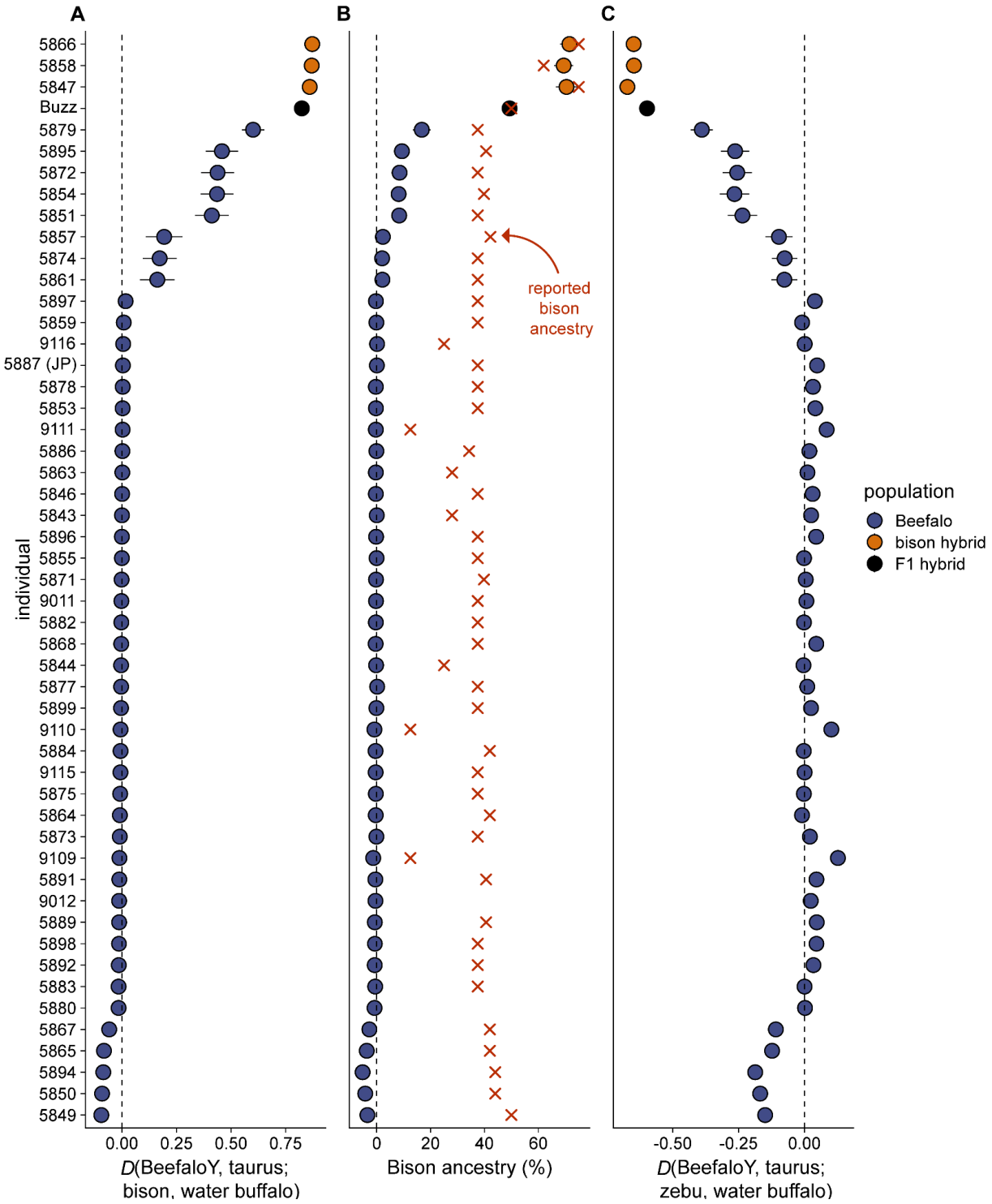


Figure 2. *Allele sharing statistics using individual bison hybrids and Beefalo. Individuals are shown on the Y axes. Beefalo are shown in blue and bison-cattle hybrids in orange,*

with a bison-cattle F1 shown in black. Error bars depict 3 standard errors. JP = Joe's Pride. ***A)*** *Testing for the presence of bison ancestry using statistics of the form* D(Beefalo, taurine cattle; bison, water buffalo)*.* ***B)*** *Quantification of the proportion of bison ancestry using f4-ratios. All bison hybrids have at least 50% bison ancestry, while only 8 Beefalo are inferred to have bison ancestry, ranging from 2-18%. Reported bison ancestry is shown in red.* ***C)*** *Testing for allele sharing between Beefalo and indicine cattle, relative to other taurine cattle. Many individual Beefalo are inferred to have significant allele sharing with indicine cattle. All Beefalo which were inferred to have bison ancestry in panels* ***A)*** *and* ***B)*** *display significantly negative values (due to the presence of bison ancestry, which is an outgroup to both cattle breeds), showing that any inferred indicine ancestry is not due to the presence of bison ancestry. For all panels, error bars depict 3 standard deviations.*

Local ancestry inference across individual Beefalo and bison-cattle hybrid genomes provides similar estimates of overall Beefalo ancestry, inferring an absence of bison ancestry across most Beefalo (Fig. 3). In Beefalo with bison ancestry, that ancestry tends to be present in large contiguous blocks, often tens of megabases in size, indicative of recent admixture (Fig. 3A,B). Bison ancestry in Beefalo is always found in a heterozygous state, consistent with a scenario in which these individuals are the product of repeated backcrossing of an initial F1 hybrid to cattle. Most Beefalo also have indicine ancestry (Fig. 3D), which, as with bison ancestry, is found in large blocks, but with many smaller blocks than seen with bison ancestry, suggesting that more generations have passed since the initial onset of hybridization (Fig. 3C).

Several lines of evidence attest to the efficacy of using these source panels for modeling ancestry. The F1 bison-taurine cattle hybrid was inferred to be almost entirely heterozygous for bison and taurine cattle across the genome (Supplementary Fig. 2) demonstrating the sensitivity of this approach to distinguishing diploid ancestry states between these groups. Further, the method assigned largely uniform and correct ancestry proportions to individual bison and cattle that were not part of the source panels, while Brangus cattle, a known taurine-indicine hybrid breed, were assigned large fractions of ancestry to both cattle groups in the expected amounts (Supplementary Fig. 3). For the seven Beefalo sequenced to high coverage, we also modeled ancestry using downsampled read counts and called genotypes, which yielded nearly identical results (Supplementary Fig. 4).

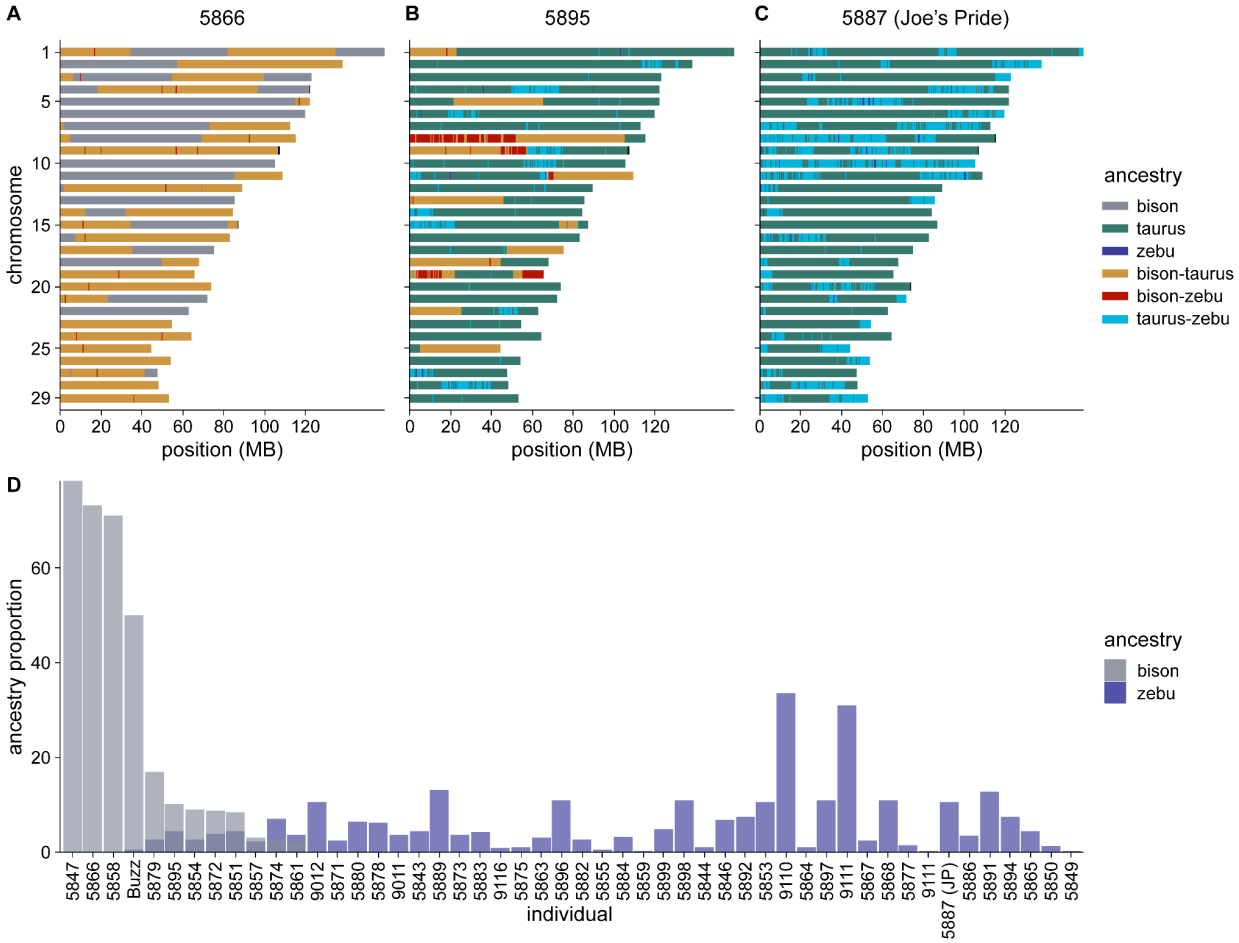


Figure 3. *Local ancestry inference of Beefalo and bison hybrid genomes, using bison and taurine and indicine cattle as potential sources. Inferred ancestry across the all autosomes are shown for* ***A****) a bison hybrid with reported 75% bison ancestry,* ***B****) a Beefalo with 37% reported, and 12% detected, bison ancestry, and* ***C****) Joe's Pride, a foundational Beefalo with 37.5% reported, and no detected, bison ancestry. Ancestry along the genome is colored by inferred source: homozygous bison, taurine, and zebu ancestry are shown in gray, green, and blue, respectively, with heterozygous bison-taurine, bison-zebu, and taurine-zebu ancestry shown in yellow, red, and light blue. Panel* ***D****) shows the inferred global bison and indicine ancestry proportions for each bison hybrid and Beefalo individual, with bison ancestry proportion shown in gray and indicine proportion shown in blue. Three bison hybrids are inferred to have ~75% bison ancestry, while eight Beefalo have detectable bison ancestry, ranging from 2-18%. Indicine ancestry is detected in most Beefalo at variable levels, ranging from 2-38%, with most Beefalo having between 2-18%.*

Beefalo sex chromosomal ancestry

Analyses of sex chromosomes reveal that the main mechanism for introducing bison ancestry into Beefalo was breeding with bison bulls. Most Beefalo have Y chromosomes associated with taurine cattle, while the bison hybrid backcrosses have bison Y chromosomes (Fig. 4B). Further, those Beefalo with evidence of bison ancestry have less bison ancestry on X chromosomes compared to autosomes (Fig. 4A,C). This is the opposite of what would be expected if bison ancestry were introduced maternally, which would result in an excess of bison rather than cattle ancestry on the X chromosome (Lenoir and Lichtenberger 1978). Unfortunately, as semen was the source of genomic DNA, we could not assemble mitochondrial genomes to assess whether the Beefalo maternal line would provide any additional insight into these patterns.

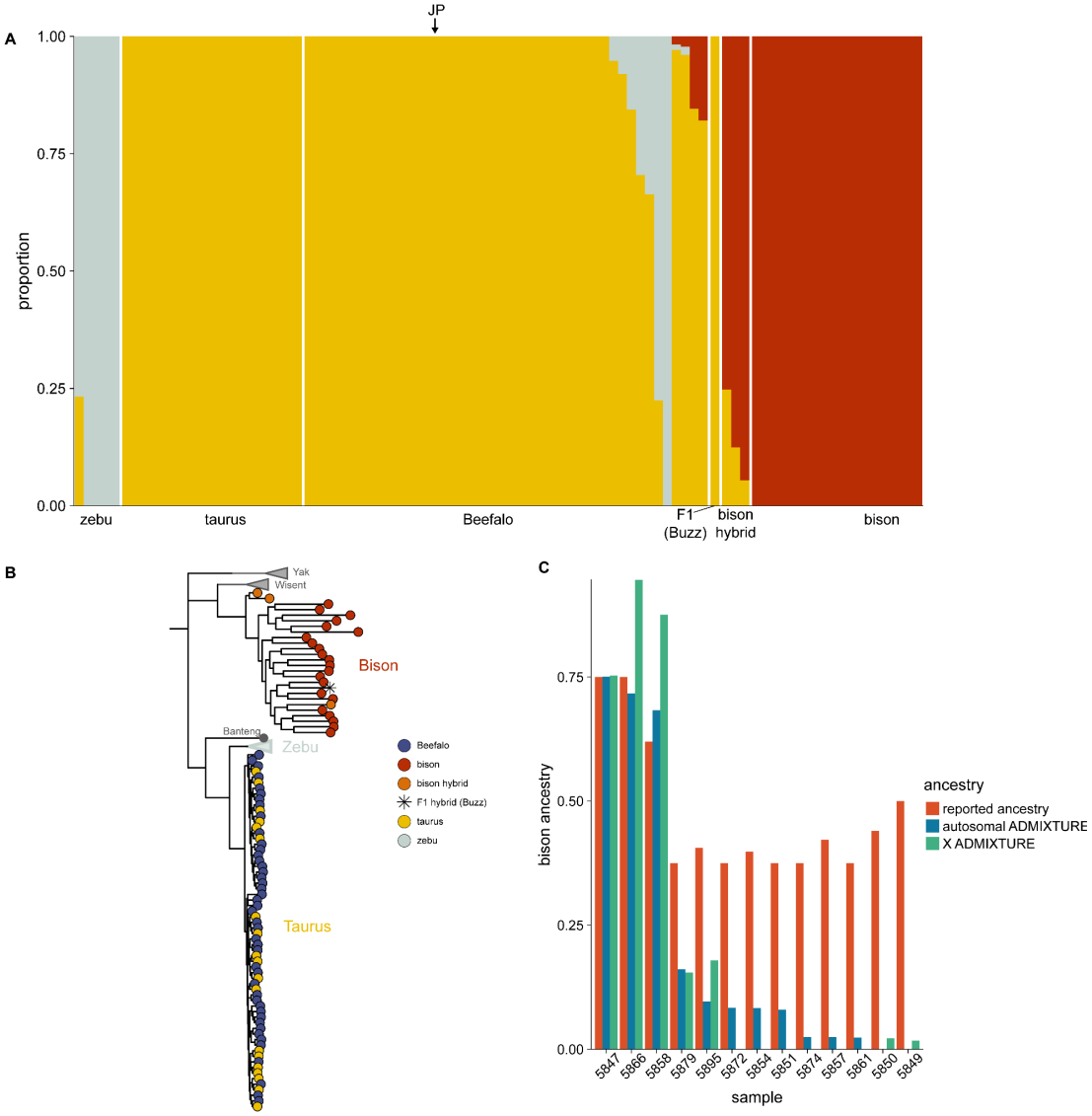


Figure 4. A) *ADMIXTURE analysis of bison, taurine and indicine cattle, bison hybrid, and Beefalo X chromosomes. Beefalo are largely inferred to have X chromosomes*

derived entirely from taurine cattle, though variable amounts of indicine and bison ancestry are also present in some Beefalo. The position of Joe's Pride (JP) is indicated with an arrow. Bison hybrids have majority bison ancestry. ***B)*** *Bovid Y chromosomal phylogeny. The indicine (zebu) clade is collapsed, as are those containing yak and wisent. Water buffalo (not shown) was used as an outgroup. Beefalo all have taurine Y chromosomes, while bison hybrids have bison Y chromosomes.* ***C)*** *Comparison of the reported (red) and ADMIXTURE inferred bison ancestry proportion on the autosomes (blue) and X chromosome (green) for the three bison hybrids and eight bison which had detectable autosomal bison ancestry.*

Discussion

Only eight of the 47 Beefalo that we sampled contained detectable bison ancestry, and those eight had substantially less (2-18%) than the 37.5% specified as the breed standard set by the ABA. Our samples represent a complete survey of individuals in the USDA NAGP biobank, and include important foundational animals, such as those from original Basolo herds. Notably, important foundational individuals, including Joe's Pride (NAGP5887), which was sold for $2.5 million to a Canadian breeders group ("Most Expensive Cattle"), lack bison ancestry. While these results show that interbreeding between bison and cattle is possible, they also prove bison ancestry encompasses a much smaller portion of the Beefalo breed than has been claimed.

Our finding of little to no bison ancestry in founding Beefalo individuals is aligned with challenges hybridizing *Bison* and *Bos* (Goodnight 1914; Murdoch 2018). Artificial insemination of taurine cows with bison semen results in 77% calf mortality and sterile male calves (Anstey 1986). In fact, Canadian government research reported that no functional males carrying more than 12.5% bison ancestry were ever observed (Brower 2008). In Beefalo individuals with bison ancestry, such ancestry was found exclusively in a heterozygous state, implying that repeated backcrossing to parental species, rather than breeding of hybrids themselves. A depletion of bison ancestry on the X chromosome and the presence of taurine Y chromosomes in all Beefalo, with bison hybrids exclusively having bison Y haplotypes, is consistent with this scenario. Therefore, it seems that extensive reproductive barriers exist to establishing a hybrid bison-cattle population.

While bison ancestry was surprisingly underrepresented in our Beefalo sample, nearly half of the Beefalo genomes we sequenced contained some zebu cattle ancestry, suggesting zebu/taurine interbreeding may have been used as a mechanism to manipulate the Beefalo phenotype. Zebu cattle are known for heat and drought tolerance (Kumar et al. 2016; Vajana et al. 2018; Hansen 2004), lower nutritional demands relative to taurine cattle (Hunter and Siebert 1985; Hennessy, Williamson, and

Darnell 2000) and their humped appearance (Heath 1979), and so share a number of desirable attributes with bison. The Beefalo breed standard does not require cattle ancestry to be of taurine origin, though if early breeders intentionally incorporated indicine ancestry while creating the breed, they left this detail out of the animals' reported pedigrees. For example, the Basolo founder individual Joe's Pride has a recorded pedigree that attests Hereford, Charolais, and bison (37.5%) ancestry, but we estimated that he had ~10.5% zebu ancestry.

This study is the first examination of Beefalo ancestry using whole genome data, but it is not the first to question claims surrounding bison ancestry in the breed. In an investigation of the paternal origins of Beefalo, Lenoir and Lichtenberger (1978) used karyotyping to determine that all 12 Beefalo bulls they examined had cattle Y chromosomes, in concordance with our results. While there are plausible scenarios for Beefalo with 37.5% bison ancestry to have cattle Y chromosomes, these require early-generation crosses to be fertile. In a later study, Stormont et al. (1986) used species-specific blood group markers to show that only one of the 148 Basolo Beefalo they tested had any alleles found predominantly in bison. However, they also note that some later Beefalo, separate from the initial founding herds, did display some bison-associated markers, perhaps agreeing with our finding of a small fraction of Beefalo possessing appreciable bison ancestry. Blood typing provides limited insight into ancestry proportions, however, demonstrating the utility of genomic information in validating specific breeding claims.

Our genome-wide analyses were primarily conducted on ~2x coverage genomes with Beefalo samples obtained from the USDA NAGP, but we believe our Beefalo ancestry estimates are robust and representative of the breed at large. We derive highly concordant ancestry estimates across a range of approaches, including ADMIXTURE, f4-ratios, and local ancestry inference techniques, and for animals whose ancestry is known, such as an F1 bison-cattle hybrid, these methods assign the correct ancestry proportions. Additionally, we derive similar ancestry estimates in high coverage (30-42x) and downsampled data for the seven individuals we sequenced more deeply. We did not sample present-day sources of Beefalo exhaustively, instead choosing to focus largely on animals that originate at or near the founding of the breed. Beefalo herds are generally established by breeding with fullblood (37.5% bison) Beefalo individuals, rather than through backcrosses with bison ("American Beefalo Association"), so surveying breed founders is an effective way of documenting the bison ancestry across the breed. However, a larger sample of Beefalo across current producers would provide a more comprehensive understanding of the breadth and variation of bison ancestry in Beefalo.

The extent of gene flow and the existence of reproductive barriers among Bos species remain incompletely explored. Breeders worked throughout the 20th century to create commercially-viable bison-cattle hybrids using bulls and cows with different ancestries. All of these efforts, including Beefalo, failed to create a stable hybrid population, implying that considerable barriers to gene flow exist between these two species. Though genomic incompatibilities among bison and cattle are hinted at by these breeding efforts themselves, which reported sterility in male hybrids even after backcrossing to cattle for several generations (Deakin et al. 1942; Brower 2008), they have not been detected directly. The general absence of bison ancestry in Beefalo also calls attention to several other examples of assumed gene flow among Bos species that have yet to be fully characterized. This includes introgression from cattle into wisent (Wecek et al. 2017; Soubrier et al. 2016) and American bison, the latter of which has been suggested to be widespread over the past two centuries and has led to the presence of cattle ancestry in all bison today (Stroupe et al. 2022; Halbert and Derr 2007), though has not been fully examined with genomic evidence.

Materials and Methods

Genomic data collection

Beefalo semen samples (n = 47) were obtained from the USDA, ARS, NAGP collection at Ft. Collins Colorado in the USA (Supplementary Table 1). The three bison hybrid semen samples were also obtained from the NAGP. Most of the NAGP samples were collected in the 1970s and 1980s, stored by breeders, and donated to NAGP circa 2007 as a geographically diverse set of Beefalo. Purebred bison meat samples (n = 10, tongue) were purchased from commercial processors and represent animals from three different commercial bison herds (Supplementary Table 1).

DNA was extracted from semen samples with a modified standard phenol-chloroform method. Briefly, one 0.5-ml straw of semen was diluted with 1.5 ml of a solution containing 10 mM TrisCl, 100 mM NaCl, 1 mM EDTA, pH 8.0) with 1% wt/vol sodium dodecyl sulfate, 1 mg proteinase K, and 40 mM dithiothreitol. This lysis solution was incubated overnight at 37°C in a 15 ml conical tube preloaded with 2 ml of high-vacuum grease in preparation organic phase extractions. The lysed and digested solution was extracted twice with 1 vol of phenol:chloroform:isoamyl alcohol (25:24:1), and once with 1 vol of chloroform. For each extraction the sample was centrifuged for 10 minutes in a swinging bucket rotor at 3210 x g at 23C to partition the organic phase below the band of high-vacuum grease, while to the aqueous phase was held above. The DNA was precipitated with 0.1 vol of 3 M sodium acetate (pH 5.2) and 2 vol of 100% ethanol. The precipitated DNA was washed once in 70% ethanol, briefly air dried, and dissolved in a solution of 10 mM TrisCl, 1 mM EDTA (pH 8.0). The bison tongue DNA was extracted

with the Qiagen Blood and Cell Culture DNA Mini Kit according to the manufacturer's instructions (Qiagen, Venlo, The Netherlands).

We sheared extracted DNA using a Covaris ultrasonicator (Covaris, Inc. Woburn, MA) prepared into Illumina sequencing libraries using either the NEB Ultra II FS kit (NEB, Ipswich, MA) or the TruSeq PCR-free DNA kit (Illumina, San Diego, CA). We quantified libraries with a Qubit fluorometer using the 1x HS kit (Thermo Fisher, Waltham, MA) or, for PCR-free libraries, via qPCR with the Kapa Complete Universal Kit (Roche Sequencing, Santa Clara, CA). Library fragment length distribution was assessed using either a Tapestation 2200 (Agilent, Santa Clara, CA) or a Fragment Analyzer (Advanced Analytical Technologies Inc., Ames, IA). For whole genome sequence (WGS), 1 µg of genomic DNA was fragmented and used to make indexed, 350 bp, paired-end libraries.

Beefalo and bison libraries were sequenced with Illumina instruments using 2 × 150 bp paired-end kits, either on the NovaSeq 6000 for bison or on the NextSeq 2000 platform for Beefalo and bison hybrids.

Variant Calling

Remnant adapter sequences were trimmed from reads using Trimmomatic (v0.39) (Bolger, Lohse, and Usadel 2014), requiring a minimum length of 30bp. Trimmed reads were then mapped to the cattle genome ARS-UCD1.2 with the Y chromosome from Btau5.1 appended (ARS-UCD1.2_Btau5.0.1Y) using BWA (v0.7.17-r1188) mem (Li 2013) with default parameters, except for 9 Beefalo samples that were mapped with BWA aln (Li and Durbin 2009). These samples had a lower fraction of reads that were properly paired, with higher rates of interchromosomal mapping within read pairs, possibly suggesting chimera formation. As BWA aln conducts end-to-end alignment, it mitigates the effects of any spurious mapping of incompletely clipped chimeric sequences. Duplicate reads were removed with Picard MarkDuplicates (https://broadinstitute.github.io/picard/).

We called genotypes in four medium-to-high coverage (>10x) gaur genomes (Heaton et al. 2016; Verdugo et al. 2019; Wu et al. 2018) using GATK HaplotypeCaller (v4.1.8.1) (DePristo et al. 2011) to ascertain a set of variants from an outgroup to both cattle and bison for use in downstream analyses. Variants were filtered variants for a minimum genotype quality of 30 and minimum and maximum depths of 1/3rd and 2x the mean coverage on a per-sample basis. We also performed mappability filtering with SNPable (https://lh3lh3.users.sourceforge.net/snpable.shtml), using a *k*-mer length of 35 and stringency of 0.5 (gen_mask -l 35 -r 0.5).This yielded a set of 5.29M high-quality autosomal variants.

To accommodate the low sequencing depth we obtained from most Beefalo, we used psuedohaploid genotypes, in which a random read covering each gaur-ascertained variant was selected to represent genotypes at those sites. This approach mitigates biases arising from differential coverage across individuals (Barlow et al. 2020). Pseudohaploid genotypes were called for all individuals using PileupCaller using samtools (v1.9) mpileup (Li et al. 2009) with BAQ disabled (-B) and pileupCaller (https://github.com/stschiff/sequenceTools.git), requiring a minimum map quality of 25 and minimum base quality of 30.

Modeling Beefalo ancestry

We visualized the relationships of Beefalo and bison hybrid individuals to bison and cattle using a Principal Component Analysis (PCA). PCA was conducted by computing principal components with *smartpca* (v18140) (Patterson, Price, and Reich 2006) using medium and high coverage bison, taurine cattle, and indicine cattle, and then projecting all Beefalo and bison hybrids onto these axes, using our set of pseudohaploid genotypes. This approach allows for directly understanding the ancestry of Beefalo individuals relative to these three groups while mitigating the effects of the low sequencing depth obtained for many Beefalo.

We used ADMIXTURE (v1.3.0) (Alexander, Novembre, and Lange 2009) to further model Beefalo ancestry. ADMIXTURE was run both in supervised and unsupervised modes, using the pseudohaploid dataset filtered for missingness (--geno 0.5) and minor allele frequency (--maf 0.05), and pruned for linkage disequilibrium (--indep-pairwise 50 5 0.5) using plink (v1.9) (Chang et al. 2015). Cross-validation was used to select the optimal value of K. For the supervised ADMIXTURE analysis, Beefalo and bison hybrid ancestry was fit as potentially coming from bison, taurine, or indicine sources.

We also sought to investigate Beefalo ancestry on the sex chromosomes. For the X chromosome, we used males from the overall set of samples and ran ADMIXTURE in haploid mode. We estimated the Y chromosomal phylogeny using males in the sample set with several additional species added as outgroups, including yaks, European bison, gaur, and banteng. The phylogeny was estimated with IQ-TREE (1.6.12) (Nguyen et al. 2015), using the GTR+ASC model and obtaining branch support values using ultrafast bootstrapping (Hoang et al. 2018).

After exploring overall autosomal Beefalo ancestry using model-free (PCA) and model-based (ADMIXTURE) approaches, we sought to specifically test for the presence of bison ancestry in Beefalo and bison hybrid genomes. The D-statistic *D(Beefalo, taurus; bison, water buffalo)* was used for this, which tests for excess allele sharing between Beefalo and bison, relative to taurine cattle, with water buffalo as an outgroup.

We had observed that Beefalo formed a cline between taurine and indicine cattle in PCA, so we also used the statistic *D(Beefalo, taurus; indicus, water buffalo)* to test for indicine ancestry among individual Beefalo. Finally, we quantified the proportion of bison ancestry in Beefalo genomes using the f4-ratio *f4(yak, water buffalo; Beefalo, taurine)/f4(yak, water buffalo; bison, taurine)*, which estimates the proportion of ancestry, α, in Beefalo that comes from a divergent lineage related to yak (the closest outgroup to bison). D- and *f*-statistics were calculated using ADMIXTOOLS2 (Maier et al. 2023) for each individual Beefalo and bison hybrid, grouping bison and taurine and indicine cattle.

Locating segments of differential ancestry within Beefalo

Diploid ancestry state across individual Beefalo genomes was modeled as coming from bison and cattle-related sources by using AncestryHMM (v1.0.2) (Corbett-Detig and Nielsen 2017). We used medium-to-high coverage genomes (>10x) from bison (n=20), taurine cattle (n=20), and zebu (n=6) each as source panels. We called genotypes for each species separately using the same filtering procedure as with our outgroup-ascertained variants. We then merged these panels, requiring sites to be called in at least 6 bison, 6 taurine cattle, and 3 zebu, have a frequency difference of at least 90% between the two source panels, and be at least 500bp from the nearest variant. This gave a final set of 2.32M ancestry informative markers. We modeled local ancestry across the genome for each Beefalo using AncestryHMM with the parameters '-e 0.02 -a 3 0.02 0.90 0.08 -p 0 -5 0.02 -p 1 10000 0.98 -p 2 -20 0.08'. This models two gene flow events, one 20 generations ago from indicine cattle at 8% and the second from bison 5 generations ago at 2%, though allows the exact timing of these admixture pulses to be inferred. We used read counts at ancestry informative markers for all Beefalo as input for ancestry inference. For seven individuals which were sequenced to high coverage, we performed additional ancestry inference using both read data downsampled to ~2x coverage and genotypes at ancestry informative positions. Genotypes at ancestry informative markers were called using GATK HaplotypeCaller. We filtered posterior ancestry probabilities using a 0.9 threshold.

Tables

Supplementary Table 1: List of samples used
Table_S1_Beefalo_sample_info

Acknowledgements

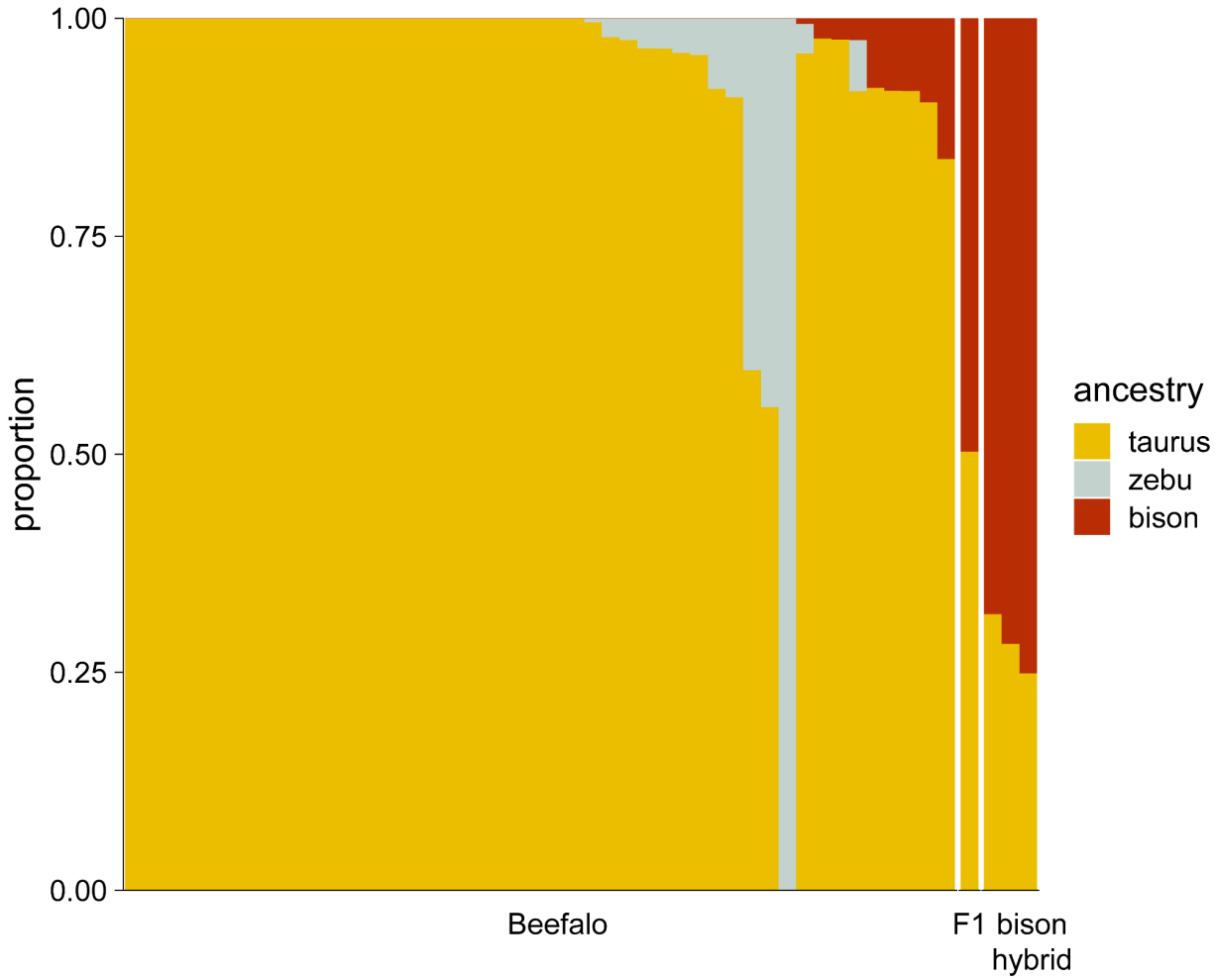
We thank USMARC technicians H.R. Sadd and N.J. Allison for assistance with laboratory procedures. Mention of trade names or commercial products in this

publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer.

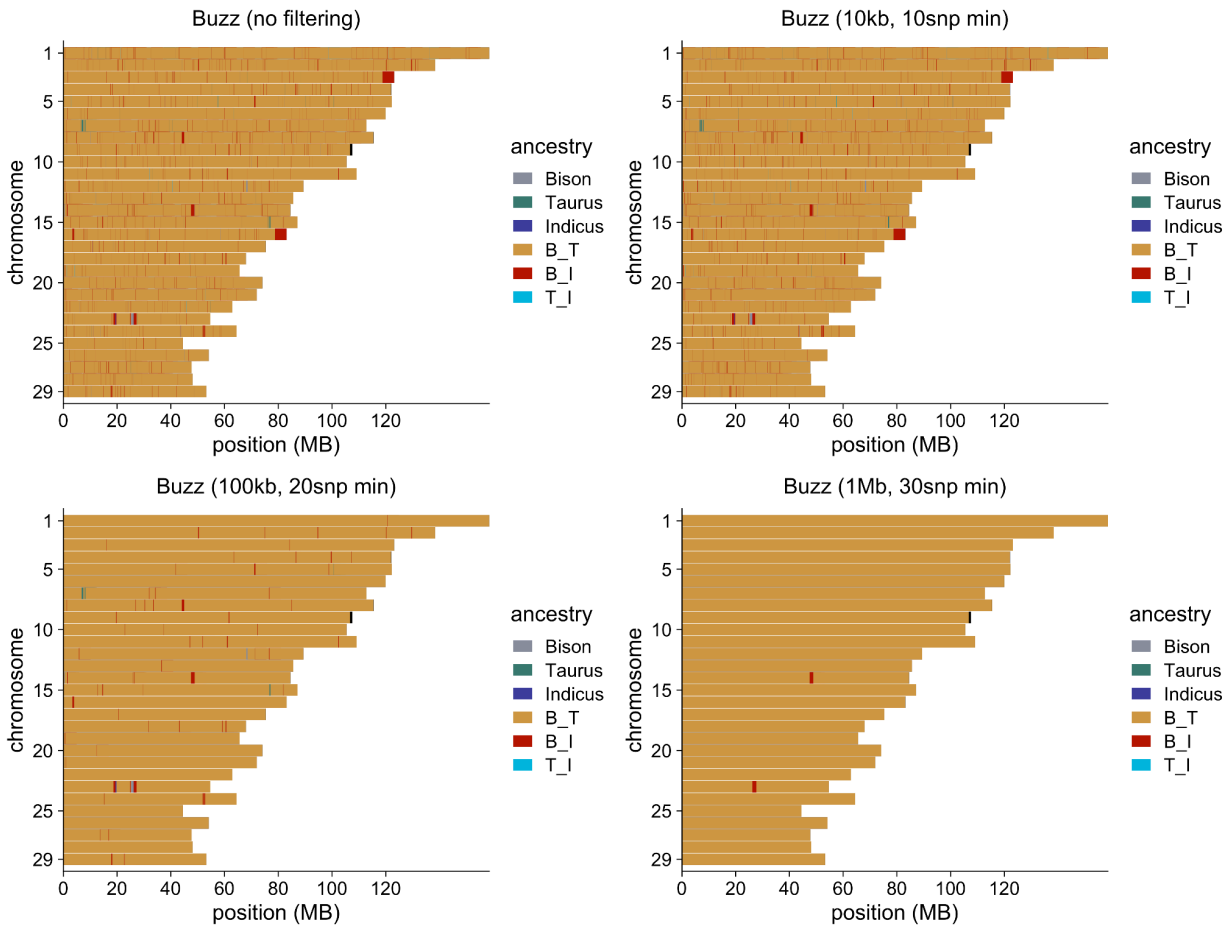
Data Availability

Beefalo, bison hybrid, and bison sequencing data generated in this study are available in NCBI BioProject PRJNA1152308.

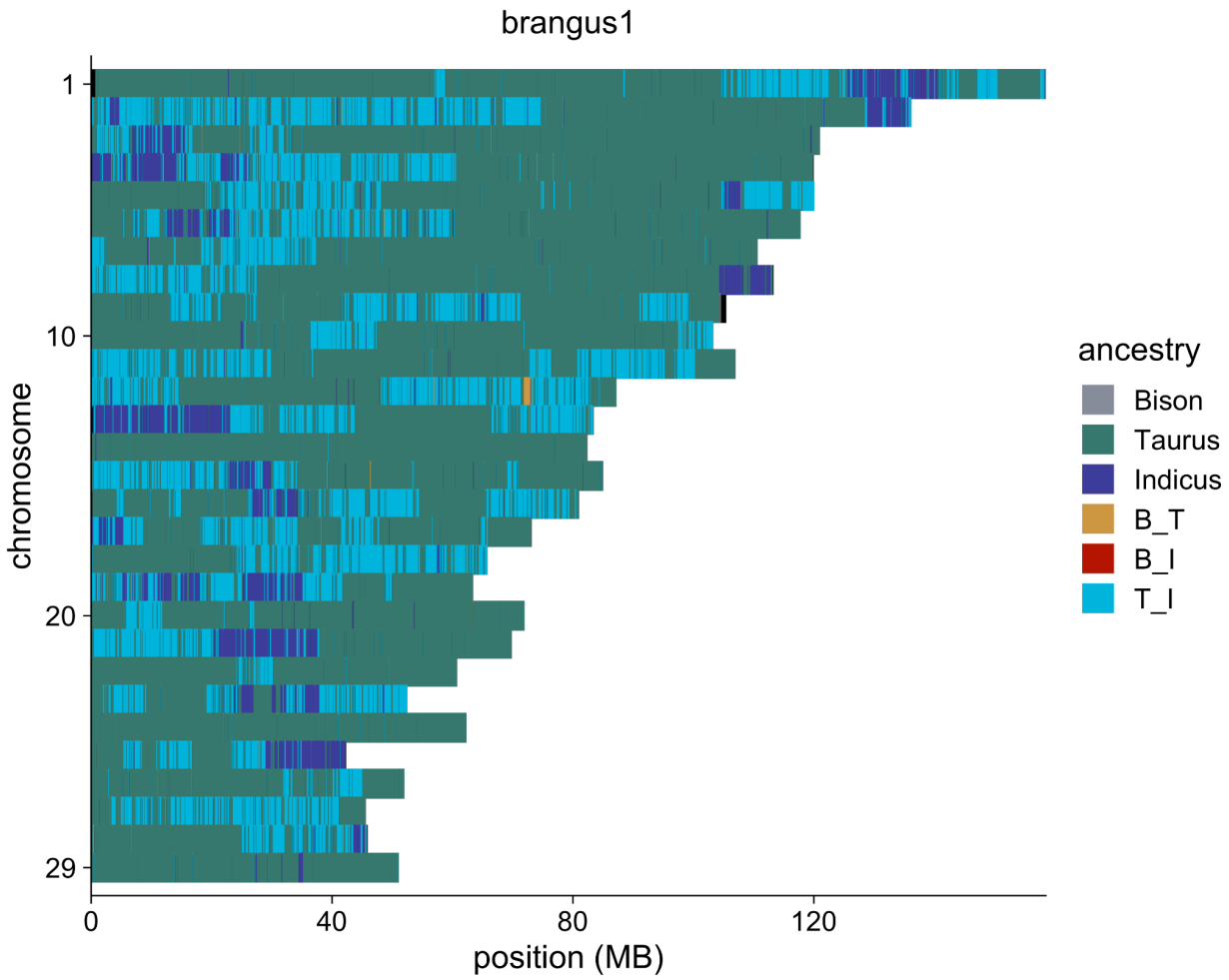
Supplementary Figures



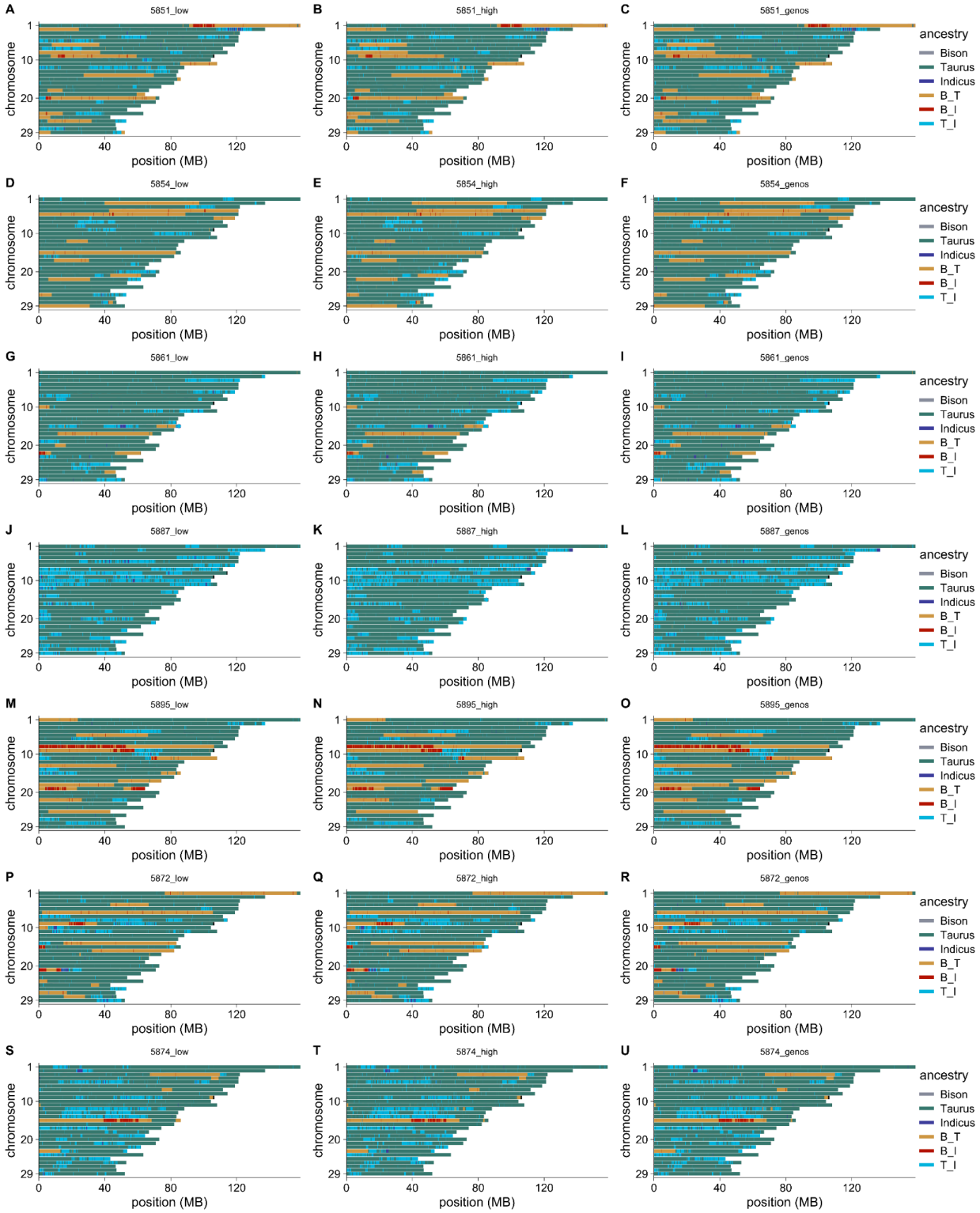
Supplementary Figure 1. *Supervised ADMIXTURE modeling of Beefalo and bison hybrid ancestry, using panels of bison and taurine and zebu cattle.*



Supplementary Figure 2. *Local ancestry inference results from a F1 bison-taurine cattle hybrid. Ancestry was correctly inferred to be almost completely heterozygous for bison and taurine ancestry across the genome, though with some small segments incorrectly assigned to be heterozygous for bison-zebu bison, likely because of the low divergence between cattle subspecies. These segments are all small and are increasingly removed with more stringent filtering (for window size and minimum SNPs to call a window, shown in panel titles).*



Supplementary Figure 3. *Local ancestry inference results from a Brangus individual, showing the presence of both taurine and zebu (indicine) ancestry.*



Supplementary Figure 4. *Comparison of local ancestry inference using either downsampled (~2x) read data (A, D, J, M, P, S), high coverage (30-42x) read data (B, E, H, K, N, Q, T), or called genotypes (C, F, L, O, R, U) on seven individual Beefalo and*

bison hybrids which were sequenced to high coverage: 5851 (A-C), 5854 (D-F), 5861 (G-I), 5887 (J-L), 5895 (M-O), 5872 (P-R), 5874 (S-U).

References

Alexander, David H., John Novembre, and Kenneth Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19 (9): 1655–64.

Anstey, T.H. 1986. "One hundred harvests: Research Branch, Agriculture Canada, 1886-1986." Historical Series No. 27. Research Branch Agriculture Canada, Ottawa, Ontario.

"American Beefalo Association." n.d. American Beefalo Association. Accessed June 1, 2024. https://americanbeefaloassociation.com/.

Barlow, Axel, Stefanie Hartmann, Javier Gonzalez, Michael Hofreiter, and Johanna L. A. Paijmans. 2020. "Consensify: A Method for Generating Pseudohaploid Genome Sequences from Palaeogenomic Datasets with Reduced Error Rates." *Genes* 11 (1).

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20.

Boyd, Mossom M. 1908. "A Short Account of an Experiment in Crossing the American Bison with Domestic Cattle." *The Journal of Heredity* 1: 324–31.

Brower, Jennifer. 2008. *Lost Tracks: National Buffalo Park, 1909-1939*. Athabasca University Press.

Chang, Christopher C., Carson C. Chow, Laurent Cam Tellier, Shashaank Vattikuti, Shaun M. Purcell, and James J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4: 7.

Chen, Ningbo, Yudong Cai, Qiuming Chen, Ran Li, Kun Wang, Yongzhen Huang, Songmei Hu, et al. 2018. "Whole-Genome Resequencing Reveals World-Wide Ancestry and Adaptive Introgression Events of Domesticated Cattle in East Asia." *Nature Communications* 9 (1): 2337.

Corbett-Detig, Russell, and Rasmus Nielsen. 2017. "A Hidden Markov Model Approach for Simultaneously Estimating Local Ancestry and Admixture Time Using Next Generation Sequence Data in Samples of Arbitrary Ploidy." *PLoS Genetics* 13 (1): e1006529.

Deakin, A., G. W. Muir, A. G. Smith, and A. S. Lellan. 1942. "Hybridisation of Domestic Cattle and Buffalo (Bison Americanus). Progress Report off the Wainwright Experiment 1935—41." Canada Department of Agriculture.

DePristo, Mark A., Eric Banks, Ryan Poplin, Kiran V. Garimella, Jared R. Maguire, Christopher Hartl, Anthony A. Philippakis, et al. 2011. "A Framework for Variation Discovery and Genotyping Using next-Generation DNA Sequencing Data." *Nature Genetics* 43 (5): 491–98.

Goodnight, Charles. 1914. "My Experience With Bison Hybrids." *The Journal of Heredity* 5 (5): 197–99.

Hansen, P. J. 2004. "Physiological and Cellular Adaptations of Zebu Cattle to Thermal Stress." *Animal Reproduction Science* 82-83 (July):349–60.

Heath, E. 1979. "Relationships of Musculus Rhomboideus, Ligamentum Nuchae and

Vertebrae Thoracis and Lumbales in Bos Indicus." *Acta Anatomica* 105 (1): 56–60.
Heaton, Michael P., Timothy P. L. Smith, Jacky K. Carnahan, Veronica Basnayake, Jiansheng Qiu, Barry Simpson, and Theodore S. Kalbfleisch. 2016. "Using Diverse U.S. Beef Cattle Genomes to Identify Missense Mutations in *EPAS1,* a Gene Associated with Pulmonary Hypertension." *F1000Research* 5 (August):2003.
Hennessy, D. W., P. J. Williamson, and R. E. Darnell. 2000. "Feed Intake and Liveweight Responses to Nitrogen And/or Protein Supplements by Steers of Bos Taurus, Bos Indicus and Bos Taurus × Bos Indicus Breed Types Offered a Low Quality Grass Hay." *The Journal of Agricultural Science* 135 (1): 35–45.
Hoang, Diep Thi, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, and Le Sy Vinh. 2018. "UFBoot2: Improving the Ultrafast Bootstrap Approximation." *Molecular Biology and Evolution* 35 (2): 518–22.
Hunter, R. A., and B. D. Siebert. 1985. "Utilization of Low-Quality Roughage by Bos Taurus and Bos Indicus Cattle. 2. The Effect of Rumen-Degradable Nitrogen and Sulphur on Voluntary Food Intake and Rumen Characteristics." *The British Journal of Nutrition* 53 (3): 649–56.
Kim, Kwondo, Taehyung Kwon, Tadelle Dessie, Dongahn Yoo, Okeyo Ally Mwai, Jisung Jang, Samsun Sung, et al. 2020. "The Mosaic Genome of Indigenous African Cattle as a Unique Genetic Resource for African Pastoralism." *Nature Genetics* 52 (10): 1099–1110.
Kumar, Anil, Syma Ashraf Waiz, T. Sridhar Goud, R. K. Tonk, Anita Grewal, S. V. Singh, B. R. Yadav, and R. C. Upadhyay. 2016. "Assessment of Adaptability of Zebu Cattle (Bos Indicus) Breeds in Two Different Climatic Conditions: Using Cytogenetic Techniques on Genome Integrity." *International Journal of Biometeorology* 60 (6): 873–82.
Kwon, Taehyung, Kwondo Kim, Kelsey Caetano-Anolles, Samsun Sung, Seoae Cho, Choongwon Jeong, Olivier Hanotte, and Heebal Kim. 2022. "Mitonuclear Incompatibility as a Hidden Driver behind the Genome Ancestry of African Admixed Cattle." *BMC Biology* 20 (1): 20.
Lenoir, F., and M. J. Lichtenberger. 1978. "The Y Chromosome of the Basolo Hybrid Beefalo Is a Y of Bos Taurus." *The Veterinary Record* 102 (19): 422–23.
Li, Heng. 2013. "Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM." *arXiv [q-bio.GN]*. arXiv. http://arxiv.org/abs/1303.3997.
Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows–Wheeler Transform." *Bioinformatics* 25 (14): 1754–60.
Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79.
Maier, Robert, Pavel Flegontov, Olga Flegontova, Ulas Isildak, Piya Changmai, and David Reich. 2023. "On the Limits of Fitting Complex Models of Population History to F-Statistics." *eLife* 12 (April). https://doi.org/10.7554/eLife.85492.
McTavish, Emily Jane, Jared E. Decker, Robert D. Schnabel, Jeremy F. Taylor, and David M. Hillis. 2013. "New World Cattle Show Ancestry from Multiple Independent Domestication Events." *Proceedings of the National Academy of Sciences of the United States of America* 110 (15): E1398–1406.

Miller, Bryan. 1982. “Beefalo: There Are Few On The Range.” *The New York Times*, January 6, 1982. https://www.nytimes.com/1982/01/06/garden/beefalo-there-are-few-on-the-range.html.

“Most Expensive Cattle.” n.d. Guinness World Records. Accessed June 1, 2024. https://www.guinnessworldrecords.com/world-records/70817-most-expensive-cattle.

Murdoch, Brenda M. 2018. “34. Meiotic Recombination and Chromosomal Defects in Hybrid Beefalo and Ruminant Livestock Spermatocytes.” *Cancer Genetics* 224 (August):63–64.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. 2015. “IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.” *Molecular Biology and Evolution* 32 (1): 268–74.

Oppenheimer, Jonas, Benjamin D. Rosen, Michael P. Heaton, Brian L. Vander Ley, Wade R. Shafer, Fred T. Schuetze, Brad Stroud, et al. 2021. “A Reference Genome Assembly of American Bison, Bison Bison Bison.” *The Journal of Heredity* 112 (2): 174–83.

Paim, T. do P., E. H. A. Hay, C. Wilson, M. G. Thomas, L. A. Kuehn, S. R. Paiva, C. McManus, and H. D. Blackburn. 2020. “Dynamics of Genomic Architecture during Composite Breed Development in Cattle.” *Animal Genetics* 51 (2): 224–34.

Patterson, Nick, Alkes L. Price, and David Reich. 2006. “Population Structure and Eigenanalysis.” *PLoS Genetics* 2 (12): e190.

Qiu, Qiang, Lizhong Wang, Kun Wang, Yongzhi Yang, Tao Ma, Zefu Wang, Xiao Zhang, et al. 2015. “Yak Whole-Genome Resequencing Reveals Domestication Signatures and Prehistoric Population Expansions.” *Nature Communications* 6 (December):10283.

Qiu, Qiang, Guojie Zhang, Tao Ma, Wubin Qian, Junyi Wang, Zhiqiang Ye, Changchang Cao, et al. 2012. “The Yak Genome and Adaptation to Life at High Altitude.” *Nature Genetics* 44 (8): 946–49.

Shirazi, Sabrina, Nasreen Broomandkhoshbacht, Jonas Oppenheimer, Jessica Z. Metcalfe, Rob Found, John W. Ives, and Beth Shapiro. 2022. “Ancient DNA-Based Sex Determination of Bison Hide Moccasins Indicates Promontory Cave Occupants Selected Female Hides for Footwear.” *Journal of Archaeological Science* 137 (January):105533.

Soubrier, Julien, Graham Gower, Kefei Chen, Stephen M. Richards, Bastien Llamas, Kieren J. Mitchell, Simon Y. W. Ho, et al. 2016. “Early Cave Art and Ancient DNA Record the Origin of European Bison.” *Nature Communications* 7 (October):13158.

Stormont, C. J., B. G. Morris, Y. Suzuki, and J. Dodd. 1986. “Blood Typing Beefalo Cattle,” 3rd World Congress on Genetics Applied to Livestock Production.

Stroupe, Sam, David Forgacs, Andrew Harris, James N. Derr, and Brian W. Davis. 2022. “Genomic Evaluation of Hybridization in Historic and Modern North American Bison (*Bison bison*).” *Scientific Reports* 12 (1): 6397.

Sun, Ting, Jiafei Shen, Alessandro Achilli, Ningbo Chen, Qiuming Chen, Ruihua Dang, Zhuqing Zheng, et al. 2020. “Genomic Analyses Reveal Distinct Genetic Architectures and Selective Pressures in Buffaloes.” *GigaScience* 9 (2).

https://doi.org/10.1093/gigascience/giz166.
Time. n.d. “Business: Have a Slice of Roast Beefalo.” July 9, 1973. https://time.com/archive/6841372/business-have-a-slice-of-roast-beefalo/.
Vajana, Elia, Mario Barbato, Licia Colli, Marco Milanesi, Estelle Rochat, Enrico Fabrizi, Christopher Mukasa, et al. 2018. “Combining Landscape Genomics and Ecological Modelling to Investigate Local Adaptation of Indigenous Ugandan Cattle to East Coast Fever.” *Frontiers in Genetics* 9 (October):385.
Verdugo, Marta Pereira, Victoria E. Mullin, Amelie Scheu, Valeria Mattiangeli, Kevin G. Daly, Pierpaolo Maisano Delser, Andrew J. Hare, et al. 2019. “Ancient Cattle Genomics, Origins, and Rapid Turnover in the Fertile Crescent.” *Science* 365 (6449): 173–76.
Wang, Kun, Johannes A. Lenstra, Liang Liu, Quanjun Hu, Tao Ma, Qiang Qiu, and Jianquan Liu. 2018. “Incomplete Lineage Sorting rather than Hybridization Explains the Inconsistent Phylogeny of the Wisent.” *Communications Biology* 1 (October):169.
Ward, James A., Gillian P. McHugo, Michael J. Dover, Thomas J. Hall, Said Ismael Ng’ang'a, Tad S. Sonstegard, Daniel G. Bradley, Laurent A. F. Frantz, Michael Salter-Townshend, and David E. MacHugh. 2022. “Genome-Wide Local Ancestry and Evidence for Mitonuclear Coadaptation in African Hybrid Cattle Populations.” *iScience* 25 (7): 104672.
Wecek, Karolina, Stefanie Hartmann, Johanna L. A. Paijmans, Ulrike Taron, Georgios Xenikoudakis, James A. Cahill, Peter D. Heintzman, et al. 2017. “Complex Admixture Preceded and Followed the Extinction of Wisent in the Wild.” *Molecular Biology and Evolution* 34 (3): 598–612.
Wu, Dong-Dong, Xiang-Dong Ding, Sheng Wang, Jan M. Wójcik, Yi Zhang, Małgorzata Tokarska, Yan Li, et al. 2018. “Pervasive Introgression Facilitated Domestication and Adaptation in the Bos Species Complex.” *Nature Ecology & Evolution* 2 (7): 1139–45.
Yang, Tianfu, Michelle Miller, David Forgacs, James Derr, and Paul Stothard. 2020. “Development of SNP-Based Genomic Tools for the Canadian Bison Industry: Parentage Verification and Subspecies Composition.” *Frontiers in Genetics* 11 (November):585999.
Zeder, Melinda A. 2006. *Documenting Domestication: New Genetic and Archaeological Paradigms*. University of California Press.
Zhang, K., J. A. Lenstra, S. Zhang, W. Liu, and J. Liu. 2020. “Evolution and Domestication of the Bovini Species.” *Animal Genetics* 51 (5): 637–57.