Genome-wide local ancestry and direct evidence for mitonuclear co-adaptation in African hybrid cattle populations (*Bos taurus/indicus*)

James A. Ward[1] | Gillian P. McHugo[1] | Michael J. Dover[1] | Thomas J. Hall[1] | Said Ismael Ng'ang'a[2, 3] | Tad S. Sonstegard[4] | Daniel G. Bradley[5] | Laurent A. F. Frantz[2, 3] | Michael Salter-Townshend[6] | David E. MacHugh[1, 7]

[1] Animal Genomics Laboratory, UCD School of Agriculture and Food Science, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland.

[2] Palaeogenomics Group, Department of Veterinary Sciences, Ludwig Maximilian University, Munich D-80539, Germany.

[3] School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom.

[4] Acceligen, 3388 Mike Collins Drive, Eagan, MN, 55121, USA.

[5] Smurfit Institute of Genetics, Trinity College Dublin, Dublin, D02 VF25, Ireland.

[6] UCD School of Mathematics and Statistics, University College Dublin, Dublin, D04 V1W8, Ireland.

[7] UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, D04 V1W8, Ireland.

Correspondence

David E. MacHugh, Animal Genomics Laboratory, UCD School of Agriculture and Food Science, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland.

Email: david.machugh@ucd.ie

Abstract

Domestic cattle have a key economic role in African societies, providing an important source of mobile wealth through supply of meat, milk, cowhide, fuel, transport, and traction. The phenotypic diversity of African cattle reflects adaptation to a wide range of agroecological conditions and complex patterns of admixture between the humpless *Bos taurus* (taurine) and humped *Bos indicus* (zebu) subspecies, which share a common ancestor 150-500 thousand years ago. Human migration and trade from Asia have left a peak of zebu nuclear ancestry in East Africa and most cattle populations across the continent have a hybrid genetic composition. Notwithstanding this, all African cattle possess taurine mitochondrial haplotypes, even populations with significant zebu nuclear ancestry. In this regard, the efficient functioning of the mitochondrion relies on a network of biochemical interactions between the products of 37 mitochondrial genes and more than one thousand nuclear genes; therefore, admixed African cattle represent ideal populations for evaluating mitonuclear interactions and mismatch between the nuclear and mitochondrial genomes. Using high-density SNP array data from 18 different cattle populations, including ten African admixed breeds, we find strong evidence for mitonuclear coevolution in hybrid African cattle with significant retention of *Bos taurus* alleles at mitochondrially-targeted nuclear genes, particularly those genes with products that directly interact with mtDNA-encoded protein subunits in OXPHOS and ribosomal complexes, or that have functions in mtDNA replication. We also show that subspecific local ancestry varies substantially across the genomes of admixed populations, with a marked signal of taurine ancestry at the major histocompatibility (MHC) gene cluster, which likely reflects adaptation to infectious disease challenges facing African livestock. Our results demonstrate that African admixed cattle represent an excellent comparative model for studying the phenotypic

47 consequences of mitonuclear mismatch and genomic introgression in humans and other large
48 mammals.

1 | INTRODUCTION

The mitochondrion is an endosymbiotic double membrane-bound eukaryotic organelle that evolved from an α-proteobacterial ancestor as consequence of the transition to an oxygen-rich atmosphere that began more than two billion years ago (Dyall *et al.* 2004; Lyons *et al.* 2014; Archibald 2015; Martin *et al.* 2015). Mitochondria harness energy from compounds such as sugars to produce, via oxidative phosphorylation (OXPHOS), adenosine triphosphate (ATP), the principal molecule used to store and transfer energy in the cell. Mitochondria are directly involved in a myriad of cellular and physiological processes, including, for example, calcium homeostasis, cell growth, cell differentiation, cell death, cellular waste management, neuronal function, and aging (McBride *et al.* 2006; Kasahara & Scorrano 2014; Spinelli & Haigis 2018). In addition, the key role of mitochondria in regulation of innate and adaptive immune responses has become more widely appreciated (Weinberg *et al.* 2015; Mills *et al.* 2017; Tiku *et al.* 2020).

Mitochondria have retained their own genome and in vertebrates this is a circular DNA molecule (mtDNA) of approximately 16 kb that encodes 37 genes and is exclusively transmitted by females to their offspring (Boore 1999; Zardoya 2020). Despite this, mitochondria also rely on more than one thousand co-adapted nuclear genes that encode proteins and protein subunits essential to mitochondrial function (Blier *et al.* 2001; Rand *et al.* 2004; Woodson & Chory 2008; Sloan *et al.* 2018). This intracellular mutualism is exemplified by the vertebrate OXPHOS system, which consists of five protein complexes, the four larger of which are chimeric—assembled using subunits encoded by the both the nuclear and mitochondrial genomes (Rand *et al.* 2004; Allen 2015; Isaac *et al.* 2018). The fundamental nature of these interactions and the importance of mitonuclear coevolution have been proposed to profoundly shape the evolution and ecology of eukaryotes (Burton *et al.* 2013; Wolff *et al.* 2014; Hill 2015; Sunnucks *et al.* 2017). Co-adapted sets of mitochondrial and nuclear genes have also been

73 proposed to underpin eukaryotic speciation and demarcation of species boundaries—the
74 mitonuclear compatibility species concept (MCSC)—thereby providing an explanatory
75 framework for using the mitochondrial cytochrome c oxidase I gene (*MT-CO1*) as an objective
76 and effective molecular barcode for taxonomic species delineation (Hill 2016; Hill 2017).

77 Functional mismatch between the mtDNA and nuclear genomes can be observed, for
78 example, in natural and human-generated hybrid populations, or in cybrids—cytoplasmic
79 hybrids—generated as cell lines, through somatic cell nuclear transfer (SCNT), or via hybrid
80 backcrosses (Burton *et al.* 2013; Chou & Leu 2015; Hill 2015; Sloan *et al.* 2017; Sunnucks *et*
81 *al.* 2017; Lechuga-Vieco *et al.* 2021). In these systems, mitonuclear incompatibilities between
82 distinct inter- and intraspecific evolutionary lineages can give rise to deleterious biochemical
83 effects associated with reduced efficacy of OXPHOS protein complexes (Blier *et al.* 2001;
84 Ellison & Burton 2006; Ellison *et al.* 2008; Ballard & Melvin 2010), which lead to lower ATP
85 production (McKenzie *et al.* 2003; McKenzie *et al.* 2004; Ellison & Burton 2008; Ellison *et al.*
86 2008) and increased levels of oxidative damage (Barreto & Burton 2013; Latorre-Pellicer *et al.*
87 2016; Du *et al.* 2017; Pichaud *et al.* 2019). At the organismal level, the physiological
88 consequences may be reduced fitness across a range of biological traits in vertebrates and
89 invertebrates, including those related to fecundity and reproductive success (Barreto & Burton
90 2013; Tourmente *et al.* 2017; Rank *et al.* 2020), developmental mortality and longevity (Ellison
91 *et al.* 2008; Latorre-Pellicer *et al.* 2016; Pichaud *et al.* 2019), metabolism and nutrition (Derr *et*
92 *al.* 2012; Latorre-Pellicer *et al.* 2016; Wang *et al.* 2017; Kong *et al.* 2020), endurance exercise
93 (Sujkowski *et al.* 2019; Spierer *et al.* 2021), and behaviour and neurobiology (Yu *et al.* 2009;
94 Gonzalez 2021). In F_3 hybrids of the marine copepod *Tigriopus californicus*, it was also
95 demonstrated that wild-type fitness could be restored through maternal backcrosses that

96 reunited intact nuclear genomes with co-evolved mitochondrial genomes (Ellison & Burton
97 2008)

98 Evidence for mitonuclear mismatches that affect biological fitness in natural vertebrate
99 populations has accumulated during the last decade and includes observations in the Iberian
100 hare (*Lepus granatensis*) subject to extensive mtDNA introgression from the mountain hare, *L.*
101 *timidus*, a northern boreal species (Seixas *et al.* 2018); the Italian sparrow (*Passer italiae*), a
102 hybrid between *P. domesticus* and *P. hispaniolensis* (Trier *et al.* 2014); hybrids of two diverged
103 Australian eastern yellow robin (*Eopsaltria australis*) populations (Morales *et al.* 2018);
104 Mediterranean chameleons (*Chamaeleo chamaeleon*) inhabiting an ancient Levantine marine
105 barrier (Bar-Yaacov *et al.* 2015); hybrids of two distinct lineages of long-toed salamander
106 (*Ambystoma macrodactylum*) in western Canada (Lee-Yaw *et al.* 2014); a hybrid population of
107 Atlantic killifish (*Fundulus heteroclitus*) with two divergent mtDNA haplotypes (Baris *et al.*
108 2017); and a hybrid fork-tongued goby (*Chaenogobius annularis*) population on the east coast
109 of Japan (Hirase *et al.* 2021).

110 The evidence for selection acting to conserve co-adapted mitonuclear interactions in
111 human populations is equivocal. Analysis of nuclear and mtDNA SNP data from the Human
112 Genome Diversity Project (HGDP) detected low levels of mitonuclear linkage disequilibria
113 (LD), even for nuclear genes encoding protein subunits that translocate to the mitochondrion
114 and form OXPHOS complexes with mtDNA-encoded proteins or that directly interact with
115 mitochondrial DNA or RNAs (Sloan *et al.* 2015). Contrary to this, however, a more recent study
116 of six distinct American admixed populations, demonstrated that mtDNA copy number is
117 inversely correlated with the level of discordance between nuclear and mtDNA ancestry (Zaidi
118 & Makova 2019). This observation was proposed to be a consequence of suboptimal regulation

119 of mtDNA replication in individuals with mismatched nuclear and mitochondrial genes. In
120 addition, there was evidence, at least in some populations, for a signature of mitonuclear
121 compatible ancestry enrichment at nuclear genes functionally associated with the
122 mitochondrion, particularly those encoding proteins that translocate to the organelle and
123 interact with mtDNA-encoded proteins in multi-subunit OXPHOS and ribosomal complexes.
124 A third study examined potential mitonuclear incompatibilities arising due to admixture with
125 Neanderthals (*Homo neanderthalensis*) and Denisovans (*Homo denisova*), which has left
126 signatures in modern human genomes of archaic hominin nuclear ancestry without mtDNA
127 introgression (Sharbrough *et al.* 2017). This analysis showed that Neandertal haplotypes for
128 nuclear gene loci associated with the mitochondrion have, as hypothesised, introgressed less
129 readily into both Asian and European human populations. This pattern, however, was not
130 observed for nuclear genomic introgression from Denisovans, likely a consequence of reduced
131 power to detect Denisovan haplotypes in populations exhibiting admixture with both archaic
132 hominins.

133 After humans, from a population genetics perspective, the most well-studied large
134 mammals are domestic cattle, which are derived from two morphologically distinct wild
135 subspecies, *Bos primigenius taurus* and *Bos primigenius indicus*, informally shortened to *B.*
136 *taurus* and *B. indicus* and colloquially referred to as "taurine" and "zebu" (or "indicine") cattle,
137 respectively (Zhang *et al.* 2020). Comprehensive genetic and archaeological evidence has
138 shown that taurine and zebu cattle were most likely domesticated separately: humpless taurine
139 cattle from the *B. p. primigenius* subspecies in the Fertile Crescent region of the Middle East
140 approximately 10 thousand years ago (kya), and humped zebu from the *B. p. namadicus*
141 subspecies two millennia later by the early Neolithic cultures located in present-day
142 Baluchistan, Pakistan (Bradley *et al.* 1998; Fuller 2006; Zeder 2011; Utsunomiya *et al.* 2019).

143 Recent estimates, based on genome-wide sequence data, place the evolutionary divergence time
144 between the two cattle subspecies at 150 to 500 kya (Chen *et al.* 2018; Wang *et al.* 2018; Wu
145 *et al.* 2018), although some analyses of complete mitogenomes have suggested a deeper split
146 (Hiendleder *et al.* 2008; Bibi 2013; Pramod *et al.* 2019).

147 The *B. taurus* and *B. indicus* lineages are interfertile and since domestication, and likely
148 before, there has been significant gene flow and admixture between humpless and humped
149 cattle across time and traversing multiple locations and environments in Eurasia and beyond
150 (Decker *et al.* 2014; Ginja *et al.* 2019; Verdugo *et al.* 2019; Gebrehiwot *et al.* 2020; Senczuk
151 *et al.* 2021). The largest and most well studied hybrid zone encompasses most of the African
152 continent and represents a complex mosaic of taurine and zebu ancestry that has been influenced
153 by a wide range of factors, including climate, physical geography, human migrations, and
154 infectious disease challenges (MacHugh *et al.* 1997; Hanotte *et al.* 2002; Kim *et al.* 2017; Kim
155 *et al.* 2020). However, even in the face of substantial zebu admixture, it has long been
156 understood that African cattle populations exclusively retain taurine mtDNA haplotypes, the
157 majority of which (within the T1 haplogroup) differ from T, T2, T3, and T4 haplotypes
158 observed in other taurine cattle populations (Loftus *et al.* 1994a; Loftus *et al.* 1994b; Bradley
159 *et al.* 1996; Troy *et al.* 2001). Therefore, as summarised in Figure 1, hybrid African cattle
160 represent excellent models for testing hypotheses concerning mitonuclear coevolution in large
161 mammals. In the present study, we leverage high-density SNP data and established
162 computational methods to characterise genome-wide local ancestry and systematically evaluate
163 mitonuclear interactions, coadaptation, and functional mismatch in ten genetically distinct
164 admixed African cattle populations.

165 ## 2 | MATERIALS AND METHODS

166 ### Animal sampling and genotyping

167 High-density genome-wide SNP array data sets (Illumina® BovineHD 777K BeadChip)
168 corresponding to a total of 605 animals were obtained from published studies (Bahbahani *et al.*
169 2015; Verdugo *et al.* 2019) and the Acceligen cattle genotyping database. For the present study,
170 new BovineHD 777K BeadChip SNP data were generated for 73 additional animals (50 Borgou
171 and 23 N'Dama) that were previously published by our group as part of microsatellite-based
172 surveys of cattle genetic diversity (MacHugh *et al.* 1997; Freeman *et al.* 2004). These new SNP
173 genotype data sets were generated by Weatherbys Scientific (Co. Kildare, Ireland) using
174 standard procedures for Illumina SNP array genotyping. In total, 18 different
175 breeds/populations were represented (Table 1), including two West African taurine breeds
176 (Muturu and N'Dama); two European taurine breeds (Holstein-Friesian and Jersey); ten West
177 and East African admixed zebu-taurine (Adamawa Gudali, Ankole, Borgou, Bunaji, East
178 African Shorthorn Zebu, Karamojong, Keteku, Nganda, Red Bororo, and Sokoto Gudali); and
179 four zebu breeds of South Asian origin (Gir, Hariana, Nelore, and Sahiwal). Table 1 also shows
180 the three- or four-letter codes used to designate each breed from this point onwards in this study.

181 ### SNP genotype data pre-processing, quality control, and filtering

182 BovineHD 777K SNP locations were remapped to the current bovine genome assembly
183 ARS-UCD1.2 (Rosen *et al.* 2020) and SNP genotype data were merged using PLINK v1.9
184 (Chang *et al.* 2015). Quality control (QC) of the combined SNP genotype data set was also
185 performed using PLINK v1.9 and autosomal SNPs with a call rate < 95% and a minor allele
186 frequency (MAF) of < 0.05 were filtered from the data.

187

188 **Principal component and population structure analyses**

189 Principal component analysis (PCA) of individual animal SNP genotype data for the
190 African taurine (MUTU and NDAG), the East African admixed (ADAG, ANKO, BORG,
191 BUNA, EASZ, KARA, KETE, NGAN, REDB, and SOKG) and two Asian indicine (GIR and
192 NELO) populations was performed using PLINK v1.9 and the results were plotted using
193 GGPLOT2 v3.3.3 (Wickham 2016) in the R v3.6.2 environment for statistical computation and
194 graphics (R Core Team 2019). The genetic structure of each population was also estimated
195 using FASTSTRUCTURE v1.0 with $K = 3$ modelled ancestries to determine mean African and
196 European *B. taurus* and Asian *B. indicus* contributions (Raj *et al.* 2014).

197 **Mitochondrial DNA haplogroup determination**

198 The BovineHD 777K BeadChip includes 346 SNPs located in the mitochondrial genome,
199 which can be used to construct haplotypes and catalogue and distinguish the mitochondrial
200 haplogroups characteristic of *B. taurus* and *B. indicus* cattle lineages. For this analysis, the
201 European JRSY and HOLS taurine breeds and the Indo-Pakistan HARI and SAHI Asian
202 indicine breeds were also included to ensure good representation of the *B. taurus* and *B. indicus*
203 mtDNA haplogroups—the 'T' and 'I' groups, respectively (Troy *et al.* 2001; Chen *et al.* 2010).
204 The mtDNA SNPs were filtered using PLINK v1.9 (Chang *et al.* 2015) such that SNPs with a
205 MAF of < 0.10, and a call rate of < 95% were removed. Individual animals with a genotype
206 missingness of > 95% were also removed. Following this, the most ancestry informative
207 mtDNA SNPs were identified using INFOCALC (Rosenberg *et al.* 2003; Rosenberg 2005), which
208 provides I_n, a general measure of the informativeness of a SNP for ancestry assignment. The 50
209 top ranked SNPs, based on I_n, were then used to generate mtDNA haplotypes with the
210 FASTPHASE v1.4 program (Scheet & Stephens 2006). Haplotype networks were constructed
211 using the POPART v1.7 package (Leigh & Bryant 2015).

212 **Local ancestry analysis of admixed populations**

213 Local ancestry across the bovine genome for each African admixed breed (ADAG,
214 ANKO, BORG, BUNA, EASZ, KARA, KETE, NGAN, REDB, and SOKG) was inferred using
215 MOSAIC v1.3.7 (Salter-Townshend & Myers 2019). The MOSAIC algorithm, unlike other
216 methods, does not require donor reference populations to be direct surrogates for the mixing
217 ancestral populations; it fits a two-layer Hidden Markov Model (HMM) that determines how
218 closely related each segment of chromosome in each admixed individual genome is to the
219 segments of chromosomes in individual genomes from potential donor populations and infers
220 a stochastic relationship between donor reference panels and mixing populations. While
221 determining local ancestry along each chromosome, MOSAIC also infers the number of
222 generations since the admixture process started for a particular population. The potential donor
223 populations used for the MOSAIC local ancestry analysis were the two West African *B. taurus*
224 breeds (MUTU and NDAG) and two of the Asian *B. indicus* breeds (GIR and NELO). The
225 MOSAIC algorithm requires phased haplotypes and a recombination rate map; therefore,
226 SHAPEIT v2 (r900) (Delaneau *et al.* 2012) was used to generate phased haplotypes and a
227 published cattle recombination map was employed (Ma *et al.* 2015). Functional enrichment
228 analysis of genes proximal to African taurine and Asian zebu local ancestry peaks was
229 performed using Ingenuity® Pathway Analysis—IPA® (version 1.1, summer 2021 release;
230 Qiagen, Redwood City, CA, USA) (Kramer *et al.* 2014).

231 **Detection of subchromosomal deviations in taurine and indicine genomic ancestry**

232 To determine if there was significant retention of *B. taurus* nuclear genes that encode
233 mitochondrially targeted proteins (*N-mito* genes) in African admixed cattle, we used an
234 approach modified from previously published surveys of mitonuclear incompatibilities in
235 modern admixed human populations (Sloan *et al.* 2015; Zaidi & Makova 2019) and because of

236 ancient gene flow from archaic hominins (*H. neanderthalensis* and *H. denisova*) (Sharbrough
237 *et al.* 2017). Firstly, the MITOCARTA 2.0 database resource (Calvo *et al.* 2016) was used to
238 obtain an inventory of genes that produce the nuclear-encoded component of the mammalian
239 mitochondrial proteome, i.e., proteins with experimental evidence for localisation in the
240 mitochondrion. Following this, the ENSEMBL BIOMART tool (Yates *et al.* 2020) was used to
241 generate a list of 1158 bovine N-mito genes, which was classified into two functional subsets
242 as defined by Sloan *et al.* (2015) and also used by Sharbrough *et al.* (2017) and Zaidi and
243 Makova (2019). These subsets were denoted as 1) high-confidence "*high-mito*" genes (HMG)
244 encoding proteins that directly interact with mtDNA-encoded protein subunits in OXPHOS and
245 ribosomal complexes, or that have functions in mtDNA replication (138 genes); and 2) lower
246 confidence "*low-mito*" genes (LMG), which encode proteins that localise to the mitochondrion,
247 but are not classified as part of the high-mito subset (701 genes). Finally, a third group of "*non-*
248 *mito*" genes (NMG) was generated, which includes the bulk of the mammalian proteome that
249 does not localise to the mitochondrion (17,834 genes). Table S1 provides further detail for the
250 functional gene subsets used to detect evidence for mitonuclear incompatibilities in African
251 admixed cattle populations.

252 The local ancestry estimates generated using MOSAIC for each SNP across the genome
253 were catalogued and the BEDTOOLS v2.18 software suite (Quinlan & Hall 2010) was then used
254 to intersect these SNPs with windows spanning 2.5 Mb upstream and downstream of genes
255 within each of the three functional gene subsets. Following this, and as described by Zaidi and
256 Makova (2019), for each of the three subsets an unweighted block bootstrap approach was used
257 to generate distributions of local ancestry deviation towards more or less *B. taurus* ancestry.
258 The first step in this methodology is subtraction of the mean ancestry fraction across the local
259 ancestry estimate for each SNP (the *expectation*), which produces the deviation in local ancestry

260 at each SNP locus. For each functional gene subset, the number of windows sampled with
261 replacement was the same as the number of HMG subset genes (n = 138). In each case, the
262 mean ancestry deviation was estimated and then averaged across all windows. Bootstrap
263 resampling (1000 replicates) was used to generate a distribution of mean deviations in local
264 ancestry for each of the three functional gene subsets. The overall statistical significance of
265 deviations from expected local ancestry (using the background NMG subset) was then assessed
266 using Welch's *t*-test in R v3.6.2 (R Core Team 2019).

267 **Transcriptomics and Cattle Gene Atlas analysis**

268 Following the methodology described by Sharbrough *et al.* (2017), under the assumption
269 that highly expressed genes are more functionally constrained, we evaluated if the local ancestry
270 estimates for HMG, LMG or NMG subsets were correlated with gene expression across a wide
271 range of bovine tissues. RNA-seq transcriptomics data in the form of fragments per kilobase
272 per million mapped reads (FPKM) was obtained from the Cattle Gene Atlas database
273 (http://cattlegeneatlas.roslin.ed.ac.uk; Fang *et al.* 2020), which consists of 723 RNA-seq data
274 sets across 84 bovine tissues. For each tissue, the expression values were averaged, resulting in
275 84 mean FPKM values for all known bovine genes. Following this, a K_{ij} matrix was generated
276 such that each cell contained mean FPKM values organised into columns (i) representing each
277 tissue and rows (j) corresponding to the 130 HMG, 638 LMG, and 14,758 NMG genes that
278 could be mapped to the Cattle Gene Atlas. An expression threshold was created such that each
279 gene was classified as detectably expressed for a particular tissue if its mean FPKM value for
280 that tissue was >7.8 (corresponds to 50% of the global mean FPKM value of 15.6) across the
281 complete data set (all genes and all tissues). A tissue expression metric (TEM) for
282 systematically examining correlated expression for the HMG, LMG and NMG gene subsets
283 was then created as follows: the mean expression of a gene across all 84 tissues was multiplied

284 by the number of cases where mean expression of a gene was greater than the global FKKM
285 mean (15.6) in any tissue. For example, if a gene exhibited a mean expression value of 10.0
286 across all 84 samples and its expression was greater than 15.6 in five of these tissues, then the
287 TEM for that gene would be 50.0. The TEM values for each gene in the HMG, LMG, and NMG
288 subsets were then plotted against the mean African taurine local ancestry scores for each gene
289 locus across the 10 African admixed breeds and Pearson correlation coefficients were estimated
290 for the LMG, HMG, and NMG subsets. The statistical analyses and data graphing for the TEM
291 and African taurine local ancestry analysis were performed using R v3.6.2 (R Core Team 2019)
292 and GGPLOT2 v3.3.3 (Wickham 2016).

293 3 | RESULTS

294 Genomic structure of African admixed cattle

295 Following filtering and QC there were 562,635 SNPs and 605 individual animals retained
296 for subsequent analyses. Figure 2 shows a PCA plot generated from SNP genotype data for
297 individual African taurine and admixed animals and with the GIR and NELO breeds
298 representing Asian *B. indicus* cattle and the HOLS and JERS representing the European *B.*
299 *taurus* group. The first principal component (PC1), which accounts for 58.4% of the variation,
300 represents the primary split between the *B. taurus* and *B. indicus* lineages and PC2 (17.9% of
301 the variation) differentiates the African and European taurine populations. The position of
302 individual admixed animals is determined by the relative proportions of African taurine and
303 zebu ancestry, with additional European gene flow evident for some East African cattle—
304 particularly within the EASZ and NGAN breeds. The results of the genetic structure analysis
305 with FASTSTRUCTURE and an inferred number of clusters of $K = 3$ are shown in Figure 3 and
306 illustrates the varying patterns of taurine and zebu ancestry across the different East and West
307 African admixed cattle populations. For the most part, the levels of admixture are relatively

308 uniform for animals within each breed; however, there are some cattle that deviate noticeably
309 from the pattern, particularly in the EASZ and NGAN where recent European taurine gene flow
310 is most evident. Table 2 provides additional detail on the taurine/zebu admixture results
311 generated using the FASTSTRUCTURE program. As generally observed in previous studies of
312 *B. taurus*/*B. indicus* admixture (MacHugh *et al.* 1997; Hanotte *et al.* 2002; Kim *et al.* 2017;
313 Gebrehiwot *et al.* 2020; Kim *et al.* 2020), the four East African admixed populations exhibit a
314 greater level of average zebu ancestry (mean = 0.653) than the six West African admixed
315 populations (mean = 0.612).

316 **_Bos taurus_ and _Bos indicus_ mitochondrial DNA haplotype network analysis: admixed**
317 **African cattle possess taurine mitochondrial genomes**

318 After filtering of the 346 mtDNA SNPs on the BovineHD 777K BeadChip and
319 identification of the most ancestry-informative SNPs, a phylogenetic haplotype network was
320 generated using 39 mtDNA SNPs and a total of 491 cattle (47 African taurine, 82 European
321 taurine, 156 East African admixed, 136 West African admixed, and 70 Asian zebu). Figure 4
322 shows this network and demonstrates that all 339 African taurine and admixed cattle surveyed
323 here possess the taurine mitochondrial genome. This observation, which uses, to the best of our
324 knowledge, the largest sample of African mtDNA genotypes, agrees with previous studies of
325 mtDNA diversity in African cattle (Loftus *et al.* 1994a; Loftus *et al.* 1994b; Bradley *et al.* 1996;
326 Troy *et al.* 2001; Dadi *et al.* 2009; Bonfiglio *et al.* 2012). It has been hypothesised that male-
327 mediated gene flow accounts for this pattern of mitonuclear discordance, which resulted due to
328 the geographical barriers between Asia and Africa and historical preferences for trade and
329 transport of *B. indicus* bulls to disseminate desirable traits (Bradley *et al.* 1994; Loftus *et al.*
330 1994a). Importantly, from the perspective of the present study, these results unambiguously
331 demonstrate that East and West admixed African cattle breeds retain taurine mitochondrial

genomes, making them ideal populations to investigate mitonuclear interactions and co-adaptation.

We also observed taurine mtDNA haplotypes in animals from the morphologically Asian zebu GIR and NELO breeds; this mitochondrial discontinuity reflects their early origins as hybrid taurine-indicine founder populations that were backcrossed to *B. indicus* bulls to recover South Asian zebu traits advantageous for the tropical production environments of Brazil (Dani *et al.* 2008). In this regard, present-day GIR and NELO cattle exhibit very low levels of autosomal taurine admixture, less than 1% when estimated using genome-wide high-density SNP data (O'Brien *et al.* 2015).

Genome-wide local ancestry analysis and deviations in African taurine local ancestry for functional mitochondrial gene subsets

Three coancestry curves (two-way African taurine and Asian zebu admixture) generated using MOSAIC are shown in Figure 5 for each of the East and West African cattle populations. The ancestry labels on top of each subplot correspond to the closest African taurine and Asian zebu donor breed as measured by the Weir and Cockerham $\hat{F}_{\mathrm{st}}$ estimator of population differentiation (Weir & Cockerham 1984). In each case, this was the NDAG *B. taurus* and GIR *B. indicus* breeds and the full set of $\hat{F}_{\mathrm{st}}$ values for each of the admixed African breeds is provided in Table S2. In addition, the coancestry curve analyses provides estimates of the time in generations since admixture commenced and these were converted to years before present (BP) using the NDAG:GIR coancestry curve values (middle subplots) and a generation interval range of between 4 and 7 years for managed domestic cattle populations (Mészáros *et al.* 2015; Abin *et al.* 2016). The estimated historical calendar range (common era – CE) for the start of the admixture process in each of the African admixed breeds is shown in Table 2. A recent study using whole-genome sequence (WGS) data, albeit with lower samples sizes, also

356 estimated the time in generations since the beginning of admixture for one of the East African
357 breeds used for the present study (ANKO). Kim *et al.* (2020) used WGS data from 10 ANKO
358 cattle to obtain linkage disequilibrium (LD) and haplotype-sharing coancestry based estimates
359 of 147.3 and 113.6 generations, respectively—more than twice the result obtained here with
360 SNP array data (54.1 generations; 1641 – 1804 CE) and pushing back the historical calendar
361 range for the start of admixture in the ANKO population to between 989 and 1566 CE.

362 For each of the East and West African admixed populations, Asian zebu, and African
363 taurine SNP-by-SNP local ancestry values were estimated for each of the 29 bovine autosomes
364 using MOSAIC. As an example of this process and the resulting output, Figure 6 shows Asian
365 zebu and African taurine local ancestry values plotted for all SNPs across bovine chromosome
366 23 (BTA23). Individual SNP local ancestry values were then averaged on a gene-by-gene basis
367 to determine African taurine local ancestry values for individual genes. These can be used to
368 systematically evaluate deviations in local ancestry for functional mitochondrial gene subsets
369 (HMG and LMG) that could indicate mitonuclear incompatibilities and coadaptation in
370 admixed African cattle. African taurine local ancestry deviations for the three different
371 functional gene subsets (HMG, LMG and NMG) are displayed as violin plots in Figure 7a. A
372 positive deviation for a breed/subset indicates that a particular functional gene subset exhibits
373 retention of taurine gene haplotypes. The HMG, LMG, and NMG African taurine local ancestry
374 deviation distributions for each breed were evaluated for normality using the Shapiro-Wilk test.
375 Following this, comparative statistical analysis of the local ancestry deviation distributions was
376 performed, and the results are shown in Table 3. This analysis demonstrated that all 10 African
377 admixed cattle breeds exhibited statistically significant and consistently increased African
378 taurine local ancestry deviation distributions for the HMG subset compared to the background
379 NMG subset. In addition, three out of four of the East African (ANKO, KARA, and NGAN)

380 and two out of six of the West African (ADAG and KETE) admixed breeds showed a
381 statistically significant increase in African taurine local ancestry deviation distributions for the
382 LMG subset compared to the background NMG subset. Finally, except for the ADAG breed,
383 all populations showed significantly increased African taurine local ancestry in the HMG subset
384 compared to the LMG subset.

385 A direct statistical comparison of the HMG African taurine local ancestry distributions
386 for the combined East and West African admixed cattle groups (Figure 7b) demonstrated that
387 the East African admixed group exhibit a significantly higher local ancestry deviation for the
388 HMG subset of nuclear genes that encode mitochondrially targeted proteins. We also tested the
389 hypothesis that mean African taurine ancestry deviation is influenced by the number of
390 generations since the start of admixture, assuming a single pulse of admixture as modelled by
391 the MOSAIC software (Salter-Townshend & Myers 2019). The results of this analysis are
392 shown in Figure 7c where a statistically significant negative correlation ($R = -0.678$; $P = 0.031$)
393 was observed between the mean African taurine local ancestry deviation values for the HMG
394 subset and the generations since the start of admixture. This result is somewhat counterintuitive
395 because it might be expected that African taurine local ancestry deviations for the HMG subset
396 would increase with time as selection acts to purge mitonuclear incompatibilities and maintain
397 co-adaption of the mitochondrial and nuclear genomes. However, several cofounding factors
398 may contribute to this result, including varying levels of genome-wide taurine-zebu ancestry
399 and the presence of European taurine ancestry in certain populations (see Figure 3 and Table
400 3). In addition, the temporal pattern and mode of admixture may not correspond to a simple
401 scenario of a single admixture event or pulse, as modelled by MOSAIC. However, the most
402 persuasive biological explanation is provided by genetic recombination acting across multiple
403 generations, which over time would erode and break down contiguous taurine haplotypes at the

404 HMG subset gene loci. Consequently, as HMG gene loci in admixed populations recombine,
405 the mean African taurine local ancestry scores averaged across multiple SNPs would decrease
406 and this would lead to dissipation of apparent positive African taurine local ancestry deviation
407 for the HMG subset compared to the background NMG subset. The gene-by-gene local ancestry
408 approach that we employ here to detect mitonuclear incompatibilities may only function
409 accurately for relatively recently admixed populations, such as the human populations with
410 African, European and Native America ancestry examined by Zaidi and Makova (2019). In
411 other words, the underlying signal of mitonuclear compatibility and co-adaption is likely
412 stronger in long established admixed African cattle breeds (*i.e.*, ADAG, BUNA, and SOKG)
413 compared to more recently admixed populations; however, it is masked by recombination and
414 may require higher-resolution local ancestry methods to detect it; for example, using larger
415 sample sizes with WGS data, and an exon-by-exon, rather than a gene-by-gene approach.

416 Taken together and bearing in mind that ten different independent admixed cattle
417 populations were examined, the results presented here provide compelling evidence supporting
418 the hypothesis that bovine nuclear genes encoding proteins directly (the HMG subset), or
419 indirectly (the LMG subset) involved with the mitochondrion have been influenced by selective
420 processes favouring mitonuclear compatibility and co-adaptation in admixed African cattle
421 populations. This is particularly evident for the HMG subset where all ten breeds exhibited
422 strong signals of African taurine local ancestry enrichment at nuclear genes that encode
423 mitochondrially targeted proteins.

424 Genome- and chromosome-wide patterns of ancestry across all African admixed 425 cattle populations

426 To illustrate genome-wide patterns of taurine and zebu admixture across all ten African
427 admixed cattle populations, Figure 8a shows mean African taurine local ancestry for each of

428 the bovine autosomes (BTA1–BTA29) across all ten African admixed breeds. The three highest
429 taurine ancestry peaks were observed on BTA7, BTA12, and BTA23 and the three lowest
430 taurine ancestry troughs (zebu peaks) were observed on BTA3, BTA5, and BTA6. Figure 8b
431 shows mean African taurine local ancestry plotted across three of these chromosomes (BTA5,
432 BTA12, and BTA23). The highest genome-wide signal of mean African taurine ancestry across
433 the ten admixed breeds was observed as a broad plateau encompassing the bovine major
434 histocompatibility complex (MHC) gene cluster on BTA23 (indicated as a purple rectangle on
435 Figure 8b), likely reflecting the key role of the MHC gene cluster in innate and adaptive
436 immunity and responses to pathogens and parasites (Behl *et al.* 2012; Ellis & Hammond 2014).

437 To further examine patterns of introgression across the genome, the top 1% of the mean
438 African taurine and Asian zebu local ancestry peaks across the 10 admixed breeds were
439 identified separately. Following this, genes 2.5 Mb up- and downstream of each peak were
440 catalogued and used for functional enrichment analysis with the IPA software tool, using the
441 default settings. Table S3 shows these input gene sets, which contained 858 and 317 genes for
442 the top 1% of African taurine and Asian zebu local ancestry peaks, respectively. A false
443 discovery rate (FDR) method Benjamini and Hochberg (1995) was used to set a $P_{adj.}$ threshold
444 of 0.05 for statistical significance of enrichment for the canonical pathways and physiological
445 system development and function categories in IPA. The top five enriched canonical pathways
446 for genes with high African taurine ancestry were *Differential Regulation of Cytokine*
447 *Production in Macrophages and T Helper Cells by IL-17A and IL-17F*, *Differential Regulation*
448 *of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F*, *Glutathione-*
449 *mediated Detoxification*, *GABA Receptor Signalling*, and *Gustation Pathway*. The top five
450 enriched physiological system development and function categories genes with high African
451 taurine ancestry were *Organismal Survival*, *Immune Cell Trafficking*, *Cardiovascular System*

452 *Development and Function*, *Haematological System Development and Function*, and *Tissue*
453 *Morphology*. The top five enriched canonical pathways for genes with high Asian zebu ancestry
454 were *NRF2-mediated Oxidative Stress Response*, *Mitotic Roles of Polo-Like Kinase*, -
455 *tocopherol Degradation*, *Methylglyoxal Degradation III*, and *Role of Osteoblasts, Osteoclasts*
456 *and Chondrocytes in Rheumatoid Arthritis*. The top five enriched physiological system
457 development and function categories for genes with high Asian zebu ancestry were *Connective*
458 *Tissue Development and Function*, *Skeletal and Muscular System Development and Function*,
459 *Nervous System Development and Function*, *Organ Morphology*, and *Organismal*
460 *Development*.

461 **Gene expression analysis of functional HMG, LMG, and NMG subsets**

462 Figure S1 show tissue expression metric (TEM) values obtained from the Cattle Gene
463 Atlas (Fang *et al.* 2020) plotted against mean gene locus taurine local ancestry values in 10
464 African admixed breeds for the HMG (130 genes), LMG (638 genes), and NMG (14,758 genes)
465 functional subsets (see Table S4 for the data used to generate these plots). Correlation analysis
466 demonstrated that no significant association was observed for the HMG subset TEM values and
467 African taurine local ancestry (Figure S1a; $P = 0.730$). Conversely, as shown in Figures S1b
468 and S1c, very weak but significant negative correlations between TEM and African taurine
469 local ancestry was observed for both the LMG subset ($R = -0.090$; $P = 0.024$) and the much
470 larger NMG subset ($R = -0.021$; $P = 0.010$). In other words, more highly expressed genes in
471 African admixed breeds tend to exhibit greater levels of *B. indicus* ancestry. If we assume that
472 zebu cattle correspond to the introgressive source population then this is the opposite pattern to
473 that observed for archaic human gene flow, where human genes with higher levels of
474 Neanderthal ancestry tend to exhibit lower average multi-tissue expression levels (Sharbrough
475 *et al.* 2017). It is important to keep in mind, however, that as was the case for Neanderthal gene

476 flow, these effects account for a trivial proportion of the variance ($R^2 < 0.01$ for both the LMG
477 and NMG data sets). In addition, direct comparison of the human and cattle results is
478 confounded by the much lower levels of Neanderthal ancestry in modern human populations
479 compared to zebu admixture in African cattle and the much longer time depth in generations
480 since *H. neanderthalensis* and *H. sapiens* interbred (Bergström *et al.* 2021).

481 **4 | DISCUSSION**

482 During the last three decades, increasingly sophisticated molecular tools and statistical
483 methods have been used to systematically survey genomic variation in African cattle
484 populations (Loftus *et al.* 1994a; MacHugh *et al.* 1997; Hanotte *et al.* 2002; Decker *et al.* 2014;
485 Kim *et al.* 2017; Pitt *et al.* 2019; Gebrehiwot *et al.* 2020; Kim *et al.* 2020). These studies have
486 revealed a complex demographic and evolutionary history fundamentally shaped by gene flow
487 and admixture between the original substrate of locally adapted humpless taurine animals (*B.*
488 *taurus*) and exogenous humped Asian zebu (*B. indicus*) introduced through human migration
489 and trade (Epstein & Mason 1971; Gifford-Gonzalez & Hanotte 2011). Intriguingly, unlike
490 other regions of the world such as Southwest Asia and East Asia where taurine and zebu cattle
491 populations also overlap and interbreed (Edwards *et al.* 2007; Chen *et al.* 2018; Xia *et al.* 2019),
492 all African cattle examined to-date retain *B. taurus* mitochondrial haplotypes (Figure 4; Loftus
493 *et al.* 1994a; Bradley *et al.* 1996; Troy *et al.* 2001; Álvarez *et al.* 2017; Mauki *et al.* 2021). In
494 this respect, admixed cattle, and animals with predominantly zebu morphology in Africa
495 represent an example of "massively discordant mitochondrial introgression (MDMI)", a term
496 introduced by Bonnet *et al.* (2017). This mitochondrial discordance is likely a result of male-
497 mediated gene flow and genetic drift through preferential dissemination of *B. indicus* genetic
498 material by a relatively small number of Asian zebu cattle, most of which were bulls (Bradley
499 *et al.* 1994; Loftus *et al.* 1994a). The marked evolutionary divergence between the *B. taurus*

500 and *B. indicus* nuclear genomes and the MDMI observed in admixed African cattle therefore
501 provide a unique testbed for exploring genomic introgression and mitonuclear interactions in a
502 numerous large mammalian species spanning diverse physical geographies and agroecological
503 zones. In addition, the presence of many discrete admixed cattle breeds across the continent
504 with varying levels of *B. indicus* ancestry (Figure 2a and Figure 3) means that incompatibilities
505 and coevolution between nuclear and mitochondrial genes can be detected and evaluated in
506 multiple independent populations.

507 **Evidence for mitonuclear incompatibilities in admixed African cattle breeds**

508 Although previous studies have examined subchromosomal admixture and local ancestry
509 in hybrid *B. taurus*/*B. indicus* animals (McTavish & Hillis 2014; Chen *et al.* 2018; Koufariotis
510 *et al.* 2018; Barbato *et al.* 2020), we present here the first evidence for mitonuclear
511 incompatibilities in admixed cattle populations. Using a high-density SNP genotyping array,
512 ten different breeds were examined with genome-wide zebu ancestries ranging between 0.370
513 (BORG) and 0.740 (KARA), and estimated dates for the start of admixture in each population
514 extending from the 14th to the 20th century (Table 2). A consistent pattern of mitonuclear
515 disequilibria was observed for the functional HMG subset in admixed African cattle (Figure 7a
516 and Table 3): African taurine local ancestry was uniformly higher for nuclear genes encoding
517 proteins that directly engage with mitochondrial-encoded gene products to form multi-subunit
518 complexes, or that directly interact with mitochondrial DNA or RNAs; this subset encompasses
519 genes that encode OXPHOS subunits, ribosomal proteins, tRNA synthetases, and DNA and
520 RNA polymerases. Although the source population divergence is substantially less in admixed
521 humans, these results are comparable to those obtained by Zaidi and Makova (2019) that
522 support the hypothesis that selection in admixed human populations has acted against
523 mitonuclear incompatibility. They observed significant enrichment of sub-Saharan African

524 ancestry for HMG subset genes in an African American population with sub-Saharan African
525 and European nuclear ancestry and predominantly sub-Saharan African mtDNA haplotypes.
526 They also observed significant enrichment of Native American ancestry at HMG subset genes
527 in a Puerto Rican population with Native American and European nuclear ancestry and
528 predominantly Native American mtDNA haplotypes. It is also noteworthy that the mean
529 ancestry deviation observed for the HMG gene subset in admixed African cattle populations
530 was approximately 2- to 4-fold larger than that observed in the African American and Puerto
531 Rican populations by Zaidi and Makova (2019). This result may reflect stronger selection
532 against mitonuclear incompatibility in admixed cattle that exhibit substantial *B. indicus* nuclear
533 ancestry in parallel with exclusive retention of *B. taurus* mtDNA haplotypes.

534 For certain admixed African cattle populations (ANKO, KARA, NGAN, ADAG, and
535 KETE) we also observed statistically significant elevated African taurine local ancestry for the
536 LMG gene subset, and except for the ADAG breed, African taurine local ancestry was
537 consistently higher in the HMG subset compared to the LMG subset (Figure 7a and Table 3).
538 A geographic and temporal aspect to potential mitonuclear incompatibility in admixed African
539 cattle populations was also observed. The East African admixed cattle group exhibit a
540 statistically significant larger mean African taurine local ancestry deviation compared to the
541 West African admixed cattle group ($P = 2.2 \times 10^{-16}$; Figure 7b). In addition, somewhat
542 counterintuitively, we observed a decline in African taurine local ancestry deviation for the
543 HMG subset in conjunction with increasing time depth to the onset of population admixture
544 (Figure 7c). The most parsimonious explanation for this pattern is breakdown of African taurine
545 haplotypes at HMG loci through genetic recombination acting across multiple generations.

546 The functional HMG and LMG subsets containing 138 and 701 genes, respectively were
547 identified, and catalogued in the present study for the purpose of evaluating mitonuclear
548 incompatibilities in admixed African cattle populations. It is also instructive, however, to
549 examine these genes in the context of recently published high-resolution surveys of African
550 cattle genomic diversity and signatures of selection. For example, the aspartyl-tRNA synthetase
551 2, mitochondrial gene (*DARS2*), an HMG subset gene on BTA16, is in the region encompassed
552 by selective sweeps detected separately in the EASZ breed and a composite sample of East
553 African zebu cattle (Bahbahani *et al.* 2017; Taye *et al.* 2018). In addition, the mitochondrial
554 ribosomal protein S33 gene (*MRPS33*), another HMG subset gene, was detected within a
555 positively selected region on BTA4 when African cattle were compared to commercial
556 European and Asian breeds (Kim *et al.* 2017) and in analyses of selective sweeps focused on
557 the evolution of thermotolerance in African cattle populations (Taye *et al.* 2017). Table S5
558 provides further information on HMG and LMG subset genes that have been detected in studies
559 of genomic selective sweeps in African cattle populations.

560 **Patterns of local ancestry across African cattle genomes**

561 The patterns of mean African taurine and Asian zebu local ancestry varied across the
562 genome for the 10 African admixed cattle populations, with notable taurine and zebu peaks on
563 BTA7, BTA12, and BTA23, and BTA3, BTA5, and BTA6, respectively. Analysis of genes
564 proximal to the top 1% of taurine and zebu local ancestry peaks demonstrated that genes with
565 high African taurine ancestry were substantial enriched for pathways and physiological systems
566 associated with immunobiology and antimicrobial host defence. Conversely, the genes with
567 high Asian zebu ancestry were overrepresented in pathways and physiological systems
568 associated with organismal growth and development. These functional differences may reflect
569 natural and human-mediated selection for retention of infectious disease resistance or resilience

570 traits and incorporation of desirable production traits in hybrid African *B. taurus*/*B. indicus*
571 populations (Mwai *et al.* 2015; Marshall *et al.* 2019).

572 A particularly notable outcome of the genome-wide analyses of local ancestry was the
573 observation of the highest signal of African taurine ancestry encompassing the bovine MHC
574 gene cluster region on BTA23 (Figure 8b). In this regard, a comparable admixture signal has
575 been observed in a composite North African and Middle Eastern human sample, which
576 exhibited elevated sub-Saharan ancestry around the MHC gene cluster on HSA6 (Salter-
577 Townshend & Myers 2019). Enrichment of sub-Saharan African MHC haplotypes has also been
578 detected in a range of admixed Latin American populations (Norris *et al.* 2020). It is also
579 noteworthy that previous analyses of genomic microevolution in African cattle populations
580 have detected selective sweeps at the MHC gene cluster (Kim *et al.* 2017; Taye *et al.* 2018;
581 Tijjani *et al.* 2019). The bovine MHC region, in common with humans and other mammals,
582 contains multiple individual genes and gene clusters associated with innate and adaptive
583 immune responses to pathogens and parasites, including the tumour necrosis factor (*TNF*),
584 lymphotoxin alpha (*LTA*), and lymphotoxin beta (*LTB*) cytokine genes; the highly polymorphic
585 genes encoding MHC class I and II molecules responsible for intracellular and extracellular
586 antigen presentation, respectively (e.g., *BOLA*, *BOLA-NC1*, *BOLA-DMA*, *BOLA-DOA*, *BOLA-*
587 *DQA1*, *BOLA-DQB*, and *BOLA-DRA*); and the MHC class I related protein genes (*MIC1*, *MIC2*,
588 and *MIC3*) (Behl *et al.* 2012; Ellis & Hammond 2014; Maccari *et al.* 2017). A key process
589 driving evolution of the vertebrate MHC gene cluster region is pathogen- and parasite-mediated
590 selection caused by infectious disease, particularly balancing selection to maintain allelic
591 diversity at MHC class I and class II genes (Ebert & Fields 2020; Radwan *et al.* 2020).
592 Consequently, a marked signal of retention of taurine MHC haplotypes in admixed breeds may
593 reflect adaptation to significant local disease challenges such as African animal trypanosomiasis

594 (AAT) caused by a range of *Trypanosoma* species, theileriosis (East Coast fever) caused by
595 *Theileria parva*, cowdriosis caused by *Ehrlichia ruminantium*, and Rift Valley fever (RVF)
596 caused by the RVF virus (Ristic & McIntyre 1981; Seifert 1996). It is also important to note
597 that patterns of genomic diversity derived from SNP array data for the bovine MHC gene cluster
598 region could be confounded by ascertainment bias or inaccurate or incomplete genome
599 annotation. However, ascertainment bias is unlikely to influence our results because the
600 Illumina® BovineHD 777K BeadChip was designed using a diverse panel representative of
601 multiple *B. taurus*, *B. indicus* and admixed *B. taurus/indicus* populations (Rincon *et al.* 2011;
602 Ventura *et al.* 2020). In addition, the 28 SNPs on the array that capture variation at the bovine
603 MHC gene cluster are located in the relatively stable and well-annotated MHC Class Ib region
604 (J.A. Hammond, *pers. comm.*).

605 **Conclusions and outlook**

606 Cattle breeding programmes in Africa are currently poised to leverage genomic selection
607 as a leapfrog technology to bypass conventional genetic evaluation and improvement of
608 production, health, and welfare traits (Ibeagha-Awemu *et al.* 2019; Marshall *et al.* 2019; Mrode
609 *et al.* 2019). Genome-enabled breeding of livestock can also contribute to mitigation of methane
610 emissions and adaptation to the consequences of climate change (Hayes *et al.* 2013; McManus
611 *et al.* 2020); in this regard, the genomic composition of African cattle populations may be
612 further impacted by the superior thermotolerance of zebu cattle (Hansen 2004; Dikmen *et al.*
613 2018). The fundamental role of the mitochondrion in the biochemistry and physiology of all
614 these traits, implies that the mitonuclear ecology framework (Hill 2015) will become
615 increasingly important for dissecting human-mediated and natural microevolution of livestock.
616 This will be particularly relevant for African admixed cattle, which as we demonstrate here,
617 exhibit strong signatures of mitonuclear coevolution. Finally, these populations can serve as

618 comparative model systems for understanding the phenotypic consequences of mitonuclear
619 interactions and adaptive and maladaptive genomic introgression in other mammals, including
620 humans.

621

622 **ACKNOWLEDGEMENTS**

623 This research was supported by Science Foundation Ireland (SFI) and
624 Acceligen/Recombinetics Inc. through the SFI CRT in Genomics Data Science under grant no.
625 18/CRT/6214. Additional support was provided by an SFI Investigator Award to D.E.M (grant
626 no. SFI/15/IA/3154).

627 **AUTHOR CONTRIBUTIONS**

628 J.A.W., M.J.D., T.S.S., D.G.B., L.A.F., M.S-T., and D.E.M. conceived and designed the
629 study; J.A.W., T.S.S., D.G.B., and D.E.M coordinated data and sample acquisition; J.A.W. and
630 G.P.M. conducted the analyses with help from T.J.H, S.I.I.N, L.A.F., and M.S-T.; and J.A.W.
631 and D.E.M. wrote the manuscript with critical input and revisions from all authors. All authors
632 gave final approval for publication and agree to be held accountable for the work performed
633 therein.

634 **CONFLICT OF INTEREST**

635 The authors declare no conflict of interest.

636 **DATA AVAILABILITY STATEMENT**

637 All Illumina® BovineHD 777K BeadChip SNP data used for this study is available from
638 the Dryad Digital Repository (https://datadryad.org).

639

640 **REFERENCES**

641 Abin S., Theron H.E. & van Marle-Koster E. (2016) Population structure and genetic trends for
642 indigenous African beef cattle breeds in South Africa. *S. Afr. J. Anim. Sci.* **46**, 152-6.

643 Allen J.F. (2015) Why chloroplasts and mitochondria retain their own genomes and genetic
644 systems: Colocation for redox regulation of gene expression. *Proc. Natl. Acad. Sci. U.*
645 *S. A.* **112**, 10231-8.

646 Álvarez I., Pérez-Pardal L., Traoré A., Koudandé D.O., Fernández I., Soudré A., Diarra S.,
647 Sanou M., Boussini H. & Goyache F. (2017) Differences in genetic structure assessed
648 using Y-chromosome and mitochondrial DNA markers do not shape the contributions
649 to diversity in African sires. *J. Anim. Breed. Genet.* **134**, 393-404.

650 Archibald J.M. (2015) Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* **25**, R911-21.

651 Bahbahani H., Clifford H., Wragg D., Mbole-Kariuki M.N., Van Tassell C., Sonstegard T.,
652 Woolhouse M. & Hanotte O. (2015) Signatures of positive selection in East African
653 Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. *Sci. Rep.* **5**,
654 11729.

655 Bahbahani H., Tijjani A., Mukasa C., Wragg D., Almathen F., Nash O., Akpa G.N., Mbole-
656 Kariuki M., Malla S., Woolhouse M., Sonstegard T., Van Tassell C., Blythe M., Huson
657 H. & Hanotte O. (2017) Signatures of selection for environmental adaptation and zebu
658 × taurine hybrid fitness in East African Shorthorn Zebu. *Front. Genet.* **8**.

659 Ballard J.W. & Melvin R.G. (2010) Linking the mitochondrial genotype to the organismal
660 phenotype. *Mol. Ecol.* **19**, 1523-39.

661 Bar-Yaacov D., Hadjivasiliou Z., Levin L., Barshad G., Zarivach R., Bouskila A. & Mishmar
662 D. (2015) Mitochondrial involvement in vertebrate speciation? The case of mito-nuclear
663 genetic divergence in chameleons. *Genome Biol. Evol.* **7**, 3322-36.

664 Barbato M., Hailer F., Upadhyay M., Del Corvo M., Colli L., Negrini R., Kim E.S., Crooijmans
665 R., Sonstegard T. & Ajmone-Marsan P. (2020) Adaptive introgression from indicine
666 cattle into white cattle breeds from Central Italy. *Sci. Rep.* **10**, 1279.

667 Baris T.Z., Wagner D.N., Dayan D.I., Du X., Blier P.U., Pichaud N., Oleksiak M.F. & Crawford
668 D.L. (2017) Evolved genetic and phenotypic differences due to mitochondrial-nuclear
669 interactions. *PLoS Genet.* **13**, e1006517.

670 Barreto F.S. & Burton R.S. (2013) Elevated oxidative damage is correlated with reduced fitness
671 in interpopulation hybrids of a marine copepod. *Proc. Biol. Sci.* **280**, 20131521.

672 Behl J.D., Verma N.K., Tyagi N., Mishra P., Behl R. & Joshi B.K. (2012) The major
673 histocompatibility complex in bovines: a review. *ISRN Vet Sci* **2012**, 872710.

674 Benjamini Y. & Hochberg Y. (1995) Controlling the false discovery rate - a practical and
675 powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Method.* **57**, 289-300.

676 Benson D.A., Cavanaugh M., Clark K., Karsch-Mizrachi I., Ostell J., Pruitt K.D. & Sayers E.W.
677 (2018) GenBank. *Nucleic Acids Res.* **46**, D41-d7.

678 Bergström A., Stringer C., Hajdinjak M., Scerri E.M.L. & Skoglund P. (2021) Origins of
679 modern human ancestry. *Nature* **590**, 229-37.

680 Bibi F. (2013) A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla,
681 Ruminantia) and the importance of the fossil record to systematics. *BMC Evol. Biol.* **13**,
682 166.

683 Blier P.U., Dufresne F. & Burton R.S. (2001) Natural selection and the evolution of mtDNA-
684 encoded peptides: evidence for intergenomic co-adaptation. *Trends Genet.* **17**, 400-6.

685 Bonfiglio S., Ginja C., De Gaetano A., Achilli A., Olivieri A., Colli L., Tesfaye K., Agha S.H.,
686 Gama L.T., Cattonaro F., Penedo M.C., Ajmone-Marsan P., Torroni A. & Ferretti L.
687 (2012) Origin and spread of *Bos taurus*: new clues from mitochondrial genomes
688 belonging to haplogroup T1. *PLoS ONE* **7**, e38601.

689 Bonnet T., Leblois R., Rousset F. & Crochet P.A. (2017) A reassessment of explanations for
690 discordant introgressions of mitochondrial and nuclear genomes. *Evolution* **71**, 2140-
691 58.

692 Boore J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Res.* **27**, 1767-80.

693 Bradley D.G., Loftus R.T., Cunningham P. & MacHugh D.E. (1998) Genetics and domestic
694 cattle origins. *Evol. Anthropol.* **6**, 79-86.

695 Bradley D.G., MacHugh D.E., Cunningham P. & Loftus R.T. (1996) Mitochondrial diversity
696 and the origins of African and European cattle. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 5131-
697 5.

698 Bradley D.G., MacHugh D.E., Loftus R.T., Sow R.S., Hoste C.H. & Cunningham E.P. (1994)
699 Zebu-taurine variation in Y chromosomal DNA: a sensitive assay for genetic
700 introgression in west African trypanotolerant cattle populations. *Anim. Genet.* **25**, 7-12.

701 Burton R.S., Pereira R.J. & Barreto F.S. (2013) Cytonuclear genomic interactions and hybrid
702 breakdown. *Annu. Rev. Ecol. Evol. Syst.* **44**, 281-302.

703 Calvo S.E., Clauser K.R. & Mootha V.K. (2016) MitoCarta2.0: an updated inventory of
704 mammalian mitochondrial proteins. *Nucleic Acids Res.* **44**, D1251-7.

705 Chang C.C., Chow C.C., Tellier L.C., Vattikuti S., Purcell S.M. & Lee J.J. (2015) Second-
706 generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**,
707 7.

708 Chen N., Cai Y., Chen Q., Li R., Wang K., Huang Y., Hu S., Huang S., Zhang H., Zheng Z.,
709 Song W., Ma Z., Ma Y., Dang R., Zhang Z., Xu L., Jia Y., Liu S., Yue X., Deng W.,
710 Zhang X., Sun Z., Lan X., Han J., Chen H., Bradley D.G., Jiang Y. & Lei C. (2018)
711 Whole-genome resequencing reveals world-wide ancestry and adaptive introgression
712 events of domesticated cattle in East Asia. *Nat. Commun.* **9**, 2337.

713 Chen S., Lin B.Z., Baig M., Mitra B., Lopes R.J., Santos A.M., Magee D.A., Azevedo M.,
714 Tarroso P., Sasazaki S., Ostrowski S., Mahgoub O., Chaudhuri T.K., Zhang Y.P., Costa
715 V., Royo L.J., Goyache F., Luikart G., Boivin N., Fuller D.Q., Mannen H., Bradley D.G.
716 & Beja-Pereira A. (2010) Zebu cattle are an exclusive legacy of the South Asia neolithic.
717 *Mol. Biol. Evol.* **27**, 1-6.

718 Chou J.Y. & Leu J.Y. (2015) The Red Queen in mitochondria: cyto-nuclear co-evolution,
719 hybrid breakdown and human disease. *Front. Genet.* **6**, 187.

720 Dadi H., Tibbo M., Takahashi Y., Nomura K., Hanada H. & Amano T. (2009) Variation in
721 mitochondrial DNA and maternal genetic ancestry of Ethiopian cattle populations.
722 *Anim. Genet.* **40**, 556-9.

723 Dani M.A., Heinneman M.B. & Dani S.U. (2008) Brazilian Nelore cattle: a melting pot
724 unfolded by molecular genetics. *Genet. Mol. Res.* **7**, 1127-37.

725 Decker J.E., McKay S.D., Rolf M.M., Kim J., Molina Alcala A., Sonstegard T.S., Hanotte O.,
726 Gotherstrom A., Seabury C.M., Praharani L., Babar M.E., Correia de Almeida Regitano
727 L., Yildiz M.A., Heaton M.P., Liu W.S., Lei C.Z., Reecy J.M., Saif-Ur-Rehman M.,
728 Schnabel R.D. & Taylor J.F. (2014) Worldwide patterns of ancestry, divergence, and
729 admixture in domesticated cattle. *PLoS Genet.* **10**, e1004254.

730 Delaneau O., Marchini J. & Zagury J.F. (2012) A linear complexity phasing method for
731 thousands of genomes. *Nat. Methods* **9**, 179-81.

732 Derr J.N., Hedrick P.W., Halbert N.D., Plough L., Dobson L.K., King J., Duncan C., Hunter
733 D.L., Cohen N.D. & Hedgecock D. (2012) Phenotypic effects of cattle mitochondrial
734 DNA in American bison. *Conserv Biol* **26**, 1130-6.

735 Dikmen S., Mateescu R.G., Elzo M.A. & Hansen P.J. (2018) Determination of the optimum
736 contribution of Brahman genetics in an Angus-Brahman multibreed herd for regulation
737 of body temperature during hot weather. *J. Anim. Sci.* **96**, 2175-83.

738 Du S.N.N., Khajali F., Dawson N.J. & Scott G.R. (2017) Hybridization increases mitochondrial
739 production of reactive oxygen species in sunfish. *Evolution* **71**, 1643-52.

740 Dyall S.D., Brown M.T. & Johnson P.J. (2004) Ancient invasions: from endosymbionts to
741 organelles. *Science* **304**, 253-7.

742 Ebert D. & Fields P.D. (2020) Host-parasite co-evolution and its genomic signature. *Nat. Rev.*
743 *Genet.* **21**, 754-68.

744 Edwards C.J., Baird J.F. & MacHugh D.E. (2007) Taurine and zebu admixture in Near Eastern
745 cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Anim.*
746 *Genet.* **38**, 520-4.

747 Ellis S.A. & Hammond J.A. (2014) The functional significance of cattle major
748 histocompatibility complex class I genetic diversity. *Annu. Rev. Anim. Biosci.* **2**, 285-
749 306.

750 Ellison C.K. & Burton R.S. (2006) Disruption of mitochondrial function in interpopulation
751 hybrids of *Tigriopus californicus*. *Evolution* **60**, 1382-91.

752 Ellison C.K. & Burton R.S. (2008) Interpopulation hybrid breakdown maps to the
753 mitochondrial genome. *Evolution* **62**, 631-8.

754 Ellison C.K., Niehuis O. & Gadau J. (2008) Hybrid breakdown and mitochondrial dysfunction
755 in hybrids of *Nasonia* parasitoid wasps. *J. Evol. Biol.* **21**, 1844-51.

756 Epstein H. & Mason I.L. (1971) *The Origin of the Domestic Animals of Africa (vol. 1)*. Africana
757 Publishing, London.

758 Fang L., Cai W., Liu S., Canela-Xandri O., Gao Y., Jiang J., Rawlik K., Li B., Schroeder S.G.,
759 Rosen B.D., Li C.J., Sonstegard T.S., Alexander L.J., Van Tassell C.P., VanRaden P.M.,
760 Cole J.B., Yu Y., Zhang S., Tenesa A., Ma L. & Liu G.E. (2020) Comprehensive
761 analyses of 723 transcriptomes enhance genetic and biological interpretations for
762 complex traits in cattle. *Genome Res.* **30**, 790-801.

763 Freeman A.R., Meghen C.M., MacHugh D.E., Loftus R.T., Achukwi M.D., Bado A.,
764 Sauveroche B. & Bradley D.G. (2004) Admixture and diversity in West African cattle
765 populations. *Mol. Ecol.* **13**, 3477-87.

766 Fuller D.Q. (2006) Agricultural origins and frontiers in South Asia: a working synthesis. *J.*
767 *World Prehist.* **20**, 1-86.

768 Gebrehiwot N.Z., Strucken E.M., Aliloo H., Marshall K. & Gibson J.P. (2020) The patterns of
769 admixture, divergence, and ancestry of African cattle populations determined from
770 genome-wide SNP data. *BMC Genomics* **21**, 869.

771 Gifford-Gonzalez D. & Hanotte O. (2011) Domesticating animals in Africa: implications of
772 genetic and archaeological findings. *J. World Prehist.* **24**, 1-23.

773 Ginja C., Gama L.T., Cortés O., Burriel I.M., Vega-Pla J.L., Penedo C., Sponenberg P., Cañón
774 J., Sanz A., do Egito A.A., Alvarez L.A., Giovambattista G., Agha S., Rogberg-Muñoz
775 A., Lara M.A.C., Delgado J.V. & Martinez A. (2019) The genetic ancestry of American
776 Creole cattle inferred from uniparental and autosomal genetic markers. *Sci. Rep.* **9**,
777 11486.

778 Gonzalez S. (2021) The role of mitonuclear incompatibility in bipolar disorder susceptibility
779 and resilience against environmental stressors. *Front. Genet.* **12**, 636294.

780 Hanotte O., Bradley D.G., Ochieng J.W., Verjee Y., Hill E.W. & Rege J.E. (2002) African
781 pastoralism: genetic imprints of origins and migrations. *Science* **296**, 336-9.

782 Hansen P.J. (2004) Physiological and cellular adaptations of zebu cattle to thermal stress. *Anim.*
783 *Reprod. Sci.* **82-83**, 349-60.

784 Hayes B.J., Lewin H.A. & Goddard M.E. (2013) The future of livestock breeding: genomic
785 selection for efficiency, reduced emissions intensity, and adaptation. *Trends Genet.* **29**,
786 206-14.

787 Hiendleder S., Lewalski H. & Janke A. (2008) Complete mitochondrial genomes of *Bos taurus*
788 and *Bos indicus* provide new insights into intra-species variation, taxonomy and
789 domestication. *Cytogenet. Genome Res.* **120**, 150-6.

790 Hill G.E. (2015) Mitonuclear ecology. *Mol. Biol. Evol.* **32**, 1917-27.

791 Hill G.E. (2016) Mitonuclear coevolution as the genesis of speciation and the mitochondrial
792 DNA barcode gap. *Ecol Evol* **6**, 5831-42.

793 Hill G.E. (2017) The mitonuclear compatibility species concept. *The Auk* **134**, 393-409.

794 Hirase S., Tezuka A., Nagano A.J., Sato M., Hosoya S., Kikuchi K. & Iwasaki W. (2021)
795 Integrative genomic phylogeography reveals signs of mitonuclear incompatibility in a
796 natural hybrid goby population. *Evolution* **75**, 176-94.

797 Ibeagha-Awemu E.M., Peters S.O., Bemji M.N., Adeleke M.A. & Do D.N. (2019) Leveraging
798 available resources and stakeholder involvement for improved productivity of African
799 livestock in the era of genomic breeding. *Front. Genet.* **10**, 357.

800 Isaac R.S., McShane E. & Churchman L.S. (2018) The multiple levels of mitonuclear
801 coregulation. *Annu. Rev. Genet.* **52**, 511-33.

802 Kasahara A. & Scorrano L. (2014) Mitochondria: from cell death executioners to regulators of
803 cell differentiation. *Trends Cell Biol.* **24**, 761-70.

804 Kim J., Hanotte O., Mwai O.A., Dessie T., Bashir S., Diallo B., Agaba M., Kim K., Kwak W.,
805 Sung S., Seo M., Jeong H., Kwon T., Taye M., Song K.D., Lim D., Cho S., Lee H.J.,
806 Yoon D., Oh S.J., Kemp S., Lee H.K. & Kim H. (2017) The genome landscape of
807 indigenous African cattle. *Genome Biol.* **18**, 34.

808 Kim K., Kwon T., Dessie T., Yoo D., Mwai O.A., Jang J., Sung S., Lee S., Salim B., Jung J.,
809 Jeong H., Tarekegn G.M., Tijjani A., Lim D., Cho S., Oh S.J., Lee H.-K., Kim J., Jeong
810 C., Kemp S., Hanotte O. & Kim H. (2020) The mosaic genome of indigenous African
811 cattle as a unique genetic resource for African pastoralism. *Nat. Genet.* **52**, 1099-110.

812 Kong M., Xiang H., Wang J., Liu J., Zhang X. & Zhao X. (2020) Mitochondrial DNA
813 haplotypes influence energy metabolism across chicken transmitochondrial cybrids.
814 *Genes (Basel)* **11**.

815 Koufariotis L., Hayes B.J., Kelly M., Burns B.M., Lyons R., Stothard P., Chamberlain A.J. &
816 Moore S. (2018) Sequencing the mosaic genome of Brahman cattle identifies historic
817 and recent introgression including polled. *Sci. Rep.* **8**, 17761.

818 Kramer A., Green J., Pollard J., Jr. & Tugendreich S. (2014) Causal analysis approaches in
819 Ingenuity Pathway Analysis. *Bioinformatics* **30**, 523-30.

820 Latorre-Pellicer A., Moreno-Loshuertos R., Lechuga-Vieco A.V., Sanchez-Cabo F., Torroja C.,
821 Acin-Perez R., Calvo E., Aix E., Gonzalez-Guerra A., Logan A., Bernad-Miana M.L.,
822 Romanos E., Cruz R., Cogliati S., Sobrino B., Carracedo A., Perez-Martos A.,
823 Fernandez-Silva P., Ruiz-Cabello J., Murphy M.P., Flores I., Vazquez J. & Enriquez

824 J.A. (2016) Mitochondrial and nuclear DNA matching shapes metabolism and healthy
825 ageing. *Nature* **535**, 561-5.

826 Lechuga-Vieco A.V., Justo-Méndez R. & Enríquez J.A. (2021) Not all mitochondrial DNAs
827 are made equal and the nucleus knows it. *IUBMB Life* **73**, 511-29.

828 Lee-Yaw J.A., Jacobs C.G. & Irwin D.E. (2014) Individual performance in relation to
829 cytonuclear discordance in a northern contact zone between long-toed salamander
830 (*Ambystoma macrodactylum*) lineages. *Mol. Ecol.* **23**, 4590-602.

831 Leigh J.W. & Bryant D. (2015) POPART: full-feature software for haplotype network
832 construction. *Methods Ecol. Evol.* **6**, 1110-6.

833 Loftus R.T., MacHugh D.E., Bradley D.G., Sharp P.M. & Cunningham P. (1994a) Evidence
834 for two independent domestications of cattle. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 2757-
835 61.

836 Loftus R.T., MacHugh D.E., Ngere L.O., Balain D.S., Badi A.M., Bradley D.G. & Cunningham
837 E.P. (1994b) Mitochondrial genetic variation in European, African and Indian cattle
838 populations. *Anim. Genet.* **25**, 265-71.

839 Lyons T.W., Reinhard C.T. & Planavsky N.J. (2014) The rise of oxygen in Earth's early ocean
840 and atmosphere. *Nature* **506**, 307-15.

841 Ma L., O'Connell J.R., VanRaden P.M., Shen B., Padhi A., Sun C., Bickhart D.M., Cole J.B.,
842 Null D.J., Liu G.E., Da Y. & Wiggans G.R. (2015) Cattle sex-specific recombination
843 and genetic control from a large pedigree analysis. *PLoS Genet.* **11**, e1005387.

844 Maccari G., Robinson J., Ballingall K., Guethlein L.A., Grimholt U., Kaufman J., Ho C.S., de
845 Groot N.G., Flicek P., Bontrop R.E., Hammond J.A. & Marsh S.G. (2017) IPD-MHC
846 2.0: an improved inter-species database for the study of the major histocompatibility
847 complex. *Nucleic Acids Res.* **45**, D860-d4.

848 MacHugh D.E., Shriver M.D., Loftus R.T., Cunningham P. & Bradley D.G. (1997)
849 Microsatellite DNA variation and the evolution, domestication and phylogeography of
850 taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics* **146**, 1071-86.

851 Marshall K., Gibson J.P., Mwai O., Mwacharo J.M., Haile A., Getachew T., Mrode R. & Kemp
852 S.J. (2019) Livestock genomics for developing countries - African examples in practice.
853 *Front. Genet.* **10**, 297.

854 Martin W.F., Garg S. & Zimorski V. (2015) Endosymbiotic theories for eukaryote origin.
855 *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370**, 20140330.

856 Mauki D.H., Adeola A.C., Ng'ang'a S.I., Tijjani A., Akanbi I.M., Sanke O.J., Abdussamad
857 A.M., Olaogun S.C., Ibrahim J., Dawuda P.M., Mangbon G.F., Gwakisa P.S., Yin T.T.,
858 Peng M.S. & Zhang Y.P. (2021) Genetic variation of Nigerian cattle inferred from
859 maternal and paternal genetic markers. *PeerJ* **9**, e10607.

McBride H.M., Neuspiel M. & Wasiak S. (2006) Mitochondria: more than just a powerhouse. *Curr. Biol.* **16**, R551-60.

McHugo G.P., Dover M.J. & MacHugh D.E. (2019) Unlocking the origins and biology of domestic animals using ancient DNA and paleogenomics. *BMC Biol.* **17**, 98.

McKenzie M., Chiotis M., Pinkert C.A. & Trounce I.A. (2003) Functional respiratory chain analyses in murid xenomitochondrial cybrids expose coevolutionary constraints of cytochrome b and nuclear subunits of complex III. *Mol. Biol. Evol.* **20**, 1117-24.

McKenzie M., Trounce I.A., Cassar C.A. & Pinkert C.A. (2004) Production of homoplasmic xenomitochondrial mice. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1685-90.

McManus C.M., Rezende Paiva S. & Faria D. (2020) Genomics and climate change. *Rev. Sci. Tech.* **39**, 481-90.

McTavish E.J. & Hillis D.M. (2014) A genomic approach for distinguishing between recent and ancient admixture as applied to cattle. *J. Hered.*

Mészáros G., Boison S.A., Pérez O'Brien A.M., Ferenčaković M., Curik I., Da Silva M.V., Utsunomiya Y.T., Garcia J.F. & Sölkner J. (2015) Genomic analysis for managing small and endangered populations: a case study in Tyrol Grey cattle. *Front. Genet.* **6**, 173.

Mills E.L., Kelly B. & O'Neill L.A.J. (2017) Mitochondria are the powerhouses of immunity. *Nat. Immunol.* **18**, 488-98.

Morales H.E., Pavlova A., Amos N., Major R., Kilian A., Greening C. & Sunnucks P. (2018) Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nat. Ecol. Evol.* **2**, 1258-67.

Mrode R., Ojango J.M.K., Okeyo A.M. & Mwacharo J.M. (2019) Genomic selection and use of molecular tools in breeding programs for indigenous and crossbred cattle in developing countries: current status and future prospects. *Front. Genet.* **9**, 694.

Mwai O., Hanotte O., Kwon Y.J. & Cho S. (2015) African indigenous cattle: unique genetic resources in a rapidly changing world. *Asian-Australas. J. Anim. Sci.* **28**, 911-21.

Norris E.T., Rishishwar L., Chande A.T., Conley A.B., Ye K., Valderrama-Aguirre A. & Jordan I.K. (2020) Admixture-enabled selection for rapid adaptive evolution in the Americas. *Genome Biol.* **21**, 29.

O'Brien A.M., Holler D., Boison S.A., Milanesi M., Bomba L., Utsunomiya Y.T., Carvalheiro R., Neves H.H., da Silva M.V., VanTassell C.P., Sonstegard T.S., Meszaros G., Ajmone-Marsan P., Garcia F. & Solkner J. (2015) Low levels of taurine introgression in the current Brazilian Nelore and Gir indicine cattle populations. *Genet. Sel. Evol.* **47**, 31.

Pichaud N., Bérubé R., Côté G., Belzile C., Dufresne F., Morrow G., Tanguay R.M., Rand D.M. & Blier P.U. (2019) Age dependent dysfunction of mitochondrial and ROS metabolism induced by mitonuclear mismatch. *Front. Genet.* **10**, 130.

897 Pitt D., Sevane N., Nicolazzi E.L., MacHugh D.E., Park S.D.E., Colli L., Martinez R., Bruford
898 M.W. & Orozco-terWengel P. (2019) Domestication of cattle: Two or three events?
899 *Evol. Appl.* **12**, 123-36.

900 Pramod R.K., Velayutham D., P K.S., P S.B., Zachariah A., Zachariah A., B C., S S.S., P G.,
901 Dhinoth Kumar B., Iype S., Gupta R., Santhosh S. & Thomas G. (2019) Complete
902 mitogenome reveals genetic divergence and phylogenetic relationships among Indian
903 cattle (*Bos indicus*) breeds. *Anim. Biotechnol.* **30**, 219-32.

904 Quinlan A.R. & Hall I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic
905 features. *Bioinformatics* **26**, 841-2.

906 R Core Team (2019) R: A language and environment for statistical computing. R Foundation
907 for Statistical Computing. http://www.R-project.org, Vienna, Austria.

908 Radwan J., Babik W., Kaufman J., Lenz T.L. & Winternitz J. (2020) Advances in the
909 evolutionary understanding of MHC polymorphism. *Trends Genet.* **36**, 298-311.

910 Raj A., Stephens M. & Pritchard J.K. (2014) fastSTRUCTURE: variational inference of
911 population structure in large SNP data sets. *Genetics* **197**, 573-89.

912 Rand D.M., Haney R.A. & Fry A.J. (2004) Cytonuclear coevolution: the genomics of
913 cooperation. *Trends Ecol. Evol.* **19**, 645-53.

914 Rank N.E., Mardulyn P., Heidl S.J., Roberts K.T., Zavala N.A., Smiley J.T. & Dahlhoff E.P.
915 (2020) Mitonuclear mismatch alters performance and reproductive success in naturally
916 introgressed populations of a montane leaf beetle. *Evolution* **74**, 1724-40.

917 Rincon G., Weber K.L., Eenennaam A.L., Golden B.L. & Medrano J.F. (2011) Hot topic:
918 performance of bovine high-density genotyping platforms in Holsteins and Jerseys. *J.*
919 *Dairy Sci.* **94**, 6116-21.

920 Ristic M. & McIntyre I. (1981) *Diseases of Cattle in the Tropics: Economic and Zoonotic*
921 *Relevance*. Martinus Nijhoff Publishers, The Hague, Netherlands.

922 Rosen B.D., Bickhart D.M., Schnabel R.D., Koren S., Elsik C.G., Tseng E., Rowan T.N., Low
923 W.Y., Zimin A., Couldrey C., Hall R., Li W., Rhie A., Ghurye J., McKay S.D., Thibaud-
924 Nissen F., Hoffman J., Murdoch B.M., Snelling W.M., McDaneld T.G., Hammond J.A.,
925 Schwartz J.C., Nandolo W., Hagen D.E., Dreischer C., Schultheiss S.J., Schroeder S.G.,
926 Phillippy A.M., Cole J.B., Van Tassell C.P., Liu G., Smith T.P.L. & Medrano J.F.
927 (2020) De novo assembly of the cattle reference genome with single-molecule
928 sequencing. *Gigascience* **9**.

929 Rosenberg N.A. (2005) Algorithms for selecting informative marker panels for population
930 assignment. *J. Comput. Biol.* **12**, 1183-201.

931 Rosenberg N.A., Li L.M., Ward R. & Pritchard J.K. (2003) Informativeness of genetic markers
932 for inference of ancestry. *Am. J. Hum. Genet.* **73**, 1402-22.

933 Salter-Townshend M. & Myers S. (2019) Fine-scale inference of ancestry segments without
934 prior knowledge of admixing groups. *Genetics* **212**, 869-89.

935 Scheet P. & Stephens M. (2006) A fast and flexible statistical model for large-scale population
936 genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J.*
937 *Hum. Genet.* **78**, 629-44.

938 Seifert H.S.H. (1996) *Tropical Animal Health*. Kluwer Academic Publishers, Dordrecht,
939 Netherlands.

940 Seixas F.A., Boursot P. & Melo-Ferreira J. (2018) The genomic impact of historical
941 hybridization with massive mitochondrial DNA introgression. *Genome Biol.* **19**, 91.

942 Senczuk G., Mastrangelo S., Ajmone-Marsan P., Becskei Z., Colangelo P., Colli L., Ferretti L.,
943 Karsli T., Lancioni H., Lasagna E., Marletta D., Persichilli C., Portolano B., Sarti F.M.,
944 Ciani E. & Pilla F. (2021) On the origin and diversification of Podolian cattle breeds:
945 testing scenarios of European colonization using genome-wide SNP data. *Genet. Sel.*
946 *Evol.* **53**, 48.

947 Sharbrough J., Havird J.C., Noe G.R., Warren J.M. & Sloan D.B. (2017) The mitonuclear
948 dimension of Neanderthal and Denisovan ancestry in modern human genomes. *Genome*
949 *Biol. Evol.* **9**, 1567-81.

950 Sloan D.B., Fields P.D. & Havird J.C. (2015) Mitonuclear linkage disequilibrium in human
951 populations. *Proc. Biol. Sci.* **282**.

952 Sloan D.B., Havird J.C. & Sharbrough J. (2017) The on-again, off-again relationship between
953 mitochondrial genomes and species boundaries. *Mol. Ecol.* **26**, 2212-36.

954 Sloan D.B., Warren J.M., Williams A.M., Wu Z., Abdel-Ghany S.E., Chicco A.J. & Havird
955 J.C. (2018) Cytonuclear integration and co-evolution. *Nat. Rev. Genet.* **19**, 635-48.

956 Spierer A.N., Yoon D., Zhu C.T. & Rand D.M. (2021) FreeClimber: automated quantification
957 of climbing performance in *Drosophila*. *J. Exp. Biol.* **224**.

958 Spinelli J.B. & Haigis M.C. (2018) The multifaceted contributions of mitochondria to cellular
959 metabolism. *Nat. Cell Biol.* **20**, 745-54.

960 Sujkowski A., Spierer A.N., Rajagopalan T., Bazzell B., Safdar M., Imsirovic D., Arking R.,
961 Rand D.M. & Wessells R. (2019) Mito-nuclear interactions modify *Drosophila* exercise
962 performance. *Mitochondrion* **47**, 188-205.

963 Sunnucks P., Morales H.E., Lamb A.M., Pavlova A. & Greening C. (2017) Integrative
964 approaches for studying mitochondrial and nuclear genome co-evolution in oxidative
965 phosphorylation. *Front. Genet.* **8**, 25.

966 Taye M., Lee W., Caetano-Anolles K., Dessie T., Cho S., Jong Oh S., Lee H.-K. & Kim H.
967 (2018) Exploring the genomes of East African Indicine cattle breeds reveals signature
968 of selection for tropical environmental adaptation traits. *Cogent Food. Agric.* **4**,
969 1552552.

970 Taye M., Lee W., Caetano-Anolles K., Dessie T., Hanotte O., Mwai O.A., Kemp S., Cho S.,
971 Oh S.J., Lee H.K. & Kim H. (2017) Whole genome detection of signature of positive
972 selection in African cattle reveals selection for thermotolerance. *Anim. Sci. J.* **88**, 1889-
973 901.

974 Tijjani A., Utsunomiya Y.T., Ezekwe A.G., Nashiru O. & Hanotte O. (2019) Genome sequence
975 analysis reveals selection signatures in endangered trypanotolerant West African
976 Muturu cattle. *Front. Genet.* **10**, 442.

977 Tiku V., Tan M.W. & Dikic I. (2020) Mitochondrial functions in infection and immunity.
978 *Trends Cell Biol.* **30**, 263-75.

979 Tourmente M., Hirose M., Ibrahim S., Dowling D.K., Tompkins D.M., Roldan E.R.S. &
980 Gemmell N.J. (2017) mtDNA polymorphism and metabolic inhibition affect sperm
981 performance in conplastic mice. *Reproduction* **154**, 341-54.

982 Trier C.N., Hermansen J.S., Sætre G.P. & Bailey R.I. (2014) Evidence for mito-nuclear and
983 sex-linked reproductive barriers between the hybrid Italian sparrow and its parent
984 species. *PLoS Genet.* **10**, e1004075.

985 Troy C.S., MacHugh D.E., Bailey J.F., Magee D.A., Loftus R.T., Cunningham P., Chamberlain
986 A.T., Sykes B.C. & Bradley D.G. (2001) Genetic evidence for Near-Eastern origins of
987 European cattle. *Nature* **410**, 1088-91.

988 Utsunomiya Y.T., Milanesi M., Fortes M.R.S., Porto-Neto L.R., Utsunomiya A.T.H., Silva M.,
989 Garcia J.F. & Ajmone-Marsan P. (2019) Genomic clues of the evolutionary history of
990 *Bos indicus* cattle. *Anim. Genet.* **50**, 557-68.

991 Ventura R.V., Brito L.F., Oliveira G.A., Daetwyler H.D., Schenkel F.S., Sargolzaei M.,
992 Vandervoort G., Silva F.F.E., Miller S.P., Carvalho M.E., Santana M.H.A., Mattos E.C.,
993 Fonseca P., Eler J.P. & Ferraz J.B.S. (2020) A comprehensive comparison of high-
994 density SNP panels and an alternative ultra-high-density panel for genomic analyses in
995 Nellore cattle. *Anim. Prod. Sci.* **60**, 333-46.

996 Verdugo M.P., Mullin V.E., Scheu A., Mattiangeli V., Daly K.G., Maisano Delser P., Hare
997 A.J., Burger J., Collins M.J., Kehati R., Hesse P., Fulton D., Sauer E.W., Mohaseb F.A.,
998 Davoudi H., Khazaeli R., Lhuillier J., Rapin C., Ebrahimi S., Khasanov M., Vahidi
999 S.M.F., MacHugh D.E., Ertugrul O., Koukouli-Chrysanthaki C., Sampson A., Kazantzis
1000 G., Kontopoulos I., Bulatovic J., Stojanovic I., Mikdad A., Benecke N., Linstadter J.,
1001 Sablin M., Bendrey R., Gourichon L., Arbuckle B.S., Mashkour M., Orton D., Horwitz
1002 L.K., Teasdale M.D. & Bradley D.G. (2019) Ancient cattle genomics, origins, and rapid
1003 turnover in the Fertile Crescent. *Science* **365**, 173-6.

1004 Wang J., Xiang H., Liu L., Kong M., Yin T. & Zhao X. (2017) Mitochondrial haplotypes
1005 influence metabolic traits across bovine inter- and intra-species cybrids. *Sci. Rep.* **7**,
1006 4179.

1007 Wang K., Lenstra J.A., Liu L., Hu Q., Ma T., Qiu Q. & Liu J. (2018) Incomplete lineage sorting
1008 rather than hybridization explains the inconsistent phylogeny of the wisent. *Commun.*
1009 *Biol.* **1**, 169.

1010 Weinberg S.E., Sena L.A. & Chandel N.S. (2015) Mitochondria in the regulation of innate and
1011 adaptive immunity. *Immunity* **42**, 406-17.

1012 Weir B.S. & Cockerham C.C. (1984) Estimating *F*-statistics for the analysis of population
1013 structure. *Evolution* **38**, 1358-70.

1014 Wickham H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer Publishing
1015 Company, New York, USA.

1016 Wolff J.N., Ladoukakis E.D., Enriquez J.A. & Dowling D.K. (2014) Mitonuclear interactions:
1017 evolutionary consequences over multiple biological scales. *Philos. Trans. R. Soc. Lond.*
1018 *B. Biol. Sci.* **369**, 20130443.

1019 Woodson J.D. & Chory J. (2008) Coordination of gene expression between organellar and
1020 nuclear genomes. *Nat. Rev. Genet.* **9**, 383-95.

1021 Wu D.D., Ding X.D., Wang S., Wojcik J.M., Zhang Y., Tokarska M., Li Y., Wang M.S.,
1022 Faruque O., Nielsen R., Zhang Q. & Zhang Y.P. (2018) Pervasive introgression
1023 facilitated domestication and adaptation in the *Bos* species complex. *Nat. Ecol. Evol.* **2**,
1024 1139-45.

1025 Xia X., Qu K., Zhang G., Jia Y., Ma Z., Zhao X., Huang Y., Chen H., Huang B. & Lei C. (2019)
1026 Comprehensive analysis of the mitochondrial DNA diversity in Chinese cattle. *Anim.*
1027 *Genet.* **50**, 70-3.

1028 Yang Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**,
1029 1586-91.

1030 Yates A.D., Achuthan P., Akanni W., Allen J., Allen J., Alvarez-Jarreta J., Amode M.R.,
1031 Armean I.M., Azov A.G., Bennett R., Bhai J., Billis K., Boddu S., Marugan J.C.,
1032 Cummins C., Davidson C., Dodiya K., Fatima R., Gall A., Giron C.G., Gil L., Grego
1033 T., Haggerty L., Haskell E., Hourlier T., Izuogu O.G., Janacek S.H., Juettemann T., Kay
1034 M., Lavidas I., Le T., Lemos D., Martinez J.G., Maurel T., McDowall M., McMahon
1035 A., Mohanan S., Moore B., Nuhn M., Oheh D.N., Parker A., Parton A., Patricio M.,
1036 Sakthivel M.P., Abdul Salam A.I., Schmitt B.M., Schuilenburg H., Sheppard D.,
1037 Sycheva M., Szuba M., Taylor K., Thormann A., Threadgold G., Vullo A., Walts B.,
1038 Winterbottom A., Zadissa A., Chakiachvili M., Flint B., Frankish A., Hunt S.E., G I.I.,
1039 Kostadima M., Langridge N., Loveland J.E., Martin F.J., Morales J., Mudge J.M.,
1040 Muffato M., Perry E., Ruffier M., Trevanion S.J., Cunningham F., Howe K.L., Zerbino
1041 D.R. & Flicek P. (2020) Ensembl 2020. *Nucleic Acids Res.* **48**, D682-D8.

1042 Yu X., Gimsa U., Wester-Rosenlöf L., Kanitz E., Otten W., Kunz M. & Ibrahim S.M. (2009)
1043 Dissecting the effects of mtDNA variations on complex traits using mouse conplastic
1044 strains. *Genome Res.* **19**, 159-65.

1045 Zaidi A.A. & Makova K.D. (2019) Investigating mitonuclear interactions in human admixed
1046 populations. *Nat. Ecol. Evol.* **3**, 213-22.

1047 Zardoya R. (2020) Recent advances in understanding mitochondrial genome diversity.
1048 *F1000Res* **9**.

1049 Zeder M.A. (2011) The origins of agriculture in the Near East. *Curr. Anthropol.* **52**, S221-S35.

1050 Zhang K., Lenstra J.A., Zhang S., Liu W. & Liu J. (2020) Evolution and domestication of the
1051 Bovini species. *Anim. Genet.* **51**, 637-57.

1052

1053 **TABLE 1** Cattle breeds/populations, geographical origins, and sources of BovineHD 777K SNP data.

Cattle breed/population	Code	Type/morphology	Country of origin	Source 1[a]	Source 2[b]	Source 3[c]	Source 4[d]	Total (*n*)
Muturu	MUTU	West African taurine	Nigeria	8	---	20	---	28
N'Dama	NDAG	West African taurine	Guinea	24	9	---	23	56
Holstein-Friesian	HOLS	European taurine	Netherlands	59	---	---	---	59
Jersey	JRSY	European taurine	United Kingdom	32	---	---	---	32
Ankole	ANKO	East African admixed	Uganda	25	---	---	---	25
East African Shorthorn Zebu	EASZ	East African admixed	Kenya	92	---	19	---	111
Karamojong	KARA	East African admixed	Uganda	16	---	---	---	16
Nganda	NGAN	East African admixed	Uganda	23	---	4	---	27
Adamawa Gudali	ADAG	West African admixed	Nigeria	23	---	2	---	25
Borgou	BORG	West African admixed	Benin	---	---	---	50	50
Bunaji	BUNA	West African admixed	Nigeria	22	---	5	---	27

Keteku	KETE	West African admixed	Nigeria	---	---	22	---	22
Red Bororo	REDB	West African admixed	Nigeria	22	---	4	---	26
Sokoto Gudali	SOKG	West African admixed	Nigeria	19	---	---	---	19
Gir	GIR	Asian zebu	India	28	---	---	---	28
Hariana	HARI	Asian zebu	India	---	10	---	---	10
Nelore	NELO	Asian zebu	Brazil	34	---	---	---	34
Sahiwal	SAHI	Asian zebu	India	---	10	---	---	10

1054 [a] Bahbahani *et al.* (2015); [b] Verdugo *et al.* (2019); [c] Acceligen cattle genotyping database; [d] genotyped for the present study.
1055

1056 **TABLE 2** Ancestry components estimated using FASTSTRUCTURE with estimated times for the start of the admixture process in African admixed
1057 cattle populations generated using MOSAIC.

Code	Type/morphology	Country of origin	African taurine ancestry	European taurine ancestry	Asian zebu ancestry	Generations since start of admixture	Start of admixture
MUTU	West African taurine	Nigeria	0.990 ± 0.019	0.010 ± 0.018	0.000 ± 0.003		----
NDAG	West African taurine	Guinea	1.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000		----
HOLS	European taurine	Netherlands	0.010 ± 0.025	0.990 ± 0.025	0.000 ± 0.000		----
JRSY	European taurine	United Kingdom	0.010 ± 0.022	0.990 ± 0.022	0.000 ± 0.000		----
ANKO	East African admixed	Uganda	0.420 ± 0.010	0.020 ± 0.016	0.550 ± 0.013	54.1	1641 – 1804 CE
EASZ	East African admixed	Kenya	0.250 ± 0.015	0.020 ± 0.065	0.730 ± 0.055	46.2	1697 – 1835 CE
KARA	East African admixed	Uganda	0.260 ± 0.008	0.000 ± 0.001	0.740 ± 0.008	65.1	1564 – 1760 CE
NGAN	East African admixed	Uganda	0.310 ± 0.024	0.100 ± 0.052	0.590 ± 0.039	15.9	1909 – 1956 CE
ADAG	West African admixed	Nigeria	0.320 ± 0.010	0.000 ± 0.001	0.680 ± 0.010	79.6	1463 – 1702 CE
BORG	West African admixed	Benin	0.610 ± 0.047	0.010 ± 0.027	0.370 ± 0.055	20.5	1877 – 1938 CE

BUNA	West African admixed	Nigeria	0.340 ± 0.076	0.000 ± 0.007	0.660 ± 0.083	87.9	1405 – 1668 CE
KETE	West African admixed	Nigeria	0.400 ± 0.093	0.000 ± 0.001	0.600 ± 0.093	26.6	1834 – 1914 CE
REDB	West African admixed	Nigeria	0.320 ± 0.009	0.000 ± 0.001	0.680 ± 0.009	70	1530 – 1740 CE
SOKG	West African admixed	Nigeria	0.320 ± 0.008	0.000 ± 0.000	0.680 ± 0.008	95.9	1349 – 1636 CE
GIR	Asian zebu	India	0.000 ± 0.000	0.000 ± 0.000	1.000 ± 0.000		----
HARI	Asian zebu	India	0.000 ± 0.000	0.000 ± 0.000	1.000 ± 0.000		----
NELO	Asian zebu	Brazil	0.000 ± 0.000	0.000 ± 0.001	1.000 ± 0.001		----
SAHI	Asian zebu	India	0.000 ± 0.000	0.000 ± 0.000	1.000 ± 0.000		----

1058 *Note.* For each admixed population, the generations since admixture started were obtained from the NDAG:GIR coancestry plots (Figure 4). A
1059 generation interval range of 4–7 years for managed domestic cattle was used. CE = Common Era.

1060 **TABLE 3** Statistical test results for comparisons of HMG, LMG, and NMG distributions of
1061 African taurine local ancestry deviations for East and West African admixed cattle.

Code	Type/morphology	*P* value (HMG > NMG)	*P* value (LMG > NMG)	*P* value (HMG > LMG)
ANKO	East African admixed	5.69×10^{-76}	2.73×10^{-20}	4.88×10^{-24}
EASZ	East African admixed	3.09×10^{-206}	0.9567	7.51×10^{-239}
KARA	East African admixed	1.52×10^{-86}	5.10×10^{-5}	2.30×10^{-61}
NGAN	East African admixed	6.41×10^{-315}	1.87×10^{-10}	4.13×10^{-250}
ADAG	West African admixed	2.02×10^{-14}	1.46×10^{-41}	1.0000
BORG	West African admixed	8.74×10^{-41}	0.6345	1.98×10^{-50}
BUNA	West African admixed	2.56×10^{-45}	0.9998	3.35×10^{-81}
KETE	West African admixed	2.41×10^{-206}	1.6×10^{-20}	3.54×10^{-110}
REDB	West African admixed	1.08×10^{-51}	1.0000	6.00×10^{-94}
SOKG	West African admixed	8.13×10^{-22}	1.0000	1.61×10^{-95}

1062

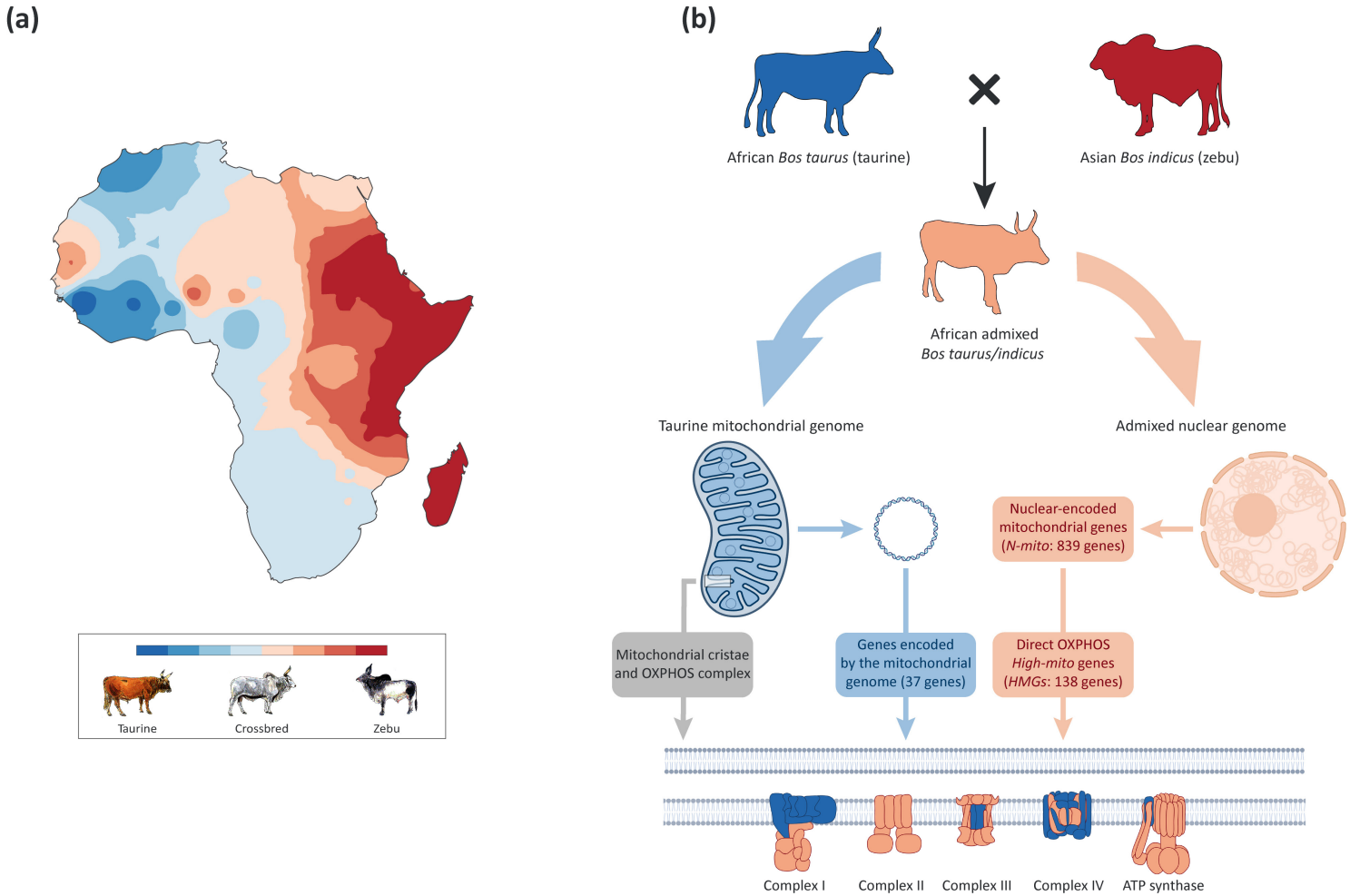


FIGURE 1 Geographical patterns of African *Bos taurus* and Asian *Bos indicus* admixture and a schematic illustrating mitonuclear interactions for hybrid African cattle that retain the taurine mitochondrial genome. (a) Map of Africa showing an interpolated synthetic map illustrating spatial distribution of African *B. taurus* and Asian *B. indicus* admixture. Admixture data was generated from the first principal component (PC1) of a principal component analysis (PCA) of microsatellite genetic variation across African cattle populations (Hanotte *et al.* 2002). Modified from (McHugo *et al.* 2019) under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0). (b) Mitonuclear interactions that can give rise to mitonuclear incompatibilities in crossbred and hybrid African cattle populations. All African cattle surveyed to-date retain the taurine mitochondrial genome (some figure components created with a BioRender.com licence).

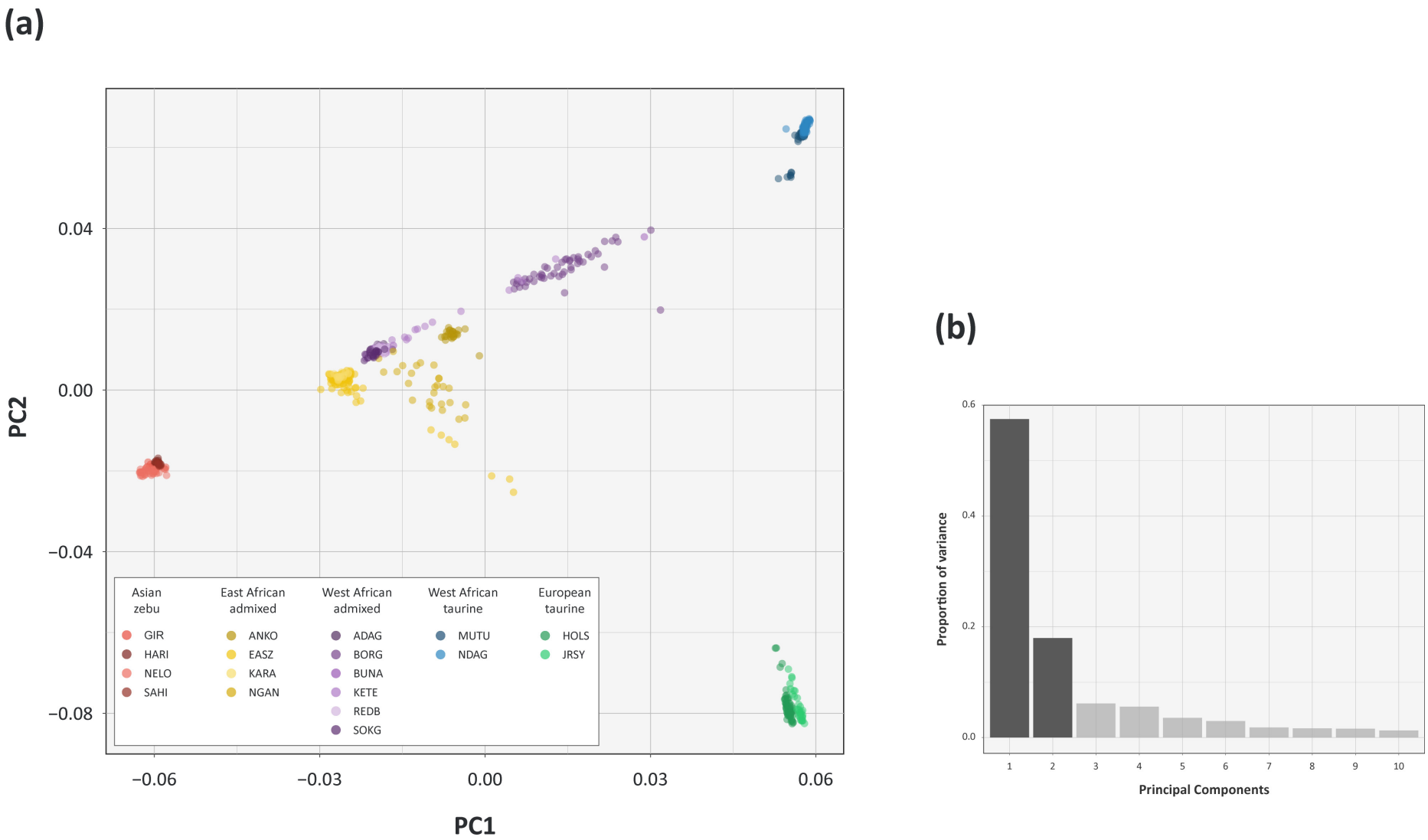


FIGURE 2 Results of the principal component analysis (PCA) for 605 animals from 18 different cattle breeds genotyped for 562,635 SNPs. (a) PCA plot showing the coordinates for each animal based on the first two principal components. Principal component 1 (PC1) differentiates the *Bos taurus* and *Bos indicus* evolutionary lineages, while PC2 separates the African and European taurine groups. (b) Histogram plot showing the relative variance contributions for the first 10 PCs with PC1 and PC2 accounting for 58.4% and 17.9% of the total variation for PC1–10, respectively.

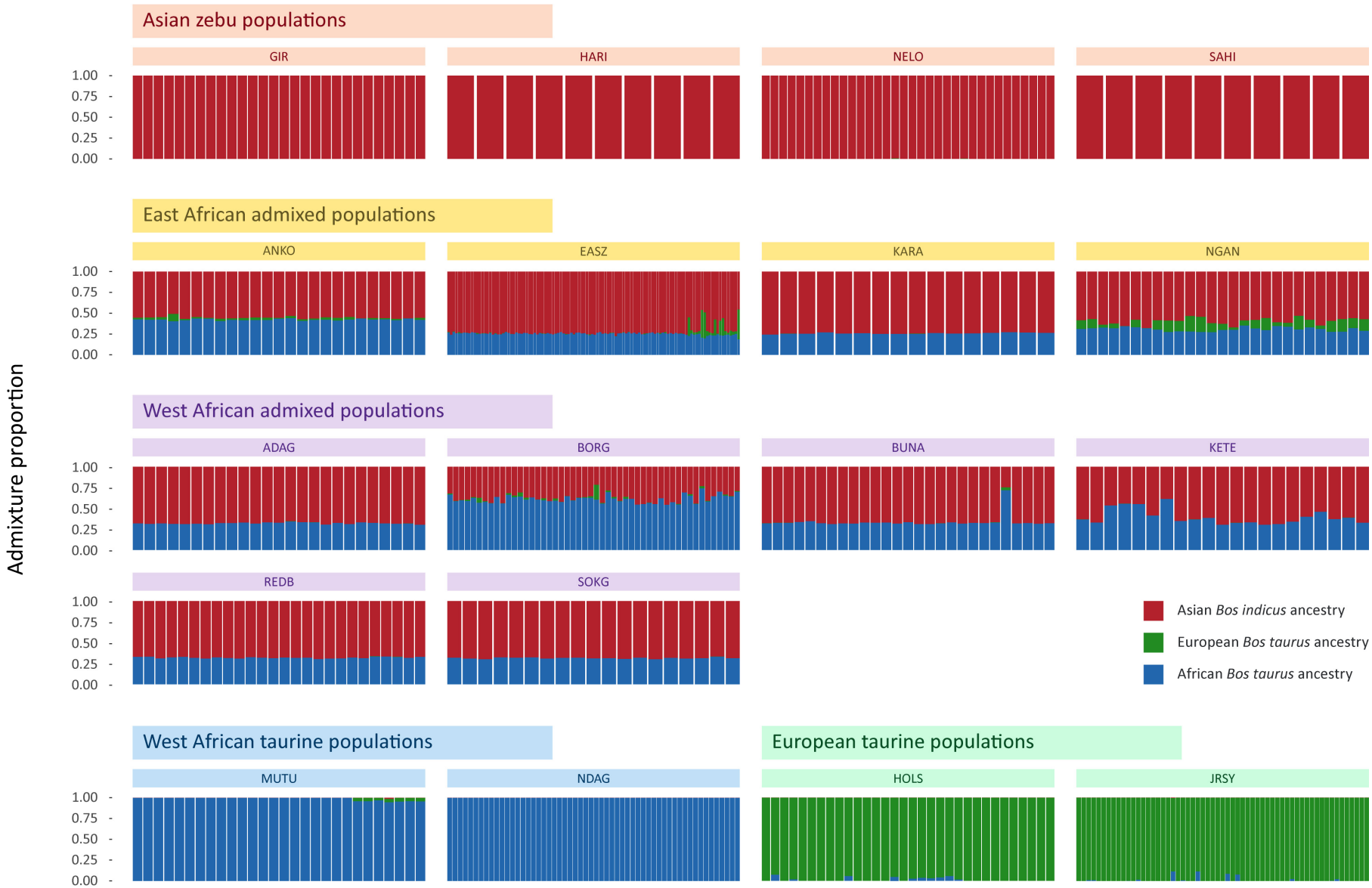
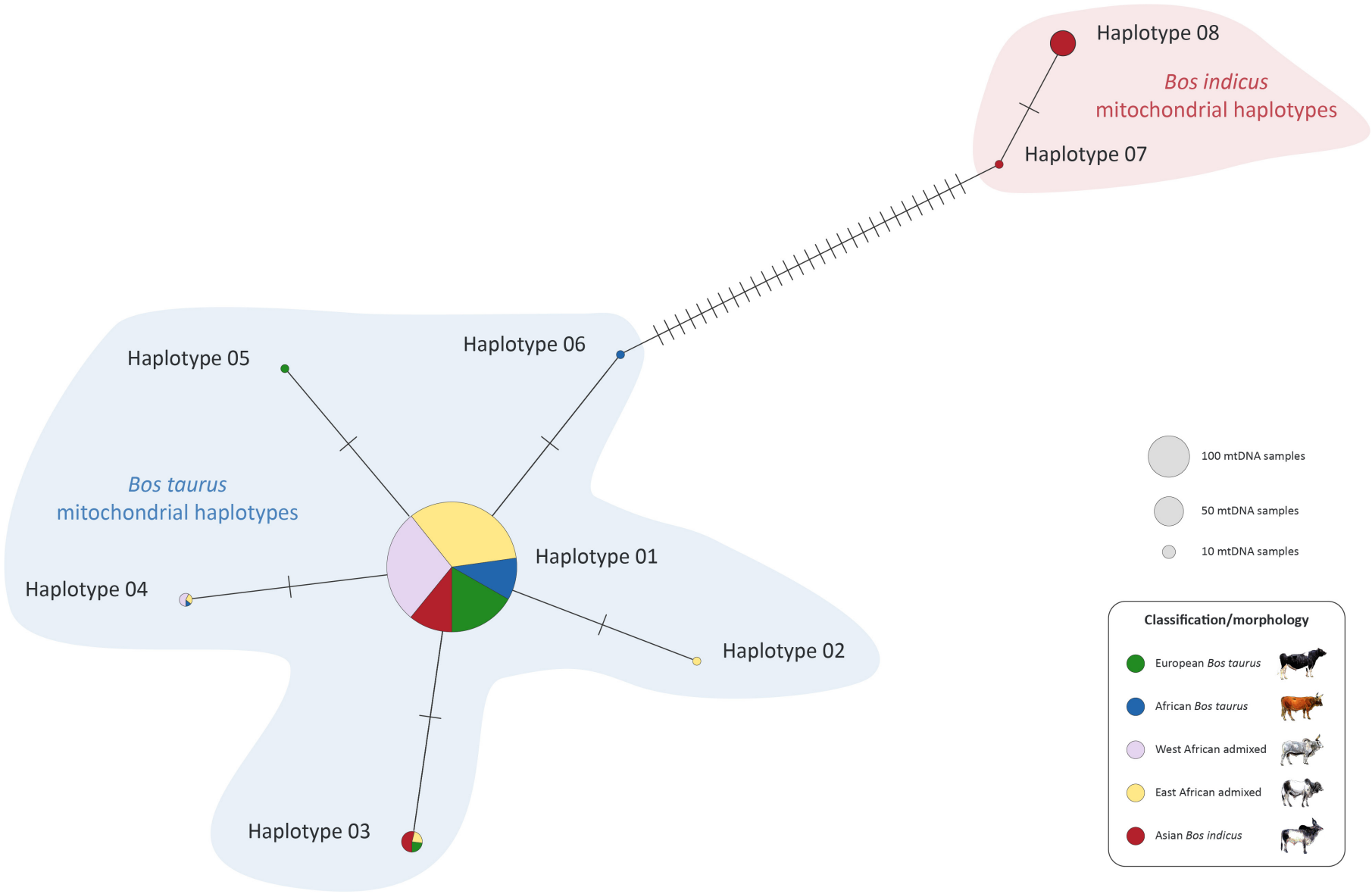


FIGURE 3 Unsupervised genetic structure plot for Asian zebu, East and West African admixed cattle, and West African and European taurine breeds. Results for an inferred number of ancestry clusters of $K = 3$ is shown, which corresponds to Asian *Bos indicus* (red), European *Bos taurus* (green), and African *B. taurus* (blue) ancestral components, respectively.



1082

1083 **FIGURE 4** Phylogenetic network of 491 cattle mtDNA haplotypes generated using 39 ancestry-informative mtDNA SNPs. This mtDNA haplotype network demonstrates that all
1084 surveyed African cattle (47 taurine, 156 East African admixed, and 136 West African admixed) possess and retain *Bos taurus* mitochondrial genomes.

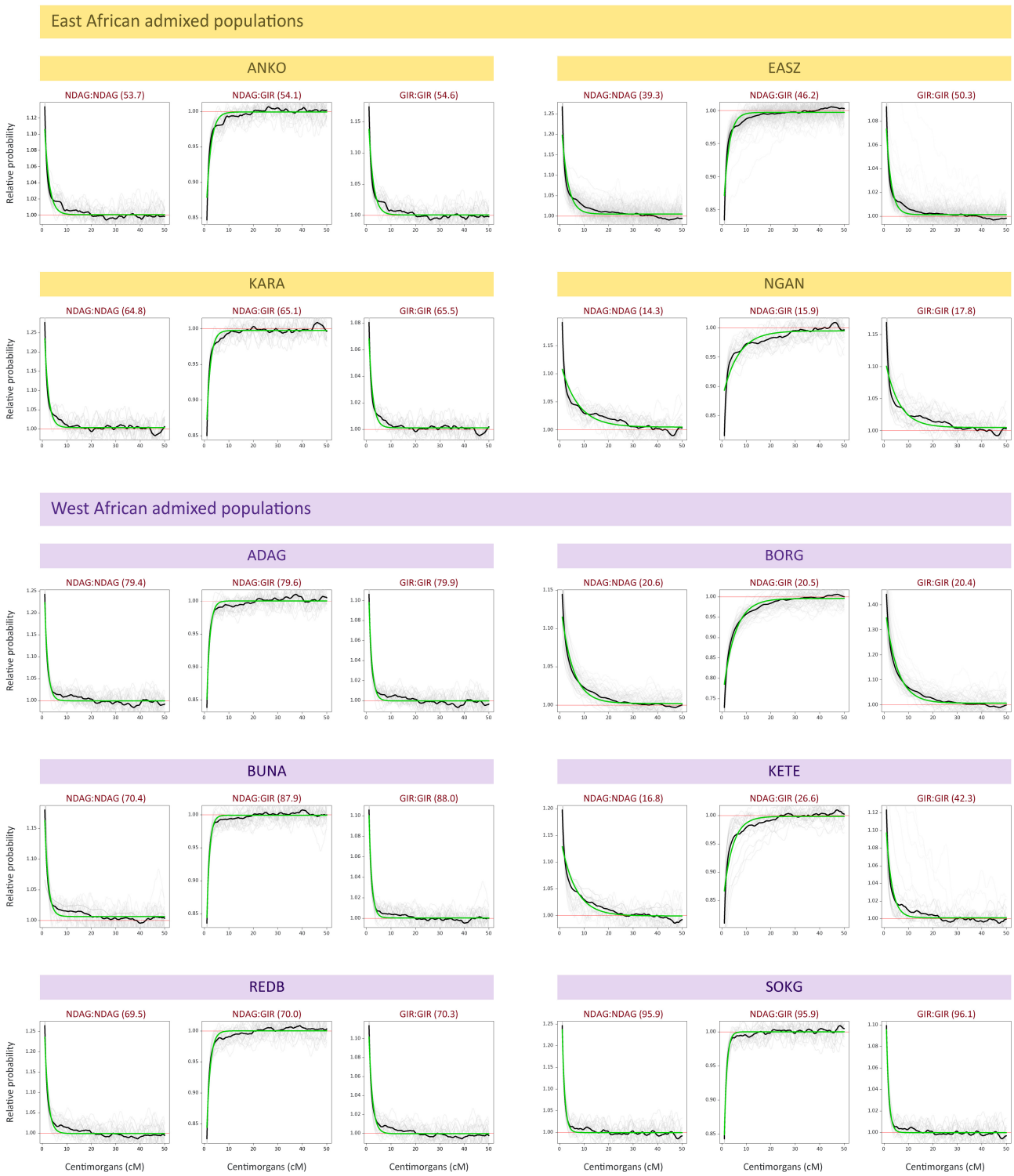


FIGURE 5 Coancestry curve plots generated using MOSAIC for 10 East and West African admixed cattle populations. These curves show the exponential decay of the ratio of probabilities of pairs of local ancestries (y-axis) as a function of genetic distance (x-axis). The pair of ancestries used for each curve is shown on the top of each plot with the estimated number of generations since the start of admixture in brackets. For each plot, the green line represents the fitted curve, the black line shows the across targets ratio, and the grey lines indicate the per target ratio (further information in Salter-Townshend & Myers 2019).

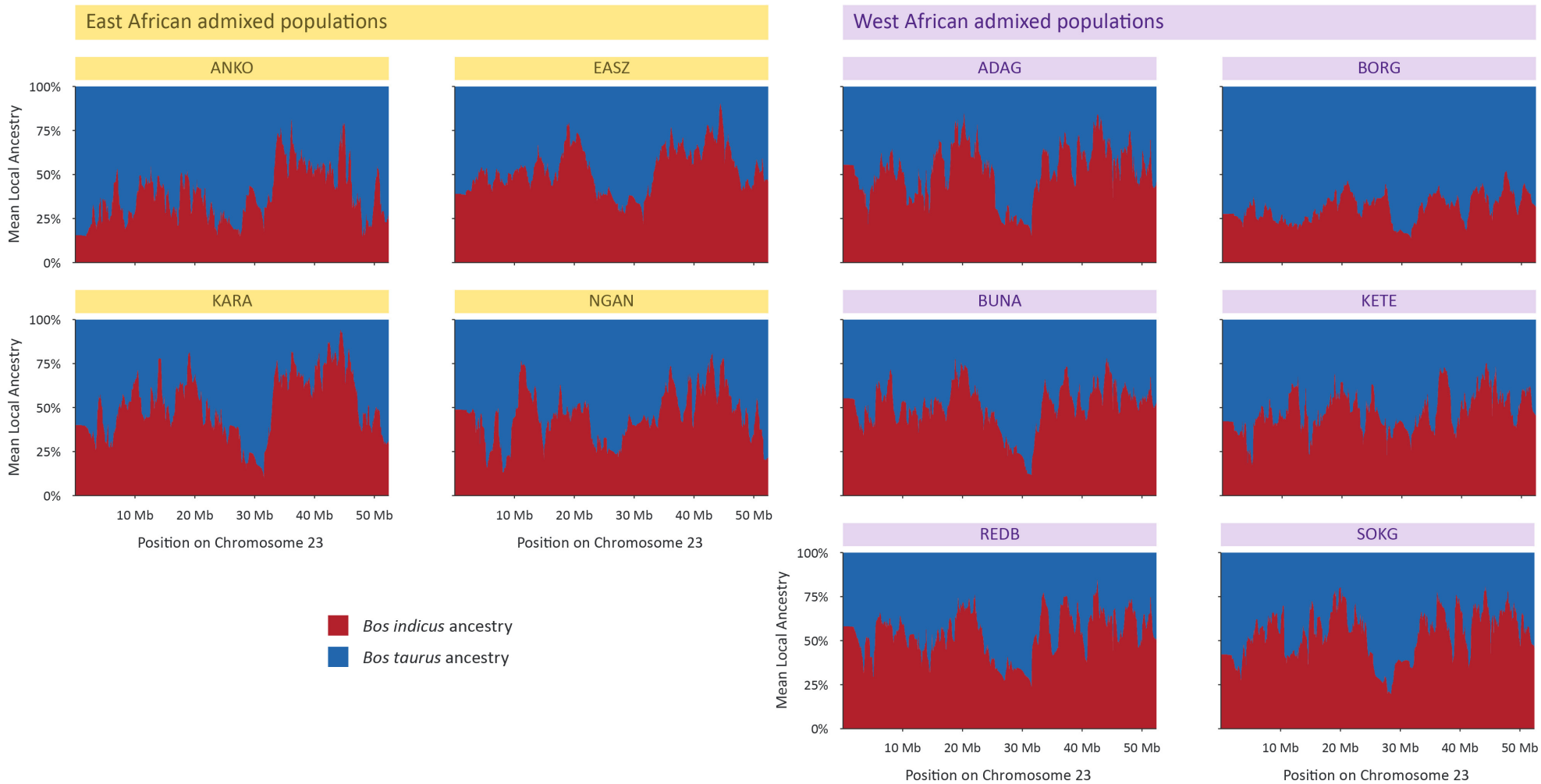


FIGURE 6 Local ancestry plots for bovine chromosome 23 (BTA23) showing Asian *Bos indicus* (red) and African *Bos taurus* (blue) subchromosomal ancestry in four East African and six West African admixed cattle populations. Physical distance along BTA23 is indicated in megabases (Mb).

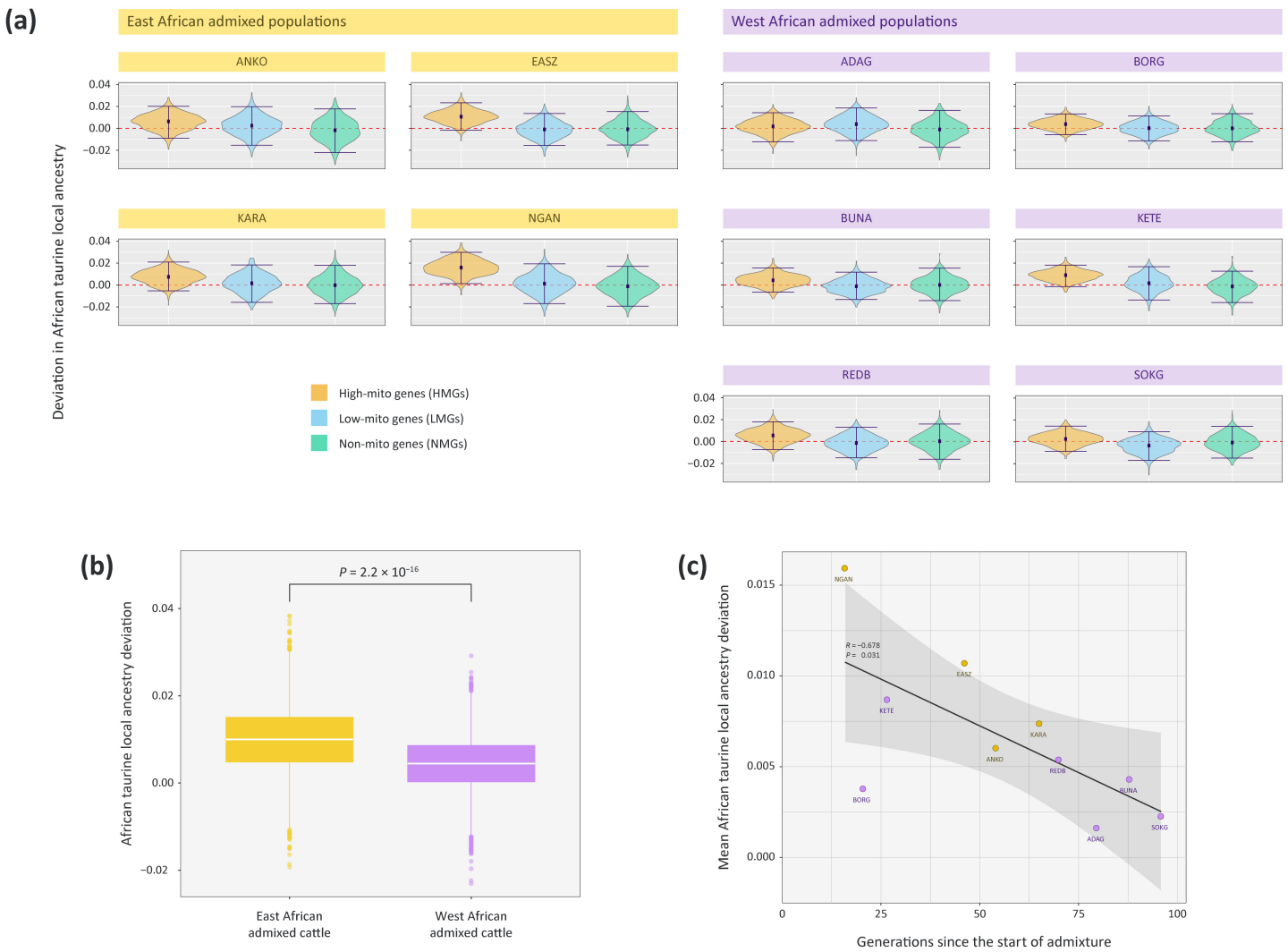


FIGURE 7 African taurine local ancestry deviations for three different functional gene sets. (a) Violin plots of African taurine local ancestry deviations for the HMG, LMG, and NMG subsets with positive deviations indicating retention of African taurine gene haplotypes. Black data points indicate the median values and horizontal lines represent the 95% confidence interval. (b) Box plot of African taurine local ancestry deviations for the HMG subset in the East African and the West African admixed groups. White lines indicate the median values and yellow and purple boxes indicate the interquartile ranges. (c) Linear regression plot of mean taurine local ancestry deviation for the HMG subset against generations since the start of admixture for four East (yellow) and six West African (purple) admixed cattle breeds. A significant negative Pearson correlation coefficient was observed for these data ($R = -0.678$; $P = 0.031$) and the grey shaded area shows the 95% confidence interval for the regression line.

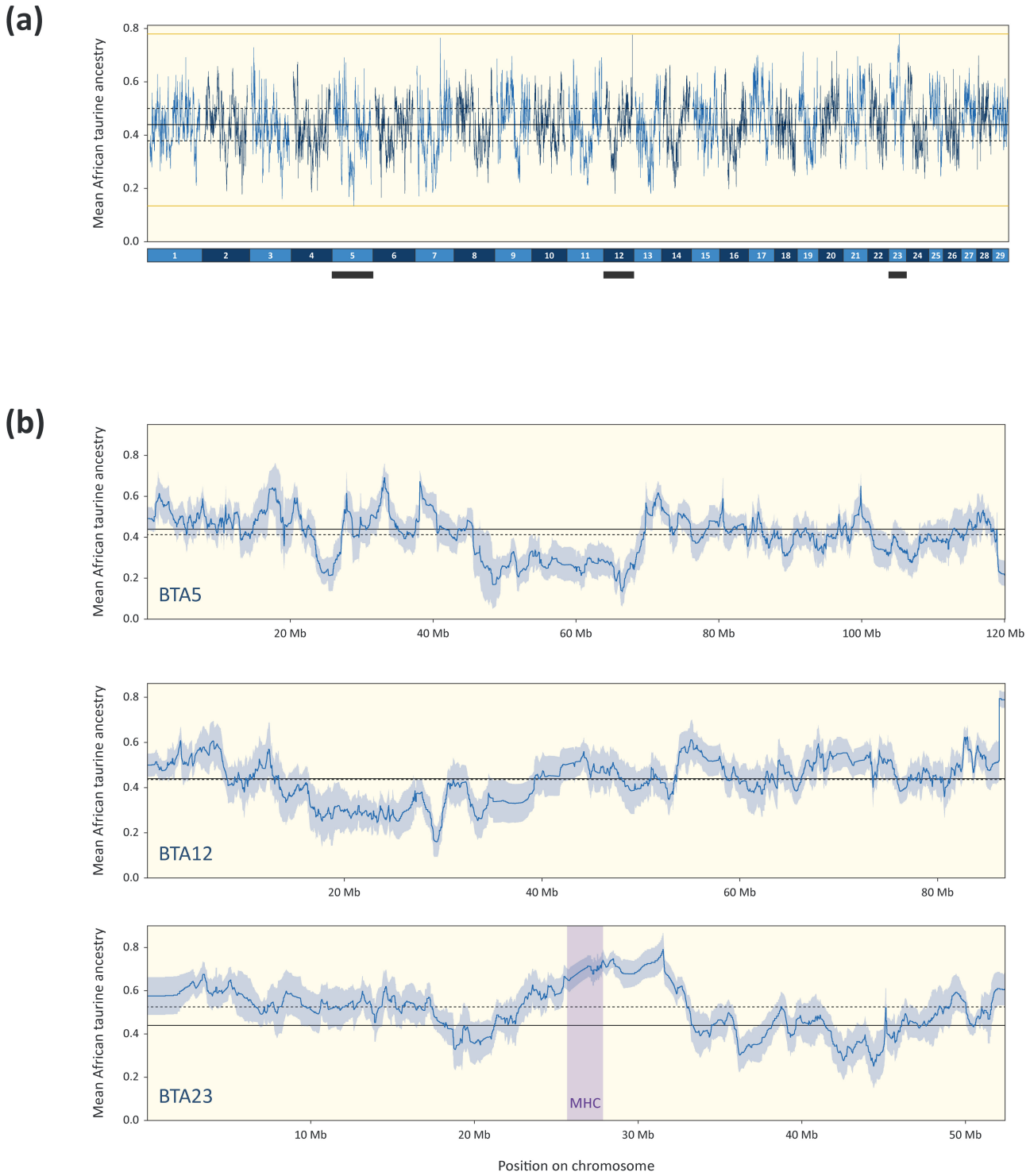


FIGURE 8 Mean African taurine local ancestry across all ten African admixed breeds shown genome-wide and individually for bovine chromosomes 5, 12 and 23 (BTA5, BTA12 and BTA23). (a) Genome-wide mean African taurine local ancestry with a solid black line indicating the global mean, dotted black lines bounding the global 95% confidence interval (CI), and orange lines indicating the global minimum and maximum. The three chromosomes shown in panel (b) are indicated with black rectangles. (b) Chromosome-wide plots for BTA5, BTA12, BTA23 with dark blue lines indicating mean African taurine local ancestry for the ten breeds, blue shading indicating the 95% CI, solid black lines indicating the global mean and dashed blue lines indicating the chromosomal means. The location of the major histocompatibility complex (MHC) gene cluster is shown on BTA23.