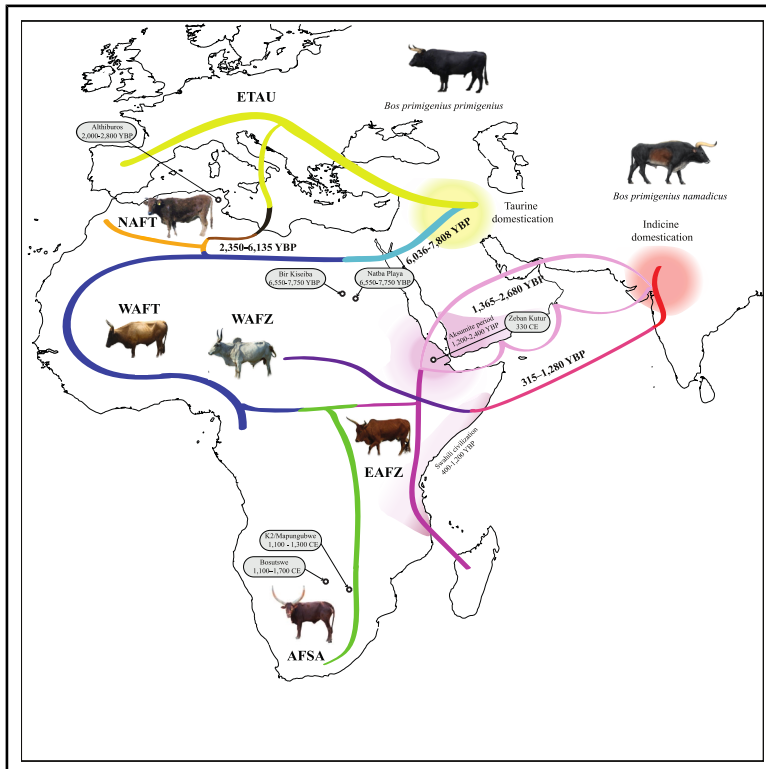


Unscrambling the history of African indicine cattle genomes

Graphical abstract



Authors

Mulusew Kassa Bitew,
Helina Solomon Woldekiros,
Christian Persichilli, ..., Olivier Hanotte,
Fabio Pilla, Gabriele Senczuk

Correspondence

o.hanotte@cgiar.org (O.H.),
pilla@unimol.it (F.P.),
g.senczuk@unimol.it (G.S.)

In brief

Zoology; paleobiology; paleogenetics

Highlights

- Demographic history of African cattle genetic diversity using ABC
- Two waves of the entrance of indicine cattle in Africa
- Integration of genomic data with archaeological and historical records



Article

Unscrambling the history of African indicine cattle genomes

Mulusew Kassa Bitew,^{1,2} Helina Solomon Woldekiros,³ Christian Persichilli,¹ Marika Di Civita,¹ Slim Ben Jemaa,^{4,5} Salvatore Mastrangelo,⁵ Joram M. Mwacharo,^{6,7} Olivier Hanotte,^{8,9,*} Fabio Pilla,^{1,*} and Gabriele Senczuk^{1,10,*}

¹Department of Agriculture, Environmental and Food Sciences, University of Molise, Campobasso, Italy

²Department of Plant Science, College of Agriculture and Natural Resource, Debre Markos University, Debre Markos, Ethiopia

³Department of Anthropology, Washington University in St. Louis, St. Louis, MO, USA

⁴Laboratoire des Productions Animales et Fourragères, Institut National de la Recherche Agronomique de Tunisie, Université de Carthage, Ariana 2049, Tunisia

⁵Dipartimento Scienze Agrarie, Alimentari e Forestali, University of Palermo, Palermo 90128, Italy

⁶Dryland Livestock Genomics, International Centre for Agricultural Research in the Dry Areas (ICARDA), Addis Ababa, Ethiopia

⁷Animal and Veterinary Sciences, SRUC and Centre for Tropical Livestock Genetics and Health (CTLGH), Roslin Institute Building, Midlothian, Scotland

⁸School of Life Sciences, University of Nottingham, Nottingham, UK

⁹CTLGH - LiveGene, International Livestock Research Institute, Addis Ababa, Ethiopia

¹⁰Lead contact

*Correspondence: o.hanotte@cgiar.org (O.H.), pilla@unimol.it (F.P.), g.senczuk@unimol.it (G.S.)

<https://doi.org/10.1016/j.isci.2026.115295>

SUMMARY

African cattle's genetic heritage reflects millennia of relationships between African and Asian human societies and civilizations. By analyzing genome-wide SNP data from 1,043 individuals representing 38 breeds, including 36 from Africa, in an approximate Bayesian computation framework, coupled with random forest classification, we show that this history is characterized by at least two independent arrivals of zebu (indicine) cattle from the Indian subcontinent. Time estimates coincide with major cultural transitions across Africa, including state formation, iron technology spread, and pastoral intensification, exemplified by developments such as the Aksumite Empire and Swahili civilization. Moreover, we show that sanga cattle are the result of admixture between native African taurine and zebu cattle from the first arrival. These findings provide a refined chronology of cattle movements and pastoralism into Africa, reveal how human cultural transitions structured livestock genomes, and highlight Africa's central role in shaping global cattle diversity.

INTRODUCTION

Following the domestication of cattle from aurochs (*Bos primigenius primigenius*) around 10,000 years ago in the Fertile Crescent and from Asian aurochs (*B. p. namadicus*) around 8,000 years ago in the Indus Valley, two lineages emerged—humpless taurine (*Bos taurus taurus*) and humped indicine or zebu (*B. t. indicus*) cattle.

The earliest cattle in Africa were most likely humpless (*B. t. taurus*). Among the oldest lines of evidence of domestic cattle on the continent are zooarchaeological findings from the Letti Desert 2 site in North Sudan (mid-8th millennium BCE/~10,000 years before present [YBP]),¹ skeletal remains dated to 7,750–6,550 YBP from Nabta Playa and Bir Kiseiba in Southern Egypt, and evidence from the Kerma civilization, Sudan (4,500–3,500 YBP).^{2–5} These animals entered Africa through Egypt and the Sinai Peninsula as fully domesticated stock and subsequently dispersed through the Sudano-Sahel zone and along the Mediterranean coast.^{6,7} Genomic data support this ancient presence, and, in fact, a distinctive taurine genetic lineage is currently present in North and West Africa.^{8,9} While West African taurine prob-

ably remained isolated to tsetse-infested sub-humid and humid agro-ecologies in West Africa,¹⁰ North African cattle underwent a more complex interaction with both European taurine and local auroch populations.¹¹

The timing and history of indicine cattle arrival on the continent remain disputed. Genetic data support the Horn of Africa as their main entry point.⁶ Today, all African cattle carry taurine mitochondrial DNA, suggesting a male-mediated indicine cattle introgression.^{12,13} Both taurine and indicine Y chromosome haplotypes occur on the continent, with two phylogeographic lineages for the latter.^{14,15} Also, two distinct patterns for the spread of indicine ancestry across Africa have been suggested, a slow diffusion of indicine genome background and a more rapid one.⁶ Cattle depictions on Egyptian tombs and pottery do not allow an unequivocal interpretation of the presence of zebu,¹⁶ suggesting that, if present, they were probably not in large numbers and, possibly, limited to Northeast Africa.^{17,18} The increased occurrence of humped cattle might have coincided with the onset of the Swahili civilization and associated Indian Ocean maritime trading network around the mid-1st millennium AD,^{19,20} which led to the movement of other animals of



Asian origin into Africa.^{21,22} Humped cattle might also have reached Africa following a terrestrial route traveling across the Southern Arabian Peninsula.¹⁹ The pre-Aksumite D'MT kingdom (c. 1,000–400 BCE) and later Aksumite Empire (c. 100–9CE 60) maintained extensive trade connections with South Arabia, as evidenced by Sabaeen script usage, monumental architecture, and cultural exchanges that could have facilitated livestock introductions alongside other traded goods.^{23,24} These Red Sea trade networks linking the Horn of Africa with the Arabian Peninsula provided established routes for the movement of cattle and other domesticated animals between regions starting from the pre-Aksumite kingdom.^{25,26} Indeed, a large proportion of modern Omani cattle is indicine.²⁷ Still, the dating of indicine genetics in Southern Arabia is unknown. It might have been relatively early, considering its geographic proximity to the center(s) of indicine domestication in the North of the Indian subcontinent. Once in Africa, indicine cattle crossed with local taurine cattle. The sanga cattle of Central and Southern Africa are the legacy of these early hybridization events, with the Bantu-speaking communities thought to be responsible, at least in part, for their dispersal and divergence into different ecotypes.^{28–31} The cradle and timing of the initial dispersion of sanga cattle remain unknown.

Despite increasing efforts to reconstruct the demographic history of African cattle using genomic tools, a model-based calibration of indicine cattle entry and admixing with native African taurine cattle on the continent remains elusive. Previous studies, while insightful, often relied on analytical models, tools, and data with limited power to resolve the frequency, timing, and demographic impact of indicine introductions. Furthermore, competing hypotheses postulating multiple indicine entry routes are yet to be tested using genome-wide SNP data within a robust statistical framework. The combination of machine learning and approximate Bayesian computation (ABC) algorithms returns robust predictions, even under computing time constraints.^{32–34} This approach has been used to unveil or confirm complex evolutionary models in plants and animals,^{35,36} including livestock³⁷ and humans.^{38,39}

Here, we used genome-wide SNP data analyzed in an ABC framework to model and test competing hypotheses (single vs. multiple introductions) relating to the entry, dispersal, and admixing of indicine African cattle on the continent. We used the best-predicted scenario to estimate effective population sizes, divergence, and admixture times, giving further insights into the complex history of African cattle pastoralism.

RESULTS AND DISCUSSION

Current diversity and genetic background for demographic inference

We uncovered substantial genetic variation within and between the breeds, as indicated by *Ho*, *He*, and *Ne* (Table 1). This may reflect several factors that have sculptured African cattle genomes,^{40–42} such as diverse historical breeding practices in heterogeneous environments, spatiotemporal demographic events, and admixture.

The first two axes of the principal-component analysis (PCA) identified several genetic clusters including West African taurine, Northern Africa taurine, and African zebu, with sanga cattle at an intermediate position (Figures 1A and S1).

Admixture at $K = 2$ separated taurine (African and European) breeds from Asian (GIR) and African zebu breeds, delineating the two domestic cattle lineages. It also showed a diverse gradient of taurine and indicine ancestries in African-admixed cattle, suggesting a complex history of hybridizations. At $K = 3$, a distinct genome background is present in West African taurine (LAG, MT, N'Dama (NDAM), Baoule [BAO], and Somba [SOM]), as reported by previous studies as an example.^{43,44} At $K = 6$, the main African cattle population groups (North African taurine [NAFT], West African taurine [WAFT], West African zebu [WAFZ], East African zebu [EAFZ], and African sanga [AFSA]) are evident (Figure 1B). African taurine cattle are separated into NAFT and WAFT, while African zebu are separated from Asian indicine. Low admixture was also observed between EAFZ and WAFZ (Figures 1B and S2A), indicating minimal/restricted interactions.

The lowest cross-validation error, which reflects the optimal number of genetic clusters in the global dataset, was at $K = 22$ (Figure S2B). This number is lower than 36 for African cattle breeds analyzed, suggesting that a mosaic of admixed genomes likely formed following distinct historical events, socio-cultural and husbandry practices, and environmental constraints.

Maximum likelihood tree on the entire dataset shows that while African zebu are separated from Asian indicine, two migration events can be observed from the latter in the direction of WAFZ (Figure 2A). In addition, considering the reduced dataset, two migration events are highlighted, the first from Asian indicine to EAFZ and a second from European taurine to NAFT (Figure 2B). The outlined historical gene exchanges conform with the significant footprint that non-African cattle have left in generating a complex genomic architecture.^{12,45}

These preliminary analyses lay a solid foundation to address some unresolved questions. How and when did indicine cattle enter Africa? Did it happen in a single wave or in two waves? To disentangle these queries, nine different demographic histories were tested in a coalescent-based Approximate Bayesian Computation using Random Forest (ABC-RF) analysis, as discussed below.

Early African taurine cattle

Using a calibration point that constrained prior information concerning the taurine/indicine admixture time based on whole-genome sequencing (WGS) estimates, the divergence between these two lineages was placed at about 150,726 generations ago (ca. 750,000 years ago). This scenario (Figure 3) augurs well with previous time estimates based on complete mitogenomes.^{46–48}

The modeled early introduction of African taurine is corroborated by osteological evidence from North Sudan and Egypt dating to approximately 10,000–7,750 YBP,^{1–4} with more conclusive evidence of established pastoral societies by the middle Neolithic in Sudan.⁴⁹ Direct dating on bones of sheep/goat and cattle from the Central Saharan site of Takarkori, Libya, provides evidence for the punctuated integration of small groups

Table 1. Sample information and summary statistics

Breed name	Group	Code	Sample	Ho	He	Ne ₁₃	Origin
Afrikaner	AFSA	AFR	4	0.217	0.192	14	South Africa
Adamawa Gudali	WAFZ	AG	24	0.259	0.251	110	Nigeria
Ankole-Watusi	AFSA	ANKW	5	0.251	0.215	17	Uganda
Ankole	AFSA	AO	25	0.262	0.258	138	Uganda
Baladi	–	BAL	30	0.297	0.301	146	Egypt
Baoule	WAFZ	BAO	20	0.230	0.220	90	Burkina Faso
Biskra	–	BIS	30	0.340	0.342	151	Algeria
Bunaji	WAFZ	BJ	22	0.253	0.255	130	Nigeria
Blonde du Cap Bon	NAFT	BLCAP	19	0.316	0.305	87	Tunisia
Boran	–	BOR1	20	0.278	0.273	100	Ethiopia
Boran	–	BOR2	20	0.245	0.236	69	Kenya
Borgou	–	BORG	158	0.278	0.279	386	Benin
Brune de l'Atlas Fauve	NAFT	BRATF	9	0.319	0.308	44	Tunisia
Brune de l'Atlas Grise	NAFT	BRATG	20	0.316	0.314	67	Tunisia
Cheurfa	NAFT	CHE	31	0.318	0.319	178	Algeria
Chélifienne	–	CHF	30	0.333	0.336	172	Algeria
East African Shorthorn Zebu	EAFZ	EASZ	112	0.247	0.254	698	Kenya
Gir	IND	GIR	28	0.171	0.169	103	India
Guelmoise	NAFT	GUEL	24	0.300	0.313	136	Algeria
Holstein	ETAU	HOL	59	0.335	0.328	92	Europe
Karamojong Zebu	EAFZ	KR	16	0.250	0.243	85	Uganda
Kuri	–	KUR	20	0.276	0.270	105	Chad
Lagune	–	LAG	20	0.196	0.192	82	Benin
Muturu	–	MT	8	0.204	0.188	35	Nigeria
N'Dama	WAFZ	NDAM	51	0.243	0.244	170	Côte d'Ivoire
Nganda	–	NG	23	0.279	0.268	74	Uganda
Red Bororo	WAFZ	OR	21	0.251	0.251	122	Nigeria
Oulmès Zaer	–	OUL	19	0.305	0.290	55	Morocco
Sheko	AFSA	SHK	17	0.266	0.259	93	Ethiopia
Sokoto Gudali	WAFZ	SO	19	0.259	0.253	104	Nigeria
Somba	WAFZ	SOM	20	0.229	0.235	105	Togo
Tidili	NAFT	TID	31	0.298	0.299	160	Morocco
Tuli	AFSA	TULI	4	0.302	0.260	15	Botswana
Yakanaji	WAFZ	YK	12	0.259	0.242	49	Nigeria
Zebu Bororo	WAFZ	ZBO	20	0.255	0.255	114	Chad
Zebu Fulani	WAFZ	ZFU	20	0.257	0.249	96	Benin
Madagascar Zebu	EAFZ	ZMA	20	0.207	0.205	105	Madagascar
Serere Zebu	EAFZ	ZS	12	0.253	0.237	51	Uganda

List of African cattle breeds used in this study with their codes, group for the ABC-RF analysis, sample sizes, observed and expected heterozygosity, effective population size at the earliest 13 generations, and geographic origin.

of herders who crossed Africa's northeastern quadrant around 8,300 years cal BP, suggesting that pastoralism emergence followed a punctuated, rather than gradual, dispersal model.⁵⁰ This punctuated dispersal pattern overcame epidemiological barriers, including trypanosomiasis and other vector-borne diseases endemic to regions with wild ungulate reservoirs, and was facilitated by the pre-existing organizational infrastructure of complex foraging societies (characterized by delayed-return economies involving food storage and seasonal resource sched-

uling) that enabled selective integration of domestic animals.^{18,50} The fragmentary and taphonomically compromised nature of faunal assemblages from early African pastoral sites precludes robust morphometric analyses necessary to differentiate autochthonous domestication events from the introduction of livestock already domesticated elsewhere.¹ The ancestral taurine cattle subsequently dispersed alongside the Neolithization process following different routes, southwards across the Nile Valley corridor and westwards along the Mediterranean

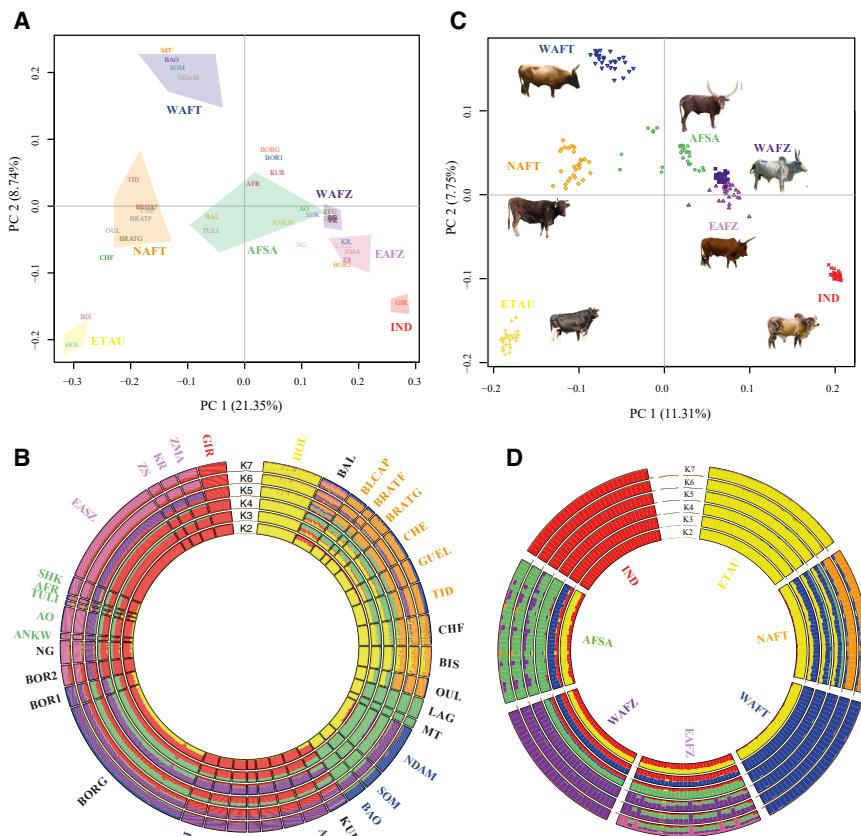


Figure 1. PCA and population structure of African cattle

(A) PCA showing the genetic relations of 38 African indigenous cattle breeds using 34,699 SNPs.

(B) Circular bar plot representation of ADMIXTURE ancestry ($K = 2-7$) of 38 cattle breeds, using 34,699 SNPs.

(C) PCA of the selected population groups used for ABC-RF (13,160 SNPs).

(D) Circular bar plot representing the proportion of ADMIXTURE ancestry ($K = 2-7$) among the population groups used for ABC-RF (13,160 SNPs).

Both PCA and ADMIXTURE showed a clear separation between lineages (taurine, indicine, and sanga at intermediate position) and among the groups of breeds according to their geographical origin: yellow, ETAU; red, IND; orange, NAFT; green, AFSA; dark purple, WAFZ; light purple, EAFZ; blue, WAFT; black, breeds not included in the ABC analysis.

coastline, with more conclusive evidence of established pastoral societies by the middle Neolithic period (5,000–4,000 YBP) in the Kerma region, Sudan.⁵¹

Modern North African cattle are characterized by mixed genetic components, including a marked fraction of European taurine (Figure 1B). Several explanations may be advanced. It could be the retention of ancestry-shared polymorphisms. If this was the case, we should have found European taurine signatures in the ancestral genomic components of other African taurine cattle. This was, however, not observed, making this postulation less plausible. A second explanation could be linked to the 19th-century introductions through European colonization of the Maghreb region.^{52,53} Yet, a recent hybridization event should have left more marked evidence, as revealed in the two Algerian breeds (CHF and BIS) in which considerable Holstein admixture is evident across the K values tested (Figures 1B and S2), underlining their known recent crossbred history.^{52,53} A further hypothesis^{11,45} is the possible role of the African auroch in shaping North African cattle genomes, which cannot be ruled out. Despite the recent increase in the availability of auroch archaeogenomes,⁴⁹ gaps still remain, especially in North Africa and the Mediterranean areas, making it difficult to unveil the extent and direction of aurochs post-domestication introgression. Iron Age genomic data of cattle from Althiburos (Tunisia) dated 2,800–2,000 YBP, including the Epipaleolithic Moroccan aurochs, fall between European and African taurine cattle, sug-

gesting that a degree of gene flow was already present at the time. Moreover, the presence of the mitochondrial R lineage, restricted to a few modern Italian breeds (i.e., Chianina, Marchigiana, and Romagnola), to Iron Age domestic cattle from Althiburos and to the Epipaleolithic aurochs of Taghit Haddouch (Morocco)^{54–56} indicates that the genetic exchange between the Italian Peninsula and North Africa was not uncommon. Early livestock gene flow across the Mediterranean Basin, dating at least from the Bronze Age is supported by the distribution and frequency of the African T1 haplotypes in Southern Europe.^{14,57} This genetic connection is also evident in our estimated effective migration surface (EEMS) method based on a stepping-stone model, which showed corridors of higher migration rates spanning Western Mediterranean Basin to African Western coastal region (Figure 4). Increasing studies on modern and ancient human genomes in North Africa are also highlighting human movements prior to historical Mediterranean conquests and commerce.^{58,59}

According to our demographic reconstruction, a secondary admixture event between African and European taurine (taN) occurred around 4,415 YBP (95% CI: 6,135–2,350 YBP). This interval spans the late Neolithic through the early Iron Age, with the point estimate coinciding with the Bronze Age. Taking into account the large confidence intervals in our estimates, a resolute time window remains uncertain, although several arguments can be advanced. The Bronze Age was a period of significant cultural and economic exchange across the Mediterranean. In particular, the late Bronze Age was a time of pronounced growth and prosperity, culminating with a major collapse during the Iron Age transition.^{60–64} This transition also coincided with the rise of the Phoenician, so that a multi-phase process driven by the Phoenician “globalization” across the Southern Mediterranean basin cannot be ruled out in explaining the genetic proximity between North African and Southern European cattle.^{65,66}

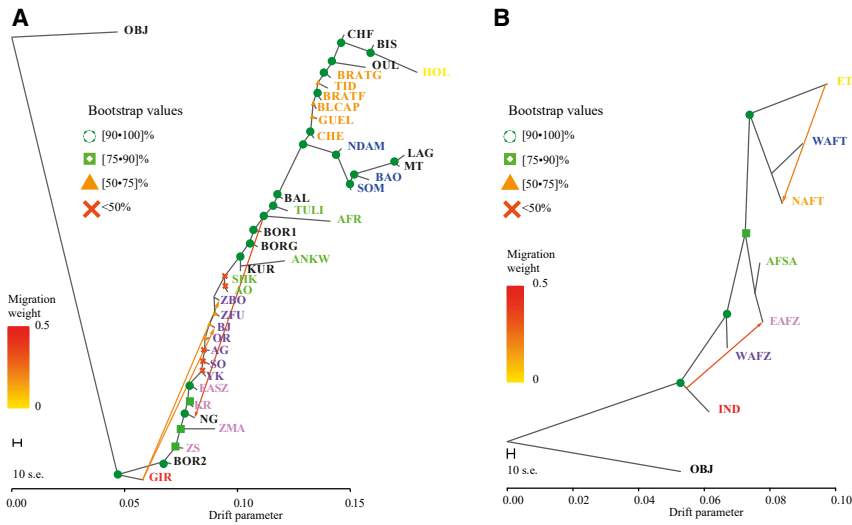


Figure 2. Maximum likelihood trees based on TreeMix

(A) Maximum likelihood tree inferred from 38 African cattle breeds, using 34,699 SNPs.

(B) Maximum likelihood tree from the six cattle metapopulations selected for ABC-RF analysis using 13,160 SNPs.

All migration events are depicted with arrows and colored according to their migration weight. Node robustness was estimated with 1,000 bootstrap replicates and is reported for each node. Migration bars indicate inferred gene flow between populations; darker bars correspond to higher migration weight (greater ancestry contribution). In both maximum likelihood trees, breeds clustered according to lineages (taurine, indicine and sanga at intermediate position) and their geographical origin: yellow, ETAU; red, IND; orange, NAFT; green, AFSA; dark purple, WAFZ; light purple, EAFZ; blue, WAF; black, breeds not included in the ABC analysis.

ETAU, European taurine; IND, Asian indicine; NAFT, North African taurine; AFSA, African sanga cattle; WAFZ, West African zebu; EAFZ, East African zebu; WAF, West African taurine; ABC, approximate Bayesian computation.

ETAU, European taurine; IND, Asian indicine; NAFT, North African taurine; AFSA, African sanga cattle; WAFZ, West African zebu; EAFZ, East African zebu; WAF, West African taurine; ABC, approximate Bayesian computation.

Two waves of indicine cattle introduction into Africa

Based on our random forest classification, the most voted hypothesis is that of at least two mutually exclusive introductions (305 votes; posterior probability = 1.00, with an error rate of 0.015 and concurrent overlap between the observed and simulated datasets, Figure S3A). Genetic evidence for two waves of arrival and dispersion of indicine cattle has been reported previously based on microsatellite data.⁶ Moreover, Pérez-Pardal et al.¹⁵ reported the presence of a Y chromosome haplotype

(Y3_B) in WAFZ cattle and the almost exclusive presence of a different one (Y3_A) in East Africa.

When we tested our nine scenarios separately, the most plausible was scenario 7 (posterior probability = 0.59, with an error rate of 0.002 and high overlap between the observed and simulated datasets; Figures S3B and S4). This scenario predicts a first entrance of early EAFZ to have occurred before the second arrival, which gave rise to the WAFZ. Moreover, AFSA are the outcome of hybridization of ancestral indicine from the first

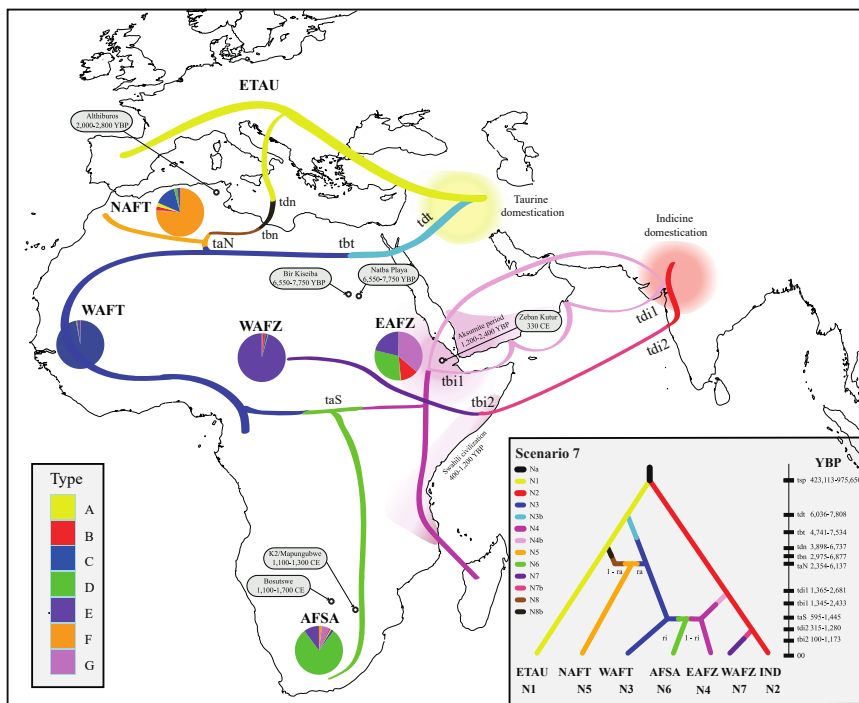


Figure 3. Schematic of the main entrances and subsequent dispersals and admixture of African cattle

Pie chart representing the genomic components at $K = 7$ of the main geographic clusters with relative possible migrations and admixture events according to the most voted scenario (scenario 7) in the ABC-RF analysis. The inferred admixture and diverging times are reported in years before present (YBP) with their relative confidence intervals. The location of key archaeological sites is also reported in the map. Breed groups are colored according to their geographic origin: yellow, ETAU; red, IND; orange, NAFT; green, AFSA; dark purple, WAFZ; light purple, EAFZ; blue, WAF.

ETAU, European taurine; IND, Asian indicine; NAFT, North African taurine; AFSA, African sanga cattle; WAFZ, West African zebu; EAFZ, East African zebu; WAF, West African taurine.

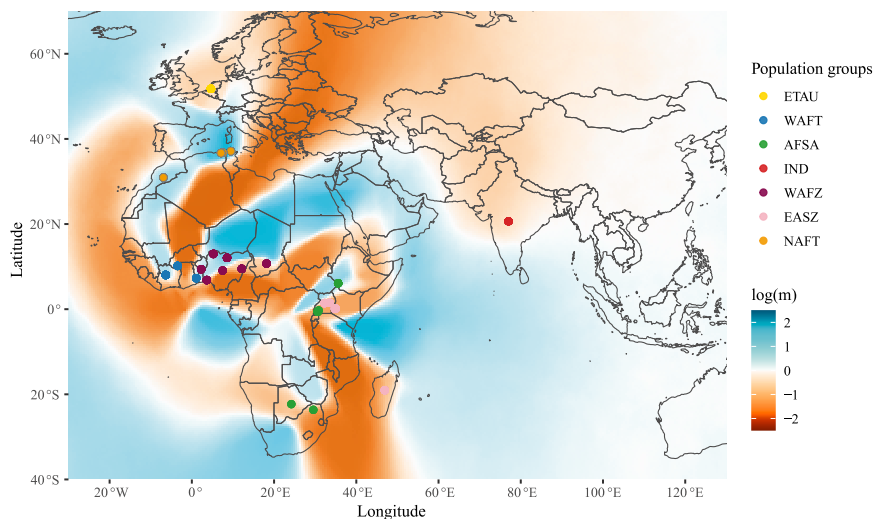


Figure 4. Estimated EEMSs of the main African cattle population groups

The posterior mean of effective migration rate (m) interpolated between grid points. Geographic locations with significantly higher migration rates, which correspond to possible corridors to gene flow, are shown in blue, while geographic areas with significantly lower migration rates, which may act as barriers, are depicted in orange. EEMSs, effective migration surfaces.

arrival with local African taurine. This scenario contrasts with the hypothesis of Pérez-Pardal et al.¹⁵ which suggested that the West African haplotype (Y3_B) represents the oldest indicine lineage introduced on the continent ca. 3,000 YBP, with the East African one (Y3_A) being the consequence of a more recent introduction. This scenario, although tested in our simulations (scenario 6), showed a smaller number of votes than scenario 7 (84 vs. 213 votes, respectively). Regarding possible entry modes, it is difficult to elucidate whether the two waves followed mainly maritime or terrestrial routes; however, the reconstruction of the EEMSs showed several possible corridors encompassing both terrestrial dispersals connecting sub-Saharan Africa regions and Arabian Peninsula and possible maritime and coastal routes from Arabian Peninsula to East African coastal regions (Figures 3 and 4).

Timing of entrances and archaeological evidence

The first entry approximating 2,190 YBP (95% confidence interval: 1,365–2,680 YBP, Table 2) coincides with the pre-Aksumite (1,600–400 BCE) and Aksumite (400 BCE–8CE 00) periods characterized by complex multi-species livestock-keeping strategies and extensive trade networks that facilitated the movement of diverse domestic species.^{25,26,67,68} Archaeological evidence demonstrates that these trade connections, including the critical Aksumite salt trade networks, linked the Horn of Africa with the Arabian Peninsula, Red Sea regions, and beyond, creating pathways for both elite goods and everyday commodities including livestock.²⁶ Zooarchaeological evidence from highland Ethiopian sites demonstrates that cattle were indeed a dominant component of pre-Aksumite and Aksumite pastoralism, with assemblages showing high frequencies of cattle remains alongside sheep and goats, and, notably, some of the earliest lines of evidence for domestic chickens in the Horn of Africa dating to the pre-Aksumite period.^{25,68} This diverse livestock portfolio indicates complex animal management systems that could have readily incorporated new bovine varieties through the established trade connections. However, morphological identification of zebu cattle from osteological findings remains challenging, as

archaeological evidence for zebu presence during this period should be confirmed using ancient DNA as it provides the most unambiguous evidence. The first indicine entrance facilitated the formation of sanga cattle through admixture with local taurine populations. Archaeological evidence from Southern Africa, including sites like K2 and Mapungubwe in South Africa and Bosutswe in Botswana, documents the subsequent expansion and diversification of these admixed sanga populations during the first and early second millennium CE, with increasing herd sizes and the development of complex societies in which livestock wealth became central to social stratification and political power.^{71,72} The introduction of zebu genetics enhanced drought resistance and heat tolerance, facilitating pastoral expansion into more arid environments and contributing to the development of the distinctive sanga cattle breeds adapted to diverse Southern African ecologies.

The second indicine entrance occurred around 770 YBP (95% confidence interval: 315–1,280 YBP; Table 2), which coincides with intensified Indian Ocean trade networks and the expansion of Swahili civilization along the East African coast. This period saw the development of powerful city states and increased connectivity across the western Indian Ocean, building on earlier trade routes that had facilitated the movement of diverse domesticated species, as evidenced by the presence of Asian-origin chickens in the Horn of Africa over a millennium earlier.⁶⁸ Archaeological evidence from Swahili sites indicates dynamic livestock economies that thrived without massive livestock importation, with cattle frequencies relative to caprines varying considerably over time and space according to local environmental conditions and economic strategies.^{73–75} At coastal sites, marine resources remained important alongside livestock, while cattle became increasingly significant at islands with suitable grazing conditions such as Pemba, Tanzania, suggesting that zebu introductions occurred selectively rather than on a massive scale despite increased connectivity with regions where indicine cattle were prevalent.⁷⁴ Archaeological evidence from Madagascar, particularly the coastal trading site of Mahilaka in Northwestern Madagascar (11th–14th century CE), documents the presence

Table 2. Parameter estimation of the most voted scenario

Parameters	Demographic event	Median	Q0.05	Q0.95
Na	ancestral population size	7718.47	3319.32	9763.22
N1	ETAU population size	8168.96	4996.26	9906.20
N2	IND population size	7785.13	4352.32	9819.78
N3	WAFZ population size	7502.80	3940.20	9381.51
N3b	bottleneck event for WAFZ	5582.38	1328.95	9651.54
N4	EAFZ population size	6099.02	1921.88	9476.64
N4b	bottleneck event for EAFZ	7061.54	2051.90	9719.33
N5	NAFT population size	6102.26	2263.61	9211.50
N6	AFSA population size	4274.35	620.893	9595.92
N7	WAFZ population size	2419.62	298.596	8055.79
N7b	bottleneck event for WAFZ	5865.89	613.889	9509.78
N8	ancestral unsampled population of NAFT	6187.35	811.257	9738.45
N8b	bottleneck event for the secondary entrance of ETAU	6118.15	529.778	9368.48
tbi2	time of the bottleneck of the second indicine entrance	109.00	20.00	234.61
tdi2	time of the entrance of the second indicine ancestors	154.00	63.00	256.00
taS	time of admixture giving rise to AFSA	210.00	119.00	289.00
tbi1	time of the bottleneck of the first indicine entrance	374.98	269.00	486.79
tdi1	time of the entrance of the first indicine ancestors	438.00	273.00	536.31
taN	time of admixture giving rise to NAFT	883.4	470.89	1227.48
tbn	time of the bottleneck of NAFT ancestors	1050.51	595.00	1375.45
tdn	time of entrance of NAFT ancestors	1115.29	779.66	1347.45
tbt	time of the bottleneck of WAFZ ancestors	1224.00	948.35	1506.89
tdt	time of entrance of WAFZ ancestors	1325.00	1207.25	1561.77
tsp	time of the split between indicine and taurine	150726.00	84622.70	195130.00

The parameters estimated and their corresponding 95% confidence intervals were determined using the 0.05 and 0.95 quantiles (Q) of the posterior distribution, focusing on the most voted scenario (scenario 7). Diverging times (t) are expressed in the number of generations.

of zebu cattle associated with intensified Indian Ocean trade during this period, though osteological identification relies primarily on archaeological contexts rather than definitive morphological criteria.⁷⁶ The genetic evidence for continued zebu dispersal to Madagascar by the 14th century CE^{77,78} aligns with this archaeological evidence of increased Indian Ocean connectivity, though the specific mechanisms and scale of livestock introductions remain to be fully understood through combined archaeological and genetic approaches.

Conclusions

The tangled history of human migrations, trading routes, and eco-socio-cultural diversity as well as pastoralism gave rise to the novel and complex mosaic of African cattle genomes that have been the cornerstone of the success of African pastoralism. After identifying the main genetic groups, nine models of diffusion were tested. Our analysis provided support for two independent waves of indicine cattle dispersal into the continent. Time parameter estimates and the overall diffusion process revealed by the most voted scenario align well with known histor-

ical epochs and archaeological findings, underscoring the central role of human-livestock interrelationships in sculpturing the genetic diversity in African livestock. Northern African cattle were the result of ancient, at least Bronze Age, intermix with European taurine introduced via maritime routes. With regard to the arrival of the indicines, the rise of the Aksumite Empire seems to have provided a major boost to the initial diffusion of ancestral zebu cattle into Africa. However, the cattle expansion appears to have occurred gradually, probably alongside the expansion of Bantu-speaking communities, which facilitated hybridization processes among African cattle. A secondary entry was also highlighted to have occurred more recently, coinciding with the emergence of Swahili civilization and maritime trade between Africa, the Arabian Peninsula, and the Indian subcontinent.

These findings not only improve our understanding of historical processes that shaped the genetic makeup of contemporary African cattle but also provide genomic resources and analytical approaches that can be used in the future to accurately characterize the ancestry and the evolution of African cattle.

Limitations of the study

The demographic picture highlighted the testing of different schemes of cattle introduction and diffusion into Africa, confirming and illustrating the complex evolutionary history of African cattle characterized by a secondary introduction of European taurine in North Africa and at least two independent indicine entrances.

It is important to note that time estimates should always be taken with caution considering the associated uncertainties with mutation rates and possible issues related to ascertainment bias when using SNP chip genotype data.⁷⁹ To minimize the effect of the latter, several precautions have been used. First, linkage disequilibrium (LD)-based pruning was used as a strategy to mitigate ascertainment, as it has been proven to produce results comparable to those obtained with WGS.⁸⁰ Second, compared with SNPs derived from WGS, pre-ascertained SNPs have been shown to produce overestimated divergence times because of retaining older SNPs.^{81,82} Therefore, to overcome this issue, information on priors was calibrated using WGS to reduce possible biased estimates. In spite of these precautions, time estimates from SNP arrays should be interpreted with caution, and future studies with WGS data and new archaeogenomes will be needed for precise dating.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Gabriele Senczuk (g.senczuk@unimol.it).

Materials availability

This study did not generate new unique materials.

Data and code availability

- This article analyzes existing, publicly available data described in detail in [STAR Methods](#), with full references to the original publications. Data subsets generated in this study as well as processed genotypes are available from the [lead contact](#) with a completed materials transfer agreement.
- All software tools and pipelines used in this study are described in detail in [STAR Methods](#), with appropriate references to their original publications. Custom R and bash scripts used for data processing are available from the [lead contact](#) with a completed materials transfer agreement.
- Any additional information required to reanalyze the data reported in this study is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

We thank Diane Gifford-Gonzalez and Mary Prendergast for their valuable suggestions on archaeological literature. We are also grateful to colleagues at ILRI and ICARDA for their support during this work. The ILRI Livestock Genetics and Genomics Program and ICARDA's Dryland Livestock Genomics Research are supported by the One CGIAR SAAF Science Program and its predecessor SAPLING, funded by contributors to the CGIAR Trust Fund (<http://www.cgiar.org/about-us/our-funders/>). This research was funded in part by the Bill & Melinda Gates Foundation and with UK aid from the UK Foreign, Commonwealth and Development Office (grant agreement: INV-040641), under the auspices of the Centre for Tropical Livestock Genetics and Health (CTLGH), established jointly by the University of Edinburgh, SRUC (Scotland's Rural College), and the International Livestock Research Institute (ILRI). The findings and conclusions presented here are those of the authors and do not necessarily reflect the positions or policies of the Gates Foundation or the UK Government.

AUTHOR CONTRIBUTIONS

G.S., F.P., and O.H. designed the study; M.K.B., C.P., and M.D.C performed research and analyzed data; M.K.B., S.B.J., and S.M. performed research; G.S., M.K.B., J.M.M., and O.H. wrote the paper. All authors subsequently edited, read, and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **METHOD DETAILS**
 - Dataset composition and filtering
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Genetic diversity and population structure
 - Breed choice and dataset preparation for the ABC-RF analysis
 - Demographic simulations using ABC-RF

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2026.115295>.

Received: October 23, 2025

Revised: January 30, 2026

Accepted: March 5, 2026

Published: March 10, 2026

REFERENCES

1. Osypińska, M., Osypiński, P., Wiktorowicz, P., Chłodnicki, M., Łopaciuk, R., Bobrowski, P., Cendrowska, M., Kokolus, J., and Madani, H.K. (2025). Cattle domestication revisited: Middle Nile evidence suggests independent origins in Africa 10 000 years ago. *J. Archaeol. Sci.* *177*, 106202. <https://doi.org/10.1016/j.jas.2025.106202>.
2. Gautier, A. (2001). The Early to Late Neolithic archeofaunas from Nabta and Bir Kiseiba. In *Holocene Settlement of the Egyptian Sahara, Vol. 1: The Archaeology of Nabta Playa*, F. Wendorf and R. Schild, eds. (New York, NY: Kluwer Academic/Plenum), pp. 609–635. https://doi.org/10.1007/978-1-4615-0653-9_23.
3. Wendorf, F., and Schild, R. (2001). *Holocene Settlement of the Egyptian Sahara, Vol. 1: The Archaeology of Nabta Playa* (New York, NY: Kluwer Academic/Plenum).
4. Jórdeczka, M., Królik, H., Masojć, M., and Schild, R. (2013). Hunter-gatherer cattle-keepers of Early Neolithic El Adam type from Nabta Playa: Latest discoveries from Site E-06-1. *Afr. Archaeol. Rev.* *30*, 253–284. <https://doi.org/10.1007/s10437-013-9136-1>.
5. Chaix, L. (2017). Cattle, a major component of the Kerma culture (Sudan). In *The Oxford Handbook of Zooarchaeology, 1*, U. Albarella, M. Rizzetto, H. Russ, K. Vickers, and S. Viner-Daniels, eds. (Oxford, UK: Oxford University Press).
6. Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y., Hill, E.W., and Rege, J.E.O. (2002). African pastoralism: Genetic imprints of origins and migrations. *Science* *296*, 336–339. <https://doi.org/10.1126/science.1069878>.
7. Gifford-Gonzalez, D. (2017). Pastoralism in sub-Saharan Africa: Emergence and ramifications. In *The Oxford Handbook of Zooarchaeology*, U. Albarella, M. Rizzetto, H. Russ, K. Vickers, and S. Viner-Daniels, eds. (Oxford, UK: Oxford University Press).

8. Decker, J.E., McKay, S.D., Rolf, M.M., Kim, J., Molina Alcalá, A., Sonstegard, T.S., Hanotte, O., Götherström, A., Seabury, C.M., Praharani, L., et al. (2014). Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genet.* 10, e1004254. <https://doi.org/10.1371/journal.pgen.1004254>.
9. Mastrangelo, S., Tolone, M., Ben Jemaa, S., Sottile, G., Di Gerlando, R., Cortés, O., and Ciani, E. (2020). Refining the genetic structure and relationships of European cattle breeds through meta-analysis of worldwide genomic SNP data, focusing on Italian cattle. *Sci. Rep.* 10, 14522. <https://doi.org/10.1038/s41598-020-71375-2>.
10. Bradley, D.G., MacHugh, D.E., Loftus, R.T., Sow, R.S., Hoste, C.H., and Cunningham, E.P. (1994). Zebu-taurine variation in Y-chromosomal DNA: A sensitive assay for genetic introgression in West African trypano-tolerant cattle populations. *Anim. Genet.* 25, 7–12.
11. Ginja, C., Guimarães, S., da Fonseca, R.R., Rasteiro, R., Rodríguez-Varrel, R., Simões, L.G., Sarmento, C., Belarte, M.C., Kallala, N., Torres, J.R., et al. (2023). Iron Age genomic data from Althiburos, Tunisia renew the debate on the origins of African taurine cattle. *iScience* 26, 107196. <https://doi.org/10.1016/j.isci.2023.107196>.
12. Hanotte, O., Tawah, C.L., Bradley, D.G., Okomo, M., Verjee, Y., Ochieng, J., and Rege, J.E. (2000). Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y-specific allele amongst sub-Saharan African cattle breeds. *Mol. Ecol.* 9, 387–396. <https://doi.org/10.1046/j.1365-294X.2000.00858.x>.
13. Kwon, T., Kim, K., Caetano-Anolles, K., Sung, S., Cho, S., Jeong, C., Hanotte, O., and Kim, H. (2022). Mitonuclear incompatibility as a hidden driver behind the genome ancestry of African admixed cattle. *BMC Biol.* 20, 20. <https://doi.org/10.1186/s12915-021-01206-x>.
14. Anderung, C., Hellborg, L., Seddon, J., Hanotte, O., and Götherström, A. (2007). Investigation of X- and Y-specific single nucleotide polymorphisms in taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle. *Anim. Genet.* 38, 595–600. <https://doi.org/10.1111/j.1365-2052.2007.01663.x>.
15. Pérez-Pardal, L., Sánchez-Gracia, A., Álvarez, I., Traoré, A., Ferraz, J.B.S., Fernández, I., Costa, V., Chen, S., Tapio, M., Cantet, R.J.C., et al. (2018). Legacies of domestication, trade and herder mobility shape extant male zebu cattle diversity in South Asia and Africa. *Sci. Rep.* 8, 18027. <https://doi.org/10.1038/s41598-018-36444-7>.
16. Grigson, C. (1991). An African origin for African cattle? Some archaeological evidence. *Afr. Archaeol. Rev.* 9, 119–144. <https://doi.org/10.1007/BF01117218>.
17. Epstein, H., and Mason, I.L. (1971). *Origin of the Domestic Animals of Africa* (New York, NY: Holmes & Meier).
18. Marshall, F., and Hildebrand, E. (2002). Cattle before crops: The beginnings of food production in Africa. *J. World Prehist.* 16, 99–143. <https://doi.org/10.1023/A:1019954903395>.
19. Gifford-Gonzalez, D., and Hanotte, O. (2011). Domesticating animals in Africa: Implications of genetic and archaeological findings. *J. World Prehist.* 24, 1–23. <https://doi.org/10.1007/s10963-010-9042-2>.
20. Kim, K., Kwon, T., Dessie, T., Yoo, D., Mwai, O.A., Jang, J., Sung, S., Lee, S., Salim, B., Jung, J., et al. (2020). The mosaic genome of indigenous African cattle as a unique genetic resource for African pastoralism. *Nat. Genet.* 52, 1099–1110. <https://doi.org/10.1038/s41588-020-0694-2>.
21. Boivin, N., Crowther, A., Prendergast, M., and Fuller, D.Q. (2014). Indian Ocean food globalisation and Africa. *Afr. Archaeol. Rev.* 31, 547–581. <https://doi.org/10.1007/s10437-014-9173-4>.
22. Prendergast, M.E., Buckley, M., Crowther, A., Frantz, L., Eager, H., Lebrasseur, O., Hutterer, R., Hulme-Beaman, A., Van Neer, W., Douka, K., et al. (2017). Reconstructing Asian faunal introductions to eastern Africa from multi-proxy biomolecular and archaeological datasets. *PLoS One* 12, e0182565. <https://doi.org/10.1371/journal.pone.0182565>.
23. Munro-Hay, S. (1991). *Aksum: An African Civilisation of Late Antiquity* (Edinburgh, UK: Edinburgh University Press).
24. Phillipson, D.W. (2000). *Archaeology at Aksum, Ethiopia, 1993–7*, 2 (London, UK: Society of Antiquaries of London).
25. Woldekiros, H.S., and D’Andrea, A.C. (2017). Early evidence for domestic chickens (*Gallus gallus domesticus*) in the Horn of Africa. *Int. J. Osteoarchaeol.* 27, 329–341. <https://doi.org/10.1002/oa.2540>.
26. Woldekiros, H.S. (2023). *The Boundaries of Ancient Trade: Kings, Commoners, and the Aksumite Salt Trade of Ethiopia* (Boulder, CO: University Press of Colorado).
27. Mahgoub, O., Babiker, H.A., Kadim, I.T., Al-Kindi, M., Hassan, S., Al-Marzooqi, W., Eltahir, Y.E., Al-Abri, M.A., Al-Khayat, A., Al-Sinani, K.R., et al. (2013). Disclosing the origin and diversity of Omani cattle. *Anim. Genet.* 44, 336–339. <https://doi.org/10.1111/j.1365-2052.2012.02399.x>.
28. Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M., et al. (2009). The genetic structure of African populations and African Americans. *Science* 324, 1031–1038. <https://doi.org/10.1126/science.1172257>.
29. Wang, K., Goldstein, S., Bleasdale, M., Clist, B., Bostoen, K., Bakwa-Lufu, P., Buck, L.T., Crowther, A., Dème, A., McIntosh, R.J., et al. (2020). Ancient genomes reveal complex patterns of population movement, interaction, and replacement in sub-Saharan Africa. *Sci. Adv.* 6, eaaz0183. <https://doi.org/10.1126/sciadv.aaz0183>.
30. Koile, E., Greenhill, S.J., Blasi, D.E., Bouckaert, R., and Gray, R.D. (2022). Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proc. Natl. Acad. Sci. USA* 119, e2112853119. <https://doi.org/10.1073/pnas.2112853119>.
31. Fortes-Lima, C.A., Burgarella, C., Hammarén, R., Eriksson, A., Vicente, M., Jolly, C., Semo, A., Gunnink, H., Pacchiarotti, S., Mundeke, L., et al. (2024). The genetic legacy of the expansion of Bantu-speaking peoples in Africa. *Nature* 625, 540–547. <https://doi.org/10.1038/s41586-023-06770-6>.
32. Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., and Robert, C.P. (2016). Reliable ABC model choice via random forests. *Bioinformatics* 32, 859–866. <https://doi.org/10.1093/bioinformatics/btv684>.
33. Fraimout, A., Debat, V., Fellous, S., Hufbauer, R.A., Foucaud, J., Pudlo, P., Marin, J.-M., Price, D.K., Cattel, J., Chen, X., et al. (2017). Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. *Mol. Biol. Evol.* 34, 980–996. <https://doi.org/10.1093/molbev/msx050>.
34. Collin, F.-D., Durif, G., Raynal, L., Lombaert, E., Gautier, M., Vitalis, R., Marin, J.M., and Estoup, A. (2021). Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. *Mol. Ecol. Resour.* 21, 2598–2613. <https://doi.org/10.1111/1755-0998.13413>.
35. Chapuis, M.-P., Raynal, L., Augé, M., Meynard, C.N., Blondin, L., Marin, J.M., Estoup, A., Maquet, M., Moquet, L., Sallé, A., et al. (2020). A young age of subspecific divergence in the desert locust inferred by ABC random forest. *Mol. Ecol.* 29, 4542–4558. <https://doi.org/10.1111/mec.15663>.
36. Campos, M., Pérez-Collazos, E., Díaz-Pérez, A., López-Alvarez, D., Oumouloud, A., Mur, L.A.J., Vogel, J.P., and Catalán, P. (2024). Repeated migration, interbreeding and bottlenecking shaped the phylogeography of the selfing grass *Brachypodium stacei*. *Mol. Ecol.* 33, e17513. <https://doi.org/10.1111/mec.17513>.
37. Senczuk, G., Mastrangelo, S., Ajmone-Marsan, P., Becskei, Z., Colan-gelo, P., Collì, L., Ferretti, L., Karsli, T., Lancioni, H., Lasagna, E., et al. (2021). On the origin and diversification of Podolian cattle breeds: Testing scenarios of European colonization using genome-wide SNP data. *Genet. Sel. Evol.* 53, 48. <https://doi.org/10.1186/s12711-021-00639-w>.
38. Estoup, A., Raynal, L., Verdu, P., Marin, J.-M., Robert, C.P., and Pudlo, P. (2018). Model choice using approximate Bayesian computation and random forests: Analyses based on model grouping to make inferences about the genetic history of Pygmy human populations. *J. Soc. Fr. Stat.* 159, 167–190. <https://hal.inrae.fr/hal-02618438>.
39. Ghirotto, S., Vizzari, M.T., Tassi, F., Barbujani, G., and Benazzo, A. (2021). Distinguishing among complex evolutionary models using

- unphased whole-genome data through random forest approximate Bayesian computation. *Mol. Ecol. Resour.* 21, 2614–2628. <https://doi.org/10.1111/1755-0998.13263>.
40. Makina, S.O., Muchadeyi, F.C., van Marle-Köster, E., MacNeil, M.D., and Maiwashe, A. (2014). Genetic diversity and population structure among six cattle breeds in South Africa using a whole-genome SNP panel. *Front. Genet.* 5, 333. <https://doi.org/10.3389/fgene.2014.00333>.
 41. Edea, Z., Bhuiyan, M.S.A., Dessie, T., Rothschild, M.F., Dadi, H., Kim, K.S., and Kim, K.S. (2015). Genome-wide genetic diversity, population structure and admixture analysis in African and Asian cattle breeds. *Animal* 9, 218–226. <https://doi.org/10.1017/S1751731114002560>.
 42. Kim, S.-J., Ka, S., Ha, J.W., Kim, J., Yoo, D., Kim, K., Lee, H.K., Lim, D., Cho, S., Hanotte, O., et al. (2017). Cattle genome-wide analysis reveals genetic signatures in trypanotolerant N'Dama. *BMC Genom.* 18, 371. <https://doi.org/10.1186/s12864-017-3742-2>.
 43. Gebrehiwot, N.Z., Strucken, E.M., Alliloo, H., Marshall, K., and Gibson, J.P. (2020). The patterns of admixture, divergence, and ancestry of African cattle populations determined from genome-wide SNP data. *BMC Genom.* 21, 869. <https://doi.org/10.1186/s12864-020-07270-x>.
 44. Friedrich, J., Bailey, R., Talenti, A., Marshall, K., Prendergast, J., and Wiener, P. (2022). Exploring genetic diversity and selection signatures in indigenous African cattle breeds. In *Proc. 12th World Congress on Genetics Applied to Livestock Production (WCGALP)*, R.F. Veerkamp and Y. de Haas, eds. (Wageningen, The Netherlands: Wageningen Academic Publishers), pp. 810–813.
 45. Kim, K., Kim, D., Hanotte, O., Lee, C., Kim, H., and Jeong, C. (2023). Inference of admixture origins in indigenous African cattle. *Mol. Biol. Evol.* 40, msad257. <https://doi.org/10.1093/molbev/msad257>.
 46. Loftus, R.T., MacHugh, D.E., Ngere, L.O., Balain, D.S., Badi, A.M., Bradley, D.G., and Cunningham, E.P. (1994). Mitochondrial genetic variation in European, African and Indian cattle populations. *Anim. Genet.* 25, 265–271. <https://doi.org/10.1111/j.1365-2052.1994.tb00203.x>.
 47. Pramod, R.K., Velayutham, D., Singh, M., Tomar, S.S., Zachariah, A., Zachariah, A., B. C., S. S., P. G., Dhinoth Kumar, B., et al. (2019). Complete mitogenome reveals genetic divergence and phylogenetic relationships among Indian cattle (*Bos indicus*) breeds. *Anim. Biotechnol.* 30, 219–232. <https://doi.org/10.1080/10495398.2018.1476376>.
 48. Sarath Kumar, T., Singh, S., Ganguly, I., and Dixit, S.P. (2025). Mitogenome diversity and evolution of *Bos indicus* cattle in India. *Gene* 936, 149096. <https://doi.org/10.1016/j.gene.2024.149096>.
 49. Rossi, C., Sinding, M.H.S., Mullin, V.E., Scheu, A., Erven, J.A.M., Verdugo, M.P., Daly, K.G., Ciucani, M.M., Mattiangeli, V., Teasdale, M.D., et al. (2024). The genomic natural history of the aurochs. *Nature* 635, 136–141. <https://doi.org/10.1038/s41586-024-08112-6>.
 50. di Lernia, S. (2021). Earliest herders of the Central Sahara (Tadrart Acacus Mountains, Libya): A punctuated model for the emergence of pastoralism in Africa. *J. World Prehist.* 34, 531–594. <https://doi.org/10.1007/s10963-021-09162-8>.
 51. Manning, K., Breeze, P.S., Drake, N., Dunne, J., Casanova, E., and Evershed, R.P. (2023). Habitat fragmentation and the sporadic spread of pastoralism in the mid-Holocene Sahara. *Quat. Sci. Rev.* 309, 108070. <https://doi.org/10.1016/j.quascirev.2023.108070>.
 52. Ben-Jemaa, S., Mastrangelo, S., Lee, S.H., Lee, J.H., and Boussaha, M. (2020). Genome-wide scan for selection signatures reveals novel insights into the adaptive capacity in local North African cattle. *Sci. Rep.* 10, 19466. <https://doi.org/10.1038/s41598-020-76576-3>.
 53. Boushaba, N., Boujenane, I., Moazami-Goudarzi, K., Flori, L., Saïdi-Meh-tar, N., Tabet-Aoul, N., and Laloë, D. (2019). Genetic diversity and relationships among six local cattle populations in semi-arid areas assessed by a bovine medium-density single nucleotide polymorphism data. *Animal* 13, 8–14. <https://doi.org/10.1017/S1751731118001179>.
 54. Bonfiglio, S., Achilli, A., Olivieri, A., Negrini, R., Colli, L., Liotta, L., Ajmone-Marsan, P., Torroni, A., and Ferretti, L. (2010). The enigmatic origin of bovine mtDNA haplogroup R: Sporadic interbreeding or an independent event of *Bos primigenius* domestication in Italy? *PLoS One* 5, e15760. <https://doi.org/10.1371/journal.pone.0015760>.
 55. Di Lorenzo, P., Lancioni, H., Ceccobelli, S., Colli, L., Cardinali, I., Karsli, T., Capodiferro, M.R., Sahin, E., Ferretti, L., Ajmone Marsan, P., et al. (2018). Mitochondrial DNA variants of Podolian cattle breeds testify for a dual maternal origin. *PLoS One* 13, e0192567. <https://doi.org/10.1371/journal.pone.0192567>.
 56. Verdugo, M.P., Mullin, V.E., Scheu, A., Mattiangeli, V., Daly, K.G., Maisano Delsler, P., Hare, A.J., Burger, J., Collins, M.J., Kehati, R., et al. (2019). Ancient cattle genomics, origins, and rapid turnover in the Fertile Crescent. *Science* 365, 173–176. <https://doi.org/10.1126/science.aav1002>.
 57. Beja-Pereira, A., Caramelli, D., Lalueza-Fox, C., Vernesi, C., Ferrand, N., Casoli, A., Goyache, F., Royo, L.J., Conti, S., Lari, M., et al. (2006). The origin of European cattle: Evidence from modern and ancient DNA. *Proc. Natl. Acad. Sci. USA* 103, 8113–8118. <https://doi.org/10.1073/pnas.0509210103>.
 58. Lucas-Sánchez, M., Serradell, J.M., and Comas, D. (2021). Population history of North Africa based on modern and ancient genomes. *Hum. Mol. Genet.* 30, R17–R23. <https://doi.org/10.1093/hmg/ddaa261>.
 59. Lipson, M., Ringbauer, H., Lucarini, G., Aouadi, N., Aoudia, L., Belhouchet, L., Cheronet, O., Dahmani, A.-R., Genchi, F., La Pastina, F., et al. (2025). High continuity of forager ancestry in the Neolithic period of the eastern Maghreb. *Nature* 623, 1–7. <https://doi.org/10.1038/s41586-025-08699-4>.
 60. Barker, G. (2006). *The Agricultural Revolution in Prehistory: Why Did Foragers Become Farmers?* (Oxford, UK: Oxford University Press).
 61. Knapp, A.B. (2016). *Mediterranean Connections: Maritime Transport Containers and Seaborne Trade in the Bronze and Early Iron Ages* (London, UK: Routledge).
 62. A. Kouremenos and J.M. Gordon, eds. (2020). *Mediterranean Archaeologies of Insularity in an Age of Globalization* (Oxford, UK: Oxbow Books).
 63. Iacono, F., Borgna, E., Cattani, M., Cavazzuti, C., Dawson, H., Galanakis, Y., Gori, M., Iaia, C., Ialongo, N., Lachenal, T., et al. (2022). Establishing the Middle Sea: The Late Bronze Age of Mediterranean Europe (1700–900 BC). *J. Archaeol. Res.* 30, 371–445. <https://doi.org/10.1007/s10814-021-09165-1>.
 64. Benattia, H., Bokbot, Y., Onrubia-Pintado, J., Benerradi, M., Bougariane, B., Bouhamidi, B., Carballo-Pérez, J., Echcherif-Baamrani, O., Elqably, A., Ghayati, N., et al. (2025). Rethinking late prehistoric Mediterranean Africa: Architecture, farming and materiality at Kach Kouch, Morocco. *Antiquity* 99, 1–21. <https://doi.org/10.15184/aaq.2025.10>.
 65. Sanmartí, J., Kallala, N., Belarte, M.C., Torres, J.R., Telmini, B.M., Niella, R.J., and Miniaoui, S. (2012). Filling gaps in the protohistory of the eastern Maghreb: The Althiburos Archaeological Project (El Kef, Tunisia). *J. Afr. Archaeol.* 10, 21–44.
 66. Bell, C. (2016). *Phoenician trade: The first 300 years. In Dynamics of Production in the Ancient Near East, 1300–500 BCE, C.A.K. Johnstone and J. Phil, eds. (Oxford, UK: Oxbow Books), pp. 57–76.*
 67. Cain, C.R. (2000). *Animals at Axum: Initial Zooarchaeological Research in the Later Prehistory of the Northern Ethiopian Highlands*. PhD Thesis (St Louis: Washington University).
 68. Woldekios, H.S., and D'Andrea, A.C. (2022). Complex (multispecies) livestock keeping: Highland agricultural strategy in the northern Horn of Africa during the Pre-Aksumite (1600 BCE–400 BCE) and Aksumite (400 BCE–CE 800) periods. *Front. Ecol. Evol.* 10, 901446. <https://doi.org/10.3389/fevo.2022.901446>.
 69. Clark, J.D., and Prince, G.R. (1978). Use-wear on later Stone Age micro-liths from Laga Oda, Haraghi, Ethiopia and possible functional interpretations. *Azania J. Br. Inst. East Afr.* 13, 101–110. <https://doi.org/10.1080/00672707809511633>.
 70. Clark, J.D., and Williams, M.A.J. (1978). Recent archaeological research in southeastern Ethiopia, 1974–1975. *ethio.* 11, 19–44. <https://doi.org/10.3406/ethio.1978.902>.

71. Plug, I., and Voigt, E.A. (1985). Archaeozoological studies of Iron Age communities in southern Africa. *Adv. World Archaeol.* 4, 189–238.
72. Huffman, T.N. (2005). *Mapungubwe: Ancient African Civilisation on the Limpopo* (Johannesburg: Wits University Press).
73. Fleisher, J., Lane, P., LaViolette, A., Horton, M., Pollard, E., Quintana Morales, E., Vernet, T., Christie, A., and Wynne-Jones, S. (2015). When did the Swahili become maritime? *Am. Anthropol.* 117, 100–115. <https://doi.org/10.1111/aman.12171>.
74. Morales, E.Q., and Prendergast, M.E. (2017). Animals and their uses in the Swahili world. In *The Swahili World*, S. Wynne-Jones and A. LaViolette, eds. (London, UK: Routledge), pp. 335–349.
75. Kusimba, C.M. (2024). *Swahili Worlds in Globalism* (Cambridge, UK: Cambridge University Press).
76. Radimilahy, C. (1998). In *Mahilaka: An archaeological investigation of an early town in northwestern Madagascar*. *Studies in African Archaeology* 15. Uppsala: Department of Archaeology and Ancient History, PhD thesis (University of Uppsala).
77. Magnier, J., Druet, T., Naves, M., Ouvrard, M., Raoul, S., Janelle, J., Moazami-Goudarzi, K., Lesnoff, M., Tillard, E., Gautier, M., et al. (2022). The genetic history of Mayotte and Madagascar cattle breeds mirrors the complex pattern of human exchanges in the Western Indian Ocean. *G3 (Bethesda)* 12, jkac029. <https://doi.org/10.1093/g3journal/jkac029>.
78. Jones, M.B., Kapumba, R., Chirikure, S., and Marshall, F. (2024). Hunting and the social lives of southern Africa's first farmers. *J. Archaeol. Res.* 32, 597–636. <https://doi.org/10.1007/s10814-023-09194-y>.
79. Quinto-Cortés, C.D., Woerner, A.E., Watkins, J.C., and Hammer, M.F. (2018). Modeling SNP array ascertainment with approximate Bayesian computation for demographic inference. *Sci. Rep.* 8, 10209. <https://doi.org/10.1038/s41598-018-28539-y>.
80. Malomane, D.K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A.R., and Simianer, H. (2018). Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genom.* 19, 22. <https://doi.org/10.1186/s12864-017-4416-9>.
81. Wollstein, A., Lao, O., Becker, C., Brauer, S., Trent, R.J., Nürnberg, P., Stoneking, M., and Kayser, M. (2010). Demographic history of Oceania inferred from genome-wide data. *Curr. Biol.* 20, 1983–1992. <https://doi.org/10.1016/j.cub.2010.10.040>.
82. Lachance, J., and Tishkoff, S.A. (2013). SNP ascertainment bias in population genetic analyses: Why it is important and how to correct it. *Bioessays* 35, 780–786. <https://doi.org/10.1002/bies.201300014>.
83. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. <https://doi.org/10.1086/519795>.
84. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>.
85. Gruber, B., Unmack, P.J., Berry, O.F., and Georges, A. (2018). DARTR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* 18, 691–699. <https://doi.org/10.1111/1755-0998.12745>.
86. Barbato, M., Orozco-terWengel, P., Tapio, M., and Bruford, M.W. (2015). SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* 6, 109. <https://doi.org/10.3389/fgene.2015.00109>.
87. Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. <https://doi.org/10.1101/gr.094052.109>.
88. Fitak, R.R. (2021). OptM: Estimating the optimal number of migration edges on population trees using TreeMix. *Biol. Methods Protoc.* 6, bpab017. <https://doi.org/10.1093/biomethods/bpab017>.
89. Milanese, M., Capomaccio, S., Vajana, E., Bomba, L., Garcia, J.F., Ajmone-Marsan, P., and Colli, L. (2017). BITE: An R package for biodiversity analyses. Preprint at bioRxiv. <https://doi.org/10.1101/181610>.
90. Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8, e1002967. <https://doi.org/10.1371/journal.pgen.1002967>.
91. Petkova, D., Novembre, J., and Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* 48, 94–100. <https://doi.org/10.1038/ng.3464>.
92. Raynal, L., Marin, J.-M., Pudlo, P., Ribatet, M., Robert, C.P., and Estoup, A. (2019). ABC random forests for Bayesian parameter inference. *Bioinformatics* 35, 1720–1728. <https://doi.org/10.1093/bioinformatics/bty867>.
93. Flori, L., Moazami-Goudarzi, K., Alary, V., Araba, A., Boujenane, I., Boushaba, N., Casabianca, F., Casu, S., Ciampolini, R., Coeur D'Acier, A., et al. (2019). A genomic map of climate adaptation in Mediterranean cattle breeds. *Mol. Ecol.* 28, 1009–1029. <https://doi.org/10.1111/mec.15004>.
94. Ben Jemaa, S., Boussaha, M., Ben Mehdi, M., Lee, J.H., and Lee, S.-H. (2015). Genome-wide insights into population structure and genetic history of Tunisian local cattle using the Illumina BovineSNP50 BeadChip. *BMC Genom.* 16, 677. <https://doi.org/10.1186/s12864-015-1638-6>.
95. Bahbahani, H., Tijjani, A., Mukasa, C., Wragg, D., Almathen, F., Nash, O., Akpa, G.N., Mbole-Kariuki, M., Malla, S., Woolhouse, M., et al. (2017). Signatures of selection for environmental adaptation and Zebu × Taurine hybrid fitness in East African Shorthorn Zebu. *Front. Genet.* 8, 68. <https://doi.org/10.3389/fgene.2017.00068>.
96. Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassel, C.P., Sonstegard, T.S., et al. (2009). A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* 10, R42. <https://doi.org/10.1186/gb-2009-10-4-r42>.
97. Gautier, M., Laloë, D., Moazami-Goudarzi, K., Navratil, V., Foglio, M., Grohs, C., Boland, A., Garnier, J.G., Boichard, D., Lathrop, G.M., et al. (2007). Genetic and haplotypic structure in 14 European and African cattle breeds. *Genetics* 177, 1059–1070. <https://doi.org/10.1534/genetics.107.075804>.
98. MacEachern, S., Hayes, B., McEwan, J., and Goddard, M. (2009). An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (*Bos taurus*) using a high-density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in domestic cattle. *BMC Genom.* 10, 181. <https://doi.org/10.1186/1471-2164-10-181>.
99. Flori, L., Thevenon, S., Dayo, G.K., Senou, M., Sylla, S., Berthier, D., Moazami-Goudarzi, K., and Gautier, M. (2014). Adaptive admixture in the West African bovine hybrid zone: Insight from the Borgou population. *Mol. Ecol.* 23, 3241–3257. <https://doi.org/10.1111/mec.12816>.
100. Molina Flores, B., Camacho Vallejo, M.E., Delgado Bermejo, J.V., Navas González, F.J., and Martínez, M.D.A. (2023). Do Pharaohs' cattle still graze the Nile Valley? Genetic characterization of the Egyptian Baladi cattle breed. *Anim. Biotechnol.* 34, 645–657. <https://doi.org/10.1080/10495398.2021.1991938>.
101. Vanvanhossou, S.F.U., Yin, T., Scheper, C., Fries, R., Dossa, L.H., and König, S. (2021). Unraveling admixture, inbreeding, and recent selection signatures in West African indigenous cattle populations in Benin. *Front. Genet.* 12, 657282. <https://doi.org/10.3389/fgene.2021.657282>.
102. Zerabruk, M., Li, M.H., Kantanen, J., Olsaker, I., Ibeagha-Awemu, E.M., Erhardt, G., and Vangen, O. (2012). Genetic diversity and admixture of indigenous cattle from North Ethiopia: Implications of historical introgressions in the gateway region to Africa. *Anim. Genet.* 43, 257–266. <https://doi.org/10.1111/j.1365-2052.2011.02245.x>.
103. Gautier, M., Flori, L., Riebler, A., Jaffrézic, F., Laloë, D., Gut, I., Moazami-Goudarzi, K., and Foulley, J.L. (2009). A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC Genom.* 10, 550. <https://doi.org/10.1186/1471-2164-10-550>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
PLINK 1.9	Purcell et al. ⁸³ and Chang et al. ⁸⁴	https://doi.org/10.1086/519795 https://doi.org/10.1186/s13742-015-0047-8
dartR	Gruber et al. ⁸⁵	https://doi.org/10.1111/1755-0998.12745
SneP V1.11	Barbato et al. ⁸⁶	https://doi.org/10.3389/fgene.2015.00109
Admixture V1.3.0	Alexander et al. ⁸⁷	https://doi.org/10.1101/gr.094052.109
OptM	Fitak. ⁸⁸	https://doi.org/10.1093/biomet/bpab017
BITEV2	Milanesi et al. ⁸⁹	https://doi.org/10.1101/181610
TreeMix-1.13	Pickrell and Pritchard. ⁹⁰	https://doi.org/10.1371/journal.pgen.1002967
Estimating Effective migration surface (EEMS)	Petkova et al. ⁹¹	https://doi.org/10.1038/ng.3464 https://github.com/dipetkov/eems/tree/master
DIYABC-RF framework (diyabc-rf-linux-v1.1.51;abcranger-linux-v1.16.67)	Pudlo et al. ³² ; Collin et al. ³⁴ and Raynal et al. ⁹²	https://doi.org/10.1111/1755-0998.13413 https://doi.org/10.1093/bioinformatics/btv684 https://doi.org/10.1093/bioinformatics/bty867 https://diyabc.github.io/cli/
Other		
Filtered whole dataset for the genetic structure analysis	This Paper	–
Filtered reduced dataset for ABC-RF	This paper	–

METHOD DETAILS

Dataset composition and filtering

Genotype data were obtained from public sources and collaborative research efforts. It included 607 individuals genotyped with the Illumina BovineSNP50 chip^{8,93,94} and 436 individuals genotyped with the BovineHD chip.⁹⁵

All genotypes' positions were updated to the ARS-UCD 2.0 reference genome assembly map⁹⁶ and then merged. SNPs with minor allele frequencies (MAF) below 0.01 and genotypic call rates under 90% were removed using PLINK 1.9,^{83,84} leaving 34,699 autosomal SNPs for analyses.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genetic diversity and population structure

A total of 1043 individuals of 38 cattle breeds were analyzed at 34,699 autosomal polymorphic SNPs (Table S1). Three population genetic parameters, observed heterozygosity (H_o), expected heterozygosity (H_e) and effective population size (N_e), were used to assess genetic diversity. For each breed, H_o and H_e were estimated using the R package dartR.⁸⁵ N_e was estimated using SNeP v1.11 software⁸⁶ at 13 generations back (N_{e13}), equivalent to 65 years ago, assuming a generation interval of 5 years for indigenous African cattle.^{20,97–99}

The genetic diversity among breeds was visualized using Principal Component Analysis (PCA) performed with the *bite.biostat* function of BITEV2 R package.⁸⁹ The most significant patterns of genetic variation were visualized by plotting the first two principal components (PC1 and PC2) of the PCA, which illustrate the global level of indicine/taurine (PC1) and European taurine/African taurine (PC2) genome backgrounds in the analyzed populations.

ADMIXTURE V1.3.0 software⁸⁷ was used to investigate admixture patterns among African cattle. The analysis assumed $2 \leq K \leq 38$ hypothetical ancestral populations. The co-ancestry coefficients were used to determine and visualize individual genome partitions at each K using BITEV2 R package while the *membercoeff.cv* function of the same package was used to determine the most optimal K based on the lowest cross-validation (CV) error value.

To infer gene flow among African cattle, the maximum likelihood (ML) algorithm implemented in TreeMix-1.13⁹⁰ was used. The stratified allele frequencies from PLINK were converted into TreeMix format using the *plink2treemix.py* script. An initial trial run was conducted with 10 repetitions, assuming 1 to 10 migration edges. The *plot_optM* function of OptM package⁸⁸ was used to calculate the ideal number of migration edges (Figure S5). These two procedures were performed following 1000 bootstrap replications. Linkage disequilibrium (LD) was accounted for within 500 SNP blocks, and *B. javanicus* (OBJ) was used as the out-group. We selected the trees with the highest likelihood, eliminated duplicates, and retained those with a unique topology to construct the consensus tree from 10 independent runs using the optimal number of predicted migrations. The ultimate consensus tree was visualized using the *treemix.bootstrap* function of BITEV2 R package.

Breed choice and dataset preparation for the ABC-RF analysis

For the analysis, we grouped breeds based on converging evidence from ADMIXTURE, PCA, TreeMix, and the breed's geographic origin. Thus, seven population groups were defined: North African taurine (NAFT) (Blonde du Cap Bon (BLCAP), Brune de l'Atlas Fauve (BRATF), Brune de l'Atlas Grise (BRATG), Cheurfa (CHE), Guelmoise (GUEL) and Tidili (TID)); West African taurine (WAFT) (Baoule (BAO), Somba (SOM) and N'Dama (NDAM)); West African zebu (WAFZ) (Bunaji (BJ), Red bororo (OR), Sokoto gudali (SO), Yakanaji (YK), Zebu Bororo (ZBO), Zebu Fulani (ZFU) and Adamawa gudali (AG)); African Sanga (AFSA) (Africander (AFR), Ankole-Watusi (ANKW), Ankole (AO), Sheko (SHK) and Tuli (TULI)); East African zebu (EAFZ) (East African Shorthorn Zebu (EASZ), Serere Zebu (ZS), Karamojong Zebu (KR) and Madagascar Zebu (ZMA)); European taurine (ETAU; Holstein) and Asian indicine (IND; Gir).

Some African breeds were excluded for the following reasons. The Egyptian Baladi (BAL) shows an admixed heterogeneous pattern. It occurs close to the Sanga breeds in Treemix (Figure 2A). A recent study using microsatellite data from 160 individuals also showed the uniqueness and unexpected genetic diversity of this breed despite massive recent introductions in Egypt's Nile Valley basin of exotic breeds.¹⁰⁰ However, due to the small sample size and the complex genomic architecture of Baladi we excluded it from the ABC-RF analysis. Also, a considerable proportion of common ancestry with European taurine (HOL) is present in two Algerian breeds (CHF and BIS), as reported in previous studies.^{52,53} It may have followed crossbreeding for higher dairy performance, resulting in their genetic dilution. Thus, we decided not to include these two Algerian breeds in the ABC-RF analysis. All African taurine breeds showed close genetic affinities among them (Figures 1A, 1C, and 2A). However, two breeds (LAG and MT) displayed a similar pattern and low genetic variability and were thus not included in the ABC-RF analysis. They showed a long branch length in the TreeMix cladogram, suggesting genetic drift could be driving their divergence. Interestingly, our results parallel those of Vanvanhosou et al.,¹⁰¹ who found high genomic inbreeding probably arising from the small number of contiguous populations scattered across the West African savanna and forest regions.

To visualize spatially heterogeneous isolation-by-distance patterns of the selected breeds in a geographic map, we adopted the method called Estimating Effective Migration Surfaces (EEMS) developed by Petkova et al.⁹¹ Under the stepping-stone model, the approach allows to detect geographic areas either of genetic permeability characterized by high effective migration or possible barriers exemplified by low gene flow. For this analysis, we run 10 million Markov Chain Monte Carlo (MCMC) iterations, discarding the initial 3 million iterations as burn-in, keeping samples every 10,000 iterations and setting a grid of 1,800 demes. All samples (665 georeferenced individuals) for the retained breeds of the aforementioned seven population groups were analyzed using 34,699 autosomal SNPs.

For the ABC-RF analysis, 28 individuals were selected from each group using the *bite.kmeans.subsampling* function implemented in BITEV2 R package, to minimize computational burden while retaining the overall genetic structure of the original dataset. After merging, the resulting dataset, which consisted of 196 individuals and 35,107 SNPs, was pruned for SNPs in LD, to strictly exclude all SNPs under selection and deviating from neutrality. In doing this, we adopted hard filtering by removing SNPs with a variance inflation factor of 1.2 (almost independent) within 50 SNPs and shifting the windows every 10 SNPs (-indep 50 10 1.2) thus retaining a final dataset of 12,699 informative SNPs.

To assess the representativeness of the dataset, we once again performed PCA, ADMIXTURE and TreeMix analyses invoking the same parameters used for the entire dataset. For each group, trends in *N_e* across generations were estimated with SNeP v1.11 (Figure S6).

Demographic simulations using ABC-RF

The genetic structure highlighted by PCA, ADMIXTURE and TreeMix using the 12,699 SNPs did not differ from the one obtained using 34,699 SNPs (Figures 1C, 1D, and 2B), suggesting that the overall structure is recovered. Again, all the analyses discriminate well among the seven population groups, with only a moderate overlap between WAFZ and EAFZ.

Based on prior published genomics information on African cattle,^{6,11,20} nine competing scenarios, which include one or two waves of introduction of indicine cattle and their subsequent admixture with taurine cattle, were modeled (Figure S7).

The one wave hypothesis considered three scenarios (1, 2 and 3). We assumed that following the split of indicine and taurine cattle (tsp), ancestral taurine cattle dispersed first into Africa (tdt) as suggested by archaeological evidence.¹⁹ After their arrival, a secondary

introduction of European taurine (tdn) with subsequent admixture (TaN) with resident North African taurine cattle was modeled. The need to model this introduction arises both from observing the current genetic diversity of North African cattle and from archaeological evidence that indicates the presence of mixed genomes during the Iron Age.¹¹ A single introduction of ancestral Asian indicine cattle was then modeled (tdi1). From this point, three scenarios assuming different hybridization hypotheses were considered. In scenarios 1 and 2, we modeled an admixture event (taS) between West African taurine (WAFt) and East African ancestral zebu (EAFZ) cattle, an event that gave rise to African sanga cattle (AFSA). We used the WAFt to model these scenarios, assuming that they represent the ancient remnants of the original taurine cattle that arrived in Africa via Northeast Africa which later hybridized with the zebu cattle. The center of early taurine – zebu hybridization, potentially involving now-extinct indigenous taurine populations, remains speculative, but it may have occurred around the Lake Victoria region and East Africa.^{6,102} The resulting AFSA subsequently dispersed to central and Southern Africa.

In scenario 1, the divergence time between East and West African zebu (tdW) is constrained to occur after the formation of AFSA, whereas in scenario 2, divergence is set to precede the admixture.

Scenario 3 presents an intentionally contrasting hypothesis, wherein AFSA is modeled as the outcome of hybridization between West and East African zebu. Although this model is unlikely given the phenotypic characteristics of the sanga and their known indicine x taurine genome background, it was included to serve as a contrasting scenario, aiding model discrimination within the ABC-RF framework. Including such edge-case scenarios enhances the capacity of the model to identify the most likely historical pathway by improving resolution across the full spectrum of demographic alternatives.

The two waves hypothesis postulated two introductions with six scenarios (4, 5, 6, 7, 8 and 9). Under these scenarios, the splitting (tsp, tdt, tdn) and admixture (taN) times leading to the formation of NAFT and WAFt, were retained in the same setting as modeled in the one wave hypothesis. From this point, the underlying assumption was that the ancestral Asian indicine cattle arrived twice on the African continent at different times (tdi1 and tdi2). Similarly, different combinations of admixture and time of entry were modeled (Figure S6). Scenario 4 assumes that all indicine entered in the first wave (tdi1), massively hybridized with resident taurine until their genomic component was completely diluted. Successively, a second indicine wave entered Africa (tdi2) and diversified into West and East African zebu (tdw). Scenario 5 assumes a similar context but in this case the indicine genome of the first entry (tdi1) persisted after admixture with the resident African taurine (taS) giving rise to sanga cattle (AFSA) and West African zebu (WAFZ). Scenario 6 instead predicts a first entry (tdi1) that gave rise to the West African indicine (WAFZ) and a second entry (tdi2) that, following hybridization with resident African taurine (WAFt), led to the formation of sanga (AFSA) and East African zebu (EAFZ). In scenario 7 we modeled a first indicine entrance (tdi1) that after hybridization with local African taurine (WAFt) gave rise to sanga (AFSA) and East African zebu (EAFZ). A secondary introduction (tdi2) was then intended to justify the differentiation of West African zebu (WAFZ). Scenario 8 is similar to scenario 4, but in this case the first entry (tdi1) gave rise to West and East African zebu while a second introduction (tdi2) dispersed, hybridized with resident taurine (taS), until it was completely diluted in sanga cattle. Finally, scenario 9 assumes an early entrance of indicine (tdi1) which gave rise to East African zebu (EAFZ) and a subsequent introduction (tdi2) that following hybridization with resident taurine (WAFt), gave rise to sanga (AFSA) and West African zebu (WAFZ).

For all nine scenarios, four time points of bottleneck events (tbt, tbn, tbi1 and tbi2), as is expected when an introduction event occurs, were modeled. The modeling also made provision for unsampled ancestral populations that may have been present on the continent but are now extinct. A detailed description of the used priors and summary statistics is reported in Figure S8 and Table S3.

All nine scenarios were drawn and simulated with the DIYABC-RF software.³² The demographic and historical parameters included seven effective population sizes: N1 (ETAU), N2 (IND), N3 (WAFt), N4 (EAFZ), N5 (NAFT), N6 (AFSA) and N7 (WAFZ); five bottleneck events (N3b, N4b, N7b, N8b and N9b) at four time points (tbt, tbn, tbi1 and tbi2); two ancestral unsampled populations (N8 and N9); six divergence times (tdi1, tdi2, tdW, tdn, tdt and tsp) and two admixture time events (taN and taS) (Table 2 and Figure S6). A pilot run was performed with permissive priors to optimize prior distribution boundary conditions. To evaluate the optimal number of datasets to simulate, we followed the recommendation by Pudlo et al.³² by checking global prior error rates from both the entire reference table and subsets of the reference table. Based on this, the trends of prior error rates were checked every 10,000 simulations, running between 20,000 and 70,000 simulations, including 1024 summaries and using 100 threads. In the absence of reliable alternative calibration points based on either fossil record or previous molecular estimates, an internal calibration point, the indicine/taurine admixture time (taS) has been incorporated. Here we constrained the lower and upper bounds of the uniform distribution at 100 and 500, as suggested by WGS estimates based on LD-based admixture time (ALDER)²⁰ while all the other priors were set by default (Table S3).

The goodness of fit of the proposed models with relative priors to the observed dataset was assessed using the projection on the first two axes of the Linear Discriminant Analysis (LDA).

After evaluating the convergence of the out-of-bag error (OOB) to the number of trees in the forest, 500 trees were used to select the optimal model and estimate parameters. To convert the number of generations to historical time, a generation interval of 5 years^{20,97,99,103} was used. Simulations and random forest (model and parameter choice) were performed using DIYABC-RF framework based on diyabc-rf-linux-v1.1.51 and abcranger-linux-v1.16.67.^{34,92}