Genetic variation of Nigerian cattle inferred from maternal and paternal genetic markers (#50178)

First submission

Guidance from your Editor

Please submit by **14 Jul 2020** for the benefit of the authors (and your $200 publishing discount) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the materials page.

5 Figure file(s)
12 Table file(s)
3 Other file(s)

Custom checks

DNA data checks

- Have you checked the authors data deposition statement?
- Can you access the deposited data?
- Has the data been deposited correctly?
- Is the deposition information noted in the manuscript?

Vertebrate animal usage checks

- Have you checked the authors ethical approval statement?
- Were the experiments necessary and ethical?
- Have you checked our animal research policies?

Field study

- Have you checked the authors field study permits?
- Are the field study permits appropriate?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genetic variation of Nigerian cattle inferred from maternal and paternal genetic markers

David H Mauki [Equal first author, 1, 2, 3], **Adeniyi C Adeola** [Equal first author, 1, 2], **Said I Ng'ang'a** [1, 2, 3], **Abdulfatai Tijjani** [4], **Akanbi I Mark** [5], **Oscar J Sanke** [6], **Abdussamad M Abdussamad** [7], **Sunday C Olaogun** [8], **Jebi Ibrahim** [9], **Philip M Dawuda** [10], **Godwin F Mangbon** [11], **Paul S Gwakisa** [12], **Ting-Ting Yin** [1], **Min-Sheng Peng** [Corresp., 1, 2, 3], **Ya-Ping Zhang** [1, 2, 3, 13, 14]

[1] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China

[2] Chinese Academy of Sciences, Sino-Africa Joint Research Center, Kunming, Yunnan, China

[3] University of Academy of Sciences, Kunming College of Life Science, Kunming, Yunnan, China

[4] School of Life Sciences, University of Nottingham, Nottingham, United Kingdom

[5] Ministry of Agriculture and Rural Development, Secretariat, Ibadan, Oyo, Nigeria

[6] Taraba State Ministry of Agriculture and Natural Resources, Jalingo, Taraba, Nigeria

[7] Department of Animal Science, Faculty of Agriculture, Bayero University, Kano, Kano, Nigeria

[8] Department of Veterinary Medicine, University of Ibadan, Ibadan, Oyo, Nigeria

[9] College of veterinary medicine, department of theriogenology, University of agriculture, Makurdi, Makurdi, Benue, Nigeria

[10] Department of Veterinary Surgery and Theriogenology, College of Veterinary Medicine, University of Agriculture Makurdi, Makurdi, Benue, Nigeria

[11] Division of Veterinary Office, Serti, Taraba, Nigeria

[12] Department of Microbiology, Parasitology and Biotechnology/ Genome Science Center, Sokoine University of Agriculture, Morogoro, Tanzania

[13] State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, School of Life Sciences,, Yunnan University, Kunming, Yunnan, China

[14] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, Yunnan, China

Corresponding Author: Min-Sheng Peng
Email address: pengminsheng@mail.kiz.ac.cn

We analyzed the genetic diversity of Nigerian cattle from 119 and 20 mitochondrial DNA (mtDNA) D-loop and Y-chromosomal sequences respectively and compared with global cattle samples. Nigerian cattle can be assigned to 80 haplotypes based on the D-loop sequences and haplotype diversity was 0.985 + 0.005. The network showed two major matrilineal clustering: the dominant cluster (with haplogroup T1, 98.3% constituting the Nigerian cattle) together with other African cattle; and that composed of Eurasian cattle. Our findings also show the presence of haplogroup T3, 1.7% in Nigerian cattle suggesting low European/West Asian cattle maternal influence. Paternal analysis indicates only Asian zebu haplogroup Y3 influence with several haplotypes in Nigerian cattle. There was no signal of maternal genetic structure in Nigerian cattle population, which may likely suggest the extensive genetic intermixing within the country. The absence of *Bos indicus* maternal signal in Nigerian cattle is attributable to vulnerability bottleneck of mtDNA lineages and concordance with the view of male zebu genetic introgression in African cattle. Our study provides insight into Nigerian cattle genetic diversity and population history in West Africa.

Genetic variation of Nigerian cattle inferred from maternal and paternal genetic markers

David H. Mauki[*†‡#], Adeniyi C. Adeola[*†#], Said I. Ng'ang'a[*†‡], Abdulfatai Tijjani[§], Akanbi I. Mark[¶], Oscar J. Sanke[ψ], Abdussamad M. Abdussamad[æ], Sunday C. Olaogun[Ж], Jebi Ibrahim[€], Philip M. Dawuda[€], Godwin F. Mangbon□, Paul S. Gwakisa[ʒ], Ting-Ting Yin[*], Min-Sheng Peng[*†‡], Ya-Ping Zhang[*†‡ρ]

[*]State Key Laboratory of Genetic Resources and Evolution& Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China.

[†]Sino-Africa Joint Research Center, Chinese Academy of Sciences, Kunming, China.

[‡]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China.

[§]School of Life Sciences, University of Nottingham, Nottingham, United Kingdom.

[¶]Ministry of Agriculture and Rural Development, Secretariat, Ibadan, Nigeria.

[ψ]Taraba State Ministry of Agriculture and Natural Resources, Jalingo, Nigeria.

[æ]Department of Animal Science, Faculty of Agriculture, Bayero University, Kano, Nigeria.

[Ж]Department of Veterinary Medicine, University of Ibadan, Ibadan, Nigeria.

[€]Department of Veterinary Surgery and Theriogenology, College of Veterinary Medicine, University of Agriculture Makurdi, Makurdi, Nigeria.

[]Division of Veterinary Office, Serti, Nigeria.

[ɜ]Sokoine University of Agriculture, Department of Microbiology, Parasitology and Biotechnology/ Genome Science Center, Morogoro, Tanzania.

[β]State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, School of Life Sciences, Yunnan University, Kunming, China.

[ķ]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

#These authors contributed equally to this work.

Corresponding Author:

Min-Sheng Peng[*†‡], Ya-Ping Zhang[*†‡βķ]

[*]State Key Laboratory of Genetic Resources and Evolution& Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China.

[†]Sino-Africa Joint Research Center, Chinese Academy of Sciences, Kunming, China.

[‡]Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, China.

[β]State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, School of Life Sciences, Yunnan University, Kunming, China.

[k]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

Email address: corresponding_ Min-Sheng Peng_pengminsheng@mail.kiz.ac.cn and Ya-Ping Zhang_zhangyp@mail.kiz.ac.cn

Abstract

We analyzed the genetic diversity of Nigerian cattle from 119 and 20 mitochondrial DNA (mtDNA) D-loop and Y-chromosomal sequences respectively and compared with global cattle samples. Nigerian cattle can be assigned to 80 haplotypes based on the D-loop sequences and haplotype diversity was 0.985±0.005. The network showed two major matrilineal clustering: the dominant cluster (with haplogroup T1, 98.3% constituting the Nigerian cattle) together with other African cattle; and that composed of Eurasian cattle. Our findings also show the presence of haplogroup T3, 1.7% in Nigerian cattle suggesting low European/West Asian cattle maternal influence. Paternal analysis indicates only Asian zebu haplogroup Y3 influence with several haplotypes in Nigerian cattle. There was no signal of maternal genetic structure in Nigerian cattle population, which may likely suggest the extensive genetic intermixing within the country. The absence of *Bos indicus* maternal signal in Nigerian cattle is attributable to vulnerability bottleneck of mtDNA lineages and concordance with the view of male zebu genetic introgression in African cattle. Our study provides insight into Nigerian cattle genetic diversity and population history in West Africa.

Keywords: mtDNA D-loop, Y-chromosome, genetic diversity, Nigerian cattle, West Africa

INTRODUCTION

The modern domestic cattle were initially domesticated about 10,000 years ago from two putative domestication centers, the Near East for *Bos taurus* and the Indian Sub-continent for *B. indicus* (*Loftus et al., 1994a*). In Africa, the modern domestic cattle were probably introduced at different times starting with *B. taurus* circa 7,000 BP and later *B. indicus* circa 4,000 BP from their putative centers of domestication (*Meghen, MacHugh & Bradley, 1994*). However, the latter were rapidly introduced around ~699 - 640 AD by the Arab traders following the death of the Prophet (*Bradley et al., 1998*). Penetration of the predominant African taurine to West Africa was amid 4,000 BP (*Marshall & Hildebrand, 2002*). According to *Meghen, MacHugh & Bradley. (1994)*, the post-introduction of the zebu cattle led to spread of zebus sporadically in Western Africa circa 1,400 years ago from eastern Africa hypothesized to be the original entry point of zebus in Africa (*Gifford-Gonzalez & Hanotte, 2011*). So far, the origin of African cattle is still under strong debate due to conflicting archaeological and genetic evidence (*Hanotte et al., 2002*), unravelling of any possible clues is still at large and with very crucial pressure.

Nigeria, a country in West Africa harbors quite a number of cattle including taurines, zebus, (*Rege, 1999*) and their crossbreds (*Loftus et al., 1994a*). Apart from being a valued source of meat, milk, skin, and wool, Nigerian cattle are observed as important idols in ceremonial rituals and utilized for drafting and ploughing during farming. Acquisition of the current knowledge regarding the genetic status of Nigerian cattle is crucial for conservation and utilization of their

genetic resources. Although there are some genetic diversity studies on West African cattle, particularly in Nigeria ~~and have~~ investigated a few populations with limited sample size (*Bradley et al., 1994; Loftus et al., 1994a; Loftus et al., 1994b; Bradley et al., 1996; Perez-Pardal et al., 2018*). This could imply that the actual extent of the genetic diversity of Nigerian cattle still remains enigma.

To ~~decipher~~ the genetic diversity in Nigerian cattle, we employed the use of both mtDNA and Y-chromosomal markers, which have been widely used in assessing the diversity and genogeoraphic structure of many domestic animals (*Ramirez et al., 2009; Wang et al., 2014*). In this study we therefore evaluated variation in mtDNA D-loop and Y-chromosome of 119 and 20 Nigerian cattle samples respectively. Due to the nature of husbandry management in most of African countries, the assessment of genetic variation in Nigerian cattle was conducted based on their sample~~d~~ locations such as from North Western (Zamfara, Kano, Katsina, Kaduna, and Sokoto States); North Eastern (Taraba, and Plateau States) and Western (Oyo State) regions of Nigeria (Fig. 1a).

MATERIALS AND METHODS

Ethical considerations

All experimental procedures in the present study were performed in accordance to Research Guidelines for the Institutional Review Board of Kunming Institute of Zoology, Chinese Academy of Sciences (SMKX2017009).

Sampling and data collection

A total of 139 Nigerian cattle individuals (119 females and 20 males) were sampled from farmer's herds in eight different States in Nigeria as follows (Fig. 1a); Kaduna State (n=19 females; n=3 males), Kano State (n=4 females; n=2 males), Katsina State (n=4 females; n=2 males), Sokoto State (n=27 females; n=4 males), Mambilla plateau in Taraba State (n=35 females; n=2 males), Zamfara State (n=8 females; n=2 males), Oyo State (n=1 female; n=2 males) and Jos city in Plateau State (n=21 females; n=2 males). During sample collection, to obtain genetically unrelated cattle from the eight states in Nigeria, ecological and geographical perspectives were considered. This included randomly sampling of at most two animals per household and only from those households located approximately 0.5 km further apart. Farmers were also interviewed on the pedigree information of their animals prior to carrying out blood collection. Blood samples were kept in 95% ethanol at room temperature before transportation to the laboratory. Samples were stored at 4^{o}C for immediate use, or at -80^{o}C for later use.

Yak (*Bos grunniens*: Accession no. MN398192, *Huang et al., 2020*) and a *B. indicus* sample from West Asia (Accession no. EU177868, *Achilli et al., 2008*) were used as outgroups in phylogenetic tree and haplotype network analyses respectively. A total of 420 sequences based on 636 bp (of the D-loop region) that included 119 individuals from Nigeria, 76 from Europe, 16 from West Asia, and 209 from other African countries (31 Egyptian cattle, 16 Mozambiquan cattle, 126 Ethiopian cattle, 34 Nguni cattle from South Africa, and two additional samples from Nigeria) were used in mtDNA analyses (more information in Figure 1 legend part b; Table S1). For Y-chromosome analyses only few representative samples of the three major haplogroups of cattle were used as detailed in the supplementary Table S2.

DNA extraction, PCR and sequencing

Genomic DNA was extracted from ~5ml of blood following phenol-chloroform method (Sambrook & Russell, 2001). We amplified 636-base pair of the D-loop region of mtDNA, using primers constructed from L27712 D-loop sequence for both forward and reverse primers (Table S3) (*Loftus et al., 1994*). The mtDNA amplification and sequencing reactions were carried out in a total of 25µl PCR reaction mixture using ~40 ng of mtDNA, 10 pmol of each primer, 2.5 mMdNTPs and 5 units of Takara Taq DNA polymerase in a 10pmol TIAGEN reaction buffer containing 1.5 mM $MgCl_2$. Amplifications were carried out in a thermocycler for at least 35 cycles as follows: 95°C for 5 min, 94°C for 45 sec, 58°C for 30 sec, 72°C for 1 min 30 sec, and a final extension of 7 min at 72°C. The quality and confirmation check were performed using 2% agarose gel and visualization under UV transilluminator. The PCR products were purified using EXO/SAP enzymes and finally sequenced using Sanger sequencing approach.

We have sequenced 286 bp of the bovine Y-chromosome (*Ginja et al., 2009*) in 20 Nigerian cattle (Table S2 and S3; GenBank accession numbers: xxx-xxx) in order to identify polymorphic sites in the X-degenerate male specific regions of the bovine Y-chromosome (MSY) (*Chang et al., 2013*). In particular, an intron 10 region of the Y-linked zinc finger protein Y (ZFY) gene (*Ginja et al., 2009*; Table S3) was used ~~to generate primers~~ for the amplification and sequencing of all 20 samples with the length of 286 bp. We applied the same PCR conditions used in the amplification of mtDNA with exception of the annealing temperature was 53 °C. The quality control checks were carried out and all sequences were assembled in SeqMan Lasergene package in DNASTAR software.

Data analysis

Sequences check and alignment

147 The assembled DNA sequences of 139 individuals were exported into MEGA X ver 10.1.7
148 software (*Kumar et al., 2018*) for alignment with other cattle populations mined from Genbank
149 for both mtDNA D-loop (Table S1 and S4) and ZFY Y-chromosomal markers (Table S2).
150 Multiple sequence alignments of the D-loop region and ZFY gene of the Y-chromosome were
151 carried out using CLUSTAL W package (*Thompson, Higgins & Gibson, 1994*) integrated in the
152 MEGA software. All sites containing alignment gaps were excluded from the analysis. Our data
153 involved amplification and sequencing of the 636 -base pair (bp) mtDNA D-loop region and 286
154 bp ZFY gene of 119 and 20 Nigerian cattle samples respectively. Variations in the D-loop region
155 were detected by assembling all forward and reverse sequences against the reference *B. taurus*
156 mtDNA sequence [(GenBank accession no. V00654); (*Anderson et al., 1982*)] using SeqMan
157 Lasergene package in DNASTAR software. Variations in the Y-chromosome of the X
158 degenerate region located at intron 10 of ZFY gene were also determined. The alignments and
159 assembly of the Y-chromosome sequences were carried out similarly following the protocol used
160 by *Ginja, Telo da Gama & Penedo, (2009)* where sequences from distantly related species were
161 used as deployed by *Gotherstrom et al. (2005)* but using *Bison bison*, *B. frontalis*, and *B.*
162 *grunniens* sequences (*Verkaar et al., 2004*). The polymorphic sites among all Nigerian
163 individuals were identified using a *B. taurus* reference genome (GenBank Accession no.
164 AF241271; *Lawson et al., 2002*).

165 **Genetic diversity and haplogroup classification**

166 MitoToolPy_Linux (*Peng et al., 2015*) was used to determine haplogroup distribution across
167 Africa by analyzing a 240 bp fragment length of the D-loop region of mtDNA involving 609
168 cattle samples in total (including the 119 Nigerian cattle sequenced in this study [(Fig. 1a; Table
169 S4); (reference data were retrieved from: *Loftus et al., 1994a; Bradley et al., 1996; Troy et al.,*

2001; Beja-Pereira et al., 2006; Dadi et al., 2009; Bonfiglio et al., 2012; Horsburgh et al., 2013; Olivieri et al., 2015)] . We used DnaSP v5 (*Librado & Rozas, 2009*) to determine the haplotypes in 420 mtDNA cattle sequences and the assignment of bovine Y-chromosome haplogroups for Nigerian cattle. ~~The extent of g~~enetic diversity, was assessed by using Arlequin v3.5 (*Excoffier & Lischer, 2010*) and expressed in terms of the number of haplotypes (H) and polymorphic sites (PS), haplotype diversity (HD), nucleotide diversity (π), the mean number of nucleotide differences (Df) and their standard deviations (SD) estimated across all African populations used in this study. Notably, the comparisons of the genetic diversity estimates were considered for only those populations with sample size above 5.

Phylogenetic tree analyses

To investigate the evolutionary relationship of Nigerian cattle with other cattle samples mined from GenBank (Table S1 and S2), the same version of MEGA software was used to construct a rooted neighbor-joining (NJ) phylogenetic tree (*Saitou & Nei 1987*) using the Maximum Composite Likelihood evolutionary distance approach *(Tamura, Nei & Kumar, 2004*) and bootstrap test was employed at 1000 replications so as to assess the confidence of each node (*Felsenstein, 1985*). To further visualize the genetic relationships between the haplotypes and identifying the number of unique mtDNA D-loop haplogroups present in the 420 dataset, the median-joining (MJ) network (*Bandelt, Forster & Rohl, 1999*) was generated by using Network v4.6 software (www.fluxus-engineering.com).

Population genetic structure and demographic dynamic profiles

To infer the matrilineal genetic variation within populations, among populations, and groups of populations, analysis of molecular variance (AMOVA) was carried out following 50,000

permutations in Arlequin v3.5 software. The analysis was conducted for Nigerian cattle at various hierarchical levels *viz* the Nigerian cattle as a single cluster, Nigerian cattle vs the other African countries but also vs cattle from Europe and West Asia. The levels of significance in each hierarchical cluster tested were evaluated using F_{ST} parameter at a significant *P* level of 0.05.

To ~~unravel~~ the population dynamics and demographic patterns of Nigerian cattle population, mismatch distribution patterns were estimated (*Rogers & Harpending, 1992*) with respect to their geographical regions for North West and North East. The chi-square test of goodness of fit and Harpending's raggedness index "*r*" (*Harpending, 1994*) statistics were also calculated to assert the significance of the deviations of the sum of squares differences (*SSD*) observed from the simulated model of demographic expansions determined by 1,000 coalescent simulations. Demographic statistical parameters for Tajima's *D* (*Tajima, 1989*) and Fu's F_S (*Fu, 1997*) were also estimated by using Arlequin v3.5 software to further complement the mismatch distribution patterns.

RESULTS

Variation in mtDNA sequences and genetic diversity

In this study we evaluated variations in the mtDNA D-loop of 119 Nigerian cattle together with 301 global cattle sequences based on 636 bp from Egypt, Ethiopia, Mozambique, South Africa, Europe and West Asia available in ~~the~~ GenBank. The sequences generated in this study have been deposited in the GenBank with accession numbers MT362777 – MT362895. There were 153 variable sites scored in all 420 cattle samples that defined 275 haplotypes (Table S5) and 80

of them assigned to Nigerian cattle sequenced in this study (Table 1). Most of the Nigerian cattle in the current study possess unique haplotypes (80%) and the remaining ones were shared with other African and European cattle. The lowest level of haplotype diversity (0.983 ± 0.009) was observed in cattle from North East while the highest (0.984 ± 0.007) was observed in the North Western region. Estimated haplotype diversity (HD) across all Nigerian individuals was 0.985+0.005 (Table 1). This value observed is lower than the haplotype diversity of Egyptian and Mozambican cattle populations but was higher compared to Ethiopian and South African cattle.

Haplogroup classification and phylogenetic trees using mtDNA

The haplogroup distribution across Africa shows ~~that~~ majority of African cattle are of *B. taurus* T1 the widely known *B. taurus* haplogroup for African cattle (Fig. 1a, Table S4). This was also evidently confirmed when considering Nigerian cattle analyzed in this study where majority of them were classified into haplogroup T1 (98.3%) (Table S1 and S4). Interestingly ~~enough~~, two of the Nigerian individuals were detected ~~with~~ haplogroup T3 ~~at rather very low~~ frequency of 1.7%, which could imply low influence of European/West Asian cattle, probably have been brought about by a way of interbreeding with exotic commercial breeds in the recent past. Apart from Nigeria, other African countries in the Northern part of Africa particularly Egypt, Libya and Morocco ~~were~~ also ~~detected with the~~ other types of *B. taurus* haplogroups T2, T3 and Q1. No trace of *B. indicus* lineage ~~was~~ observed in Nigerian cattle samples sequenced in this study.

To obtain further insights into the haplotype relationships, the network analysis (Fig. 1b) and phylogenetic tree (Fig. S1) were constructed using the 636 bp sequences of 119 Nigerian cattle and 301 other sequences retrieved from the GeneBank. The phylogenetic tree ~~has indicated~~ two major lineages of cattle, *B. taurus* (T1, T2, T3, T5, and Q1) and *B. indicus* (I1 and I2) lineages as

237 expected. All Nigerian cattle have been placed together with the rest of taurine cattle individuals
238 separately from the zebu cattle. The MJ network depicted similar pattern where haplogroup I1
239 for zebu lineage (considered herein as the outgroup) was separated from all the taurine
240 haplogroups. Furthermore, the network revealed two major clustering: the first cluster showed
241 grouping of Nigerian cattle with African and European cattle; while the second cluster did not
242 contain Nigerian individuals. Nigerian cattle in the first cluster exhibit a star-like pattern, which
243 signifies a signature of population expansion.

244 **Population genetic structure and historical demographic dynamics**

245 AMOVA analysis incorporating the eight populations from Nigeria showed that more than 99%
246 of the total genetic variation present in Nigerian cattle occurred within individuals (Table S6).
247 Furthermore, when analyzing the genetic variation between Nigerian cattle and other cattle
248 populations AMOVA showed that, 62 to 93% of the total variation between Nigerian cattle and
249 other African cattle populations occurs within individuals, with the highest variation observed
250 between Nigerian and Mozambican cattle (93.46%). Generally, all attempts to explore the
251 differentiation between Nigeria and other populations were attributable to within-population
252 variance ($P < 0.05$) with ~~exceptional~~ of variation among populations within groups for Nigerian
253 and European cattle (F = 0.07375, $P = 0.00004$). The F_{ST} distance values between most Nigerian
254 cattle sub-populations were low (Table S7). However, it was considerably high between Nigerian
255 cattle and other African cattle populations showing significant genetic variation between them at
256 $P < 0.05$ with ~~exceptional for~~ cattle populations between Ethiopia and Mozambique (FST = -
257 0.00548, $P = 0.57129$). Our AMOVA results complemented by genetic distance estimates
258 suggest general absence of maternal genetic structuring in Nigerian cattle sub-populations, likely
259 due to extensive genetic intermixing within the country.

To elucidate the demographic dynamics of Nigerian cattle, mismatch distribution patterns, for each geographical region in Nigeria were assessed (Fig. S2). The mismatch distribution patterns were unimodal, however the pattern deviated significantly from expected under a null hypothesis model of either spatial or demographic expansion due to significant values obtained for Sum of Squared deviation (SSD) and Harpending's Raggedness index (HRI) (Table S8). The significant values for SSD and HRI indicate a bad goodness of fit test, which does not support the scenario of population expansion. The values for Tajima's D -2.315 ($P < 0.05$) and Fu's F_S statistics -26.203 ($P < 0.05$) on the other hand were both negative and significant indicating an agreement of recent population growth and expansion respectively. Cattle population from Western region part of Nigeria showed Tajima's D value of 0 possibly due to only 3 samples having being used, but could indicate that this population in this region evolved as per mutation drift equilibrium with no evidence of selection.

Genetic polymorphism of cattle Y- chromosome and haplogroup distribution

In supplementary Figure 3, we provide detailed information on haplogroup distribution based on Y-chromosomal markers in Africa. Our results show that all Nigerian cattle cluster in haplogroup Y3 solely a *B. indicus* haplogroup due to similar mutations (Table 2, Fig. S3). Notably, in addition to previously reported mutations that further classify zebu cattle into haplotype Y3 families as pointed out by *Chen et al. (2018)* and *Perez-Pardal et al. (2018)*, we also show new mutations in Nigerian cattle possibly not previously reported. These include those SNPs observed between g.784 and g.805bp (Table 2). Multiple sequence alignment with haplotypes defining the three major haplogroups of cattle (Y1, Y2 and Y3) (*Nijman et al., 2008*; *Ginja, Telo da Gama & Penedo, 2009* and Table S2) revealed the existence of seven mutations including mutations A>G, T>C and T>G that distinguish Y3a (or Y_C) and Y3b (or $Y3_A$, $Y3_B$) haplotype

283 families (Table 2) grouping Nigerian cattle into several zebu Y3 sub-haplotype families (Table
284 3). Majority of the Nigerian cattle shared haplotype Hap_7Y3 together with cattle from Asia.
285 Our findings indicate non-evidence influence of European Y1 or Y2 haplogroups in Nigerian
286 cattle. The phylogenetic tree showed distinct Nigerian individuals with some of them clustering
287 together with distantly related ~~species~~ illustrating possible ancient auroch contribution into the
288 gene pool of Nigerian cattle (Fig. S4).

289 The overall genetic diversity for Nigerian cattle measured by the haplotype diversity and the
290 mean number of pairwise difference is 1.000 ± 0.016 and 1.679 ± 1.027 respectively (Table 3).
291 The genetic diversity detected in some African cattle populations/breeds as per *Perez-Pardal et*
292 *al. (2018)* and *Ginja, Telo da Gama & Penedo, (2019)* based on Y-chromosomal markers are
293 shown in Table S9. Both of these studies employed low sampling coverage for the assessment of
294 Nigerian cattle genetic diversity compared to the sampling coverage used in this study.

295

296 **DISCUSSION**

297 In this study we examined the genetic variation across eight Nigerian cattle populations, which
298 represent samples from one country in West Africa. The results based on mtDNA D-loop
299 revealed 80 haplotypes from 119 Nigerian cattle sequences, which showed a haplotype diversity
300 of 0.985 ± 0.005. This value is lower compared to those observed in Egyptian cattle (H = 1.0 ±
301 0.008) in the study conducted by *Olivieri et al. (2015)*. These findings signify a high level of
302 maternal genetic variation in Nigerian cattle. AMOVA on the other hand suggests a general
303 absence of maternal genetic structuring in Nigerian cattle that may have been as a result of recent
304 past extensive genetic intermixing within the country. Studies show that genetic intermixing is

caused by ~~rigorous~~ transportation of domestic animals like goats, cattle, and sheep from place to place as a result of their valuable resource for economic trade or cultural exchange (*Tarekegn et al., 2018*). We observed similar genetic diversity pattern based on Y-chromosome analysis where the haplotype diversity detected was higher compared to the haplotype diversity 0.800 ± 0.213 reported by *Perez-Pardal et al. (2018)* for Nigerian cattle possibly due to a single breed population having been used, albeit it was similar to haplotype diversity of cattle from India and Central Asia (*Perez-Pardal et al., 2018*). Generally, in comparison with other cattle populations in Africa, our study indicated the highest genetic diversity of 1.000 ± 0.016 (Table 3) with the lowest in Zebu_Peul breed population from Burkina Faso (Table S9).

The present-day zebu-like cattle in West Africa were products of crossbreeding events between the West African taurine cattle such as N'Dama and imported zebu cattle from Asia (*MacHugh et al. 1997*). We portrayed the genetic relationships among Nigerian cattle by using MJ network (Fig. 1b) and NJ tree (Fig. S1). We observed all Nigerian cattle to have an exclusive taurine specific genetic haplogroup T, notwithstanding some of the Nigerian cattle displayed some features of *B. indicus* morphological resemblance such as the presence of the humps. This scenario is in accordance with similar observations by *Loftus et al. (1994a)*. The phylogenetic tree indicated two major lineages of cattle: *B. taurus* haplogroup T lineage (containing T1, T2, T3, T5 haplogroups) and the novel Q1 haplogroup and *B. indicus* haplogroup I lineage with exclusive I1 and I2 haplogroups ~~expectedly~~ (*Achilli et al., 2008*). All Nigerian cattle have been placed together with the rest of taurine cattle individuals separated from the zebu cattle with high frequency of T1 (98.3%) dominating in African cattle (*Troy et al., 2001*) and very low frequency of T3 (1.7%) likely due to recent interbreeding events from commercial exotic breeds. Numerous mtDNA phylogeny studies cluster the African cattle within the European clade such as that

328 observed in this study (*Loftus et al., 1994a*; *Loftus et al., 1994b*). This reflects that the
329 predominant African cattle were solely of taurine origin. The MJ network depicted similar
330 pattern where haplogroup I1 for zebu lineage (considered herein as the out-group) was separated
331 from the taurine lineage haplogroups. The network also revealed two major clustering: the first
332 cluster showed grouping of Nigerian cattle with African and European cattle; while the second
333 cluster did not contain Nigerian individuals. The complete absence of Asian zebu mtDNA in
334 Nigerian cattle samples suggests that crossbreeding events were mainly through the imported
335 Asian male zebus (*Loftus et al., 1994a*; *Loftus et al., 1994b*; *Bradley et al., 1996*). Nonetheless,
336 this scenario could have been attributed by the continued adoption of trypanotolerant breeds that
337 probably led to total loss of the *B. indicus* mtDNA lineages or its vulnerability towards
338 population bottleneck. This has also been observed elsewhere by several studies particularly
339 cattle from North-East Asia (*Mannen et al., 2004*). *Bradley et al. (1994)* and *Perez-Pardal et al.*
340 *(2018)* further stressed availability of male zebu Y-haplotypes in West African cattle
341 extrapolating the importation of probably only male zebus into this region.

342 The mtDNA and Y-chromosome analyses conducted in this study have confirmed that Nigerian
343 cattle are an influence of both taurine and indicine lineages from female and male genetic
344 contributions respectively. Previously, mtDNA studies (*Loftus et al., 1994a*) have founded that
345 African cattle are of taurine background. However, studies conducted by using Y-chromosomal
346 markers (*Bradley et al., 1994*) founded that African cattle had been genetically introgressed with
347 zebu cattle from Asia. Based on Y-chromosomal studies, it is observed that majority of African
348 cattle belong to haplogroup Y3 (a *B. indicus* haplogroup) especially for cattle in East and Central
349 Africa (Fig. S3). However, the frequency of zebu haplogroup Y3 is seen to decrease gradually as
350 you move towards West and South of Africa likely due to the presence of *B. taurus* haplogroups

Y1 and Y2 in these regions. A study by *Decker et al. (2014)* had previously demonstrated similar observations using autosomal single nucleotide polymorphisms (SNPs) markers. Some studies have further indicated that *B. indicus* haplogroup Y3 contain several haplotype families such as Y3b (or $Y3_A$, $Y3_B$) and Y3a (or Y_C) (*Chen et al., 2018; Perez-Pardal et al., 2018*), with $Y3_A$ and $Y3_B$ being the dominant haplogroups in East and West African cattle respectively. Similarly, we have observed the same, however, we point out to a possibility of having several other distinct zebu individuals in Nigerian cattle that may have not been previously investigated (Table 3, Fig. S4). Moreover, the phylogenetic analysis of evolutionary relationship infers the possibility of local ancient wild aurochs introgression that may have had occurred in Africa post introduction of zebus adding to an effect of the current several Y3 strains present in Nigeria or rather the legacy of local adaptation. These findings shade some possible novel insights regarding Nigerian cattle perspectives of their origins. The classification of Nigerian cattle into both as taurine and as indicine confirms the ongoing hybridization process of cattle species (*B. taurus* and *B. indicus*) that occurred since ancient time spontaneously or rather through recent breeding events. In the light of this argument with respect to Nigerian cattle in West Africa comes as most farmers in this region preferred zebu of male lineage due to their massive huge size but also being resistant to rinderpest that taurines are not, even though them being less vigorous towards trypanosomiasis, a disease prevalent in tsetse regions of both West and Central Africa that taurines are resistant to (*Grigson, 1991; Ibeagha-Awemu et al., 2004*). Because of their ability to withstand rinderpest disease and the huge size of the bulls for meat, Asian zebus (trypanosusceptible) were preferentially selected for domestication that subsequently led to the modern West African cattle as hybrid product of both the predominant local African taurines (trypanotolerant) (or the earliest introduced taurines from Eurasia) and Asian zebus.

374 Inference of demographic profiles and population dynamics were extrapolated by mismatch
375 distribution patterns. We observed a unimodal mismatch distribution pattern in North East, North
376 West and the entire Nigerian cattle population. This pattern suggests the evidence of population
377 growth and expansion of Nigerian cattle. Although the goodness of fit model displayed contrary
378 results to the support of evidence for population expansion, the negative values in Tajima's *D*
379 and FU's *FS* show excess of low and high frequency polymorphisms indicative of population
380 size increase from a recent population bottleneck or genetic hitchhiking. Moreover, the MJ
381 network showed a star-like pattern that further signifies the scenario of population expansion of
382 Nigerian cattle that could have had emerged from a single founder sample or several founder
383 samples that had very little variation between them. Similar findings were observed by previous
384 studies, where they detected unimodal mismatch distribution patterns as evidence for
385 demographic expansion (*Bradley et al., 1996*; *Mannen et al., 2004*).

386

387 **CONCLUSIONS**
388 This study reported the current genetic status and some possible new insights about the origin of
389 cattle in West Africa using Nigerian samples from matrilineal and patrilineal perspectives. High
390 level of maternal and paternal genetic diversity was observed in Nigerian cattle, with lack of
391 phylogeographic structure possibly due to human~~ly~~ mediated interventions that usually enhance
392 severe intermixing as a result of random mating. Despite the ensue of the two contrary findings,
393 we tend to believe that Nigerian cattle would be more of indicine background due to their
394 morphological appearance matching that of zebus observed during sampling and also as
395 confirmed by the Y-chromosome analysis conducted in this study. Therefore, deep population

396 genetic autosomal studies are required using high-throughput technologies to provide
397 comprehensive demographic history of cattle in West Africa.

398

399 **Authors' contributions**

400 Y.-P.Z., D.H.M, A.C.A. and M.-S.P. led the project, and designed and conceived the study.
401 D.H.M. and A.C.A. prepared the manuscript. M.-S.P. and Y.-P.Z. revised the manuscript.
402 D.H.M., S.I.N. and A.C.A performed the data analysis. A.C.A., A.I.M., O.J.S., S.C.O., J.I. and
403 G.F.M. performed sampling. D.H.M. carried out experiments. All authors revised and approved
404 the final manuscript.

405

406 **Acknowledgments**

407 We appreciate all those who assisted in the study. This work was supported by the Sino-Africa
408 Joint Research Center, Chinese Academy of Sciences (SAJC201611) and the Animal Branch of
409 the Germplasm Bank of Wild Species, Chinese Academy of Sciences (the Large Research
410 Infrastructure Funding). The Chinese Academy of Sciences President's International Fellowship
411 Initiative provided support to A.C.A (2018FYB0003). D.H.M and S.I.N acknowledge the
412 support of the Chinese Academy of Sciences-The World Academy of Sciences (CAS-TWAS)
413 President's Fellowship Program for Doctoral students.

414

415 **Competing interests**

The authors declare that they have no competing interests.

References

Achilli A, Olivieri A, Pellecchia M, Uboldi C, Colli L, Al-Zahery N, Accetturo M, Pala M, Hooshiar Kashani B, Perego UA, Battaglia V, Fornarino S, Kalamati J, Houshmand M, Negrini R, Semino O, Richards M, Macaulay V, Ferretti L, Bandelt HJ, Ajmone-Marsan P, Torroni A. 2008. Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Current Biology* **18**(4):157–158.

Anderson SMHL, De Bruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG. 1982. Complete sequence of bovine mitochondrial DNA conserved features of the mammalian mitochondrial genome. *Journal of molecular biology* **156**(4):683–717.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**:37–48.

Beja-Pereira A, Caramelli D, Lalueza-Fox C, Vernesi C, Ferrand N, Casoli A, Goyache F, Royo LJ, Conti S, Lari M, Martini A, Ouragh L, Magid A, Atash A, Zsolnai A, Boscato P, Triantaphylidis C, Ploumi K, Sineo L, Mallegni F, Taberlet P, Erhardt G, Sampietro L, Bertranpetit J, Barbujani G, Luikart G, Bertorelle G. 2006. The origin of European cattle: evidence from modern and ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**:8113–8118.

435 Bradley DG, Loftus RT, Cunningham P, Machugh DE. 1998. Genetics and domestic cattle
436 origins. *Evolutionary Anthropology* **6**:79–86.

437 Bradley DG, MacHugh DE, Cunningham P, Loftus RT. 1996. Mitochondrial diversity and the
438 origins of African and European cattle. *Proceedings of the National Academy of Sciences of*
439 *the United States of America* **93**:5131–5.

440 Bradley DG, MacHugh DE, Loftus RT, Sow RS, Hoste CH, Cunningham EP. 1994. Zebu-
441 taurine variation in Y chromosomal DNA: a sensitive assay for genetic introgression in west
442 African trypanotolerant cattle populations. *Animal Genetics* **25**:7–12.

443 Bonfiglio S, Ginja C, De Gaetano A, Achilli A, Olivieri A, Colli L, Penedo MCT. 2012. Origin
444 and spread of Bos taurus: new clues from mitochondrial genomes belonging to haplogroup
445 T1. *PloS one* **7**(6).

446 Chang TC, Yang Y, Retzel EF, Liu WS. 2013. Male-specific region of the bovine Y
447 chromosome is gene rich with a high transcriptomic activity in testis development.
448 *Proceedings of the National Academy of Sciences of the United States of America*
449 **110**:12373–12378.

450 Chen N, Cai Y, Chen Q, Li R, Wang K, Huang Y, Hu S, Huang S, Zhang H, Zheng Z. Song W.
451 2018. Whole-genome resequencing reveals world-wide ancestry and adaptive introgression
452 events of domesticated cattle in East Asia. *Nature communications* **9**(1):1–13.

453 Dadi H, Tibbo M, Takahashi Y, Nomura K, Hanada H, Amano T. 2009. Variation in
454 mitochondrial DNA and maternal genetic ancestry of Ethiopian cattle populations. *Animal*
455 *genetics* **40**(4):556–559.

456 Decker JE, Mckay SD, Rolf MM, Kim J, Molina Alcala A, Sonstegard TS, Hanotte O,
457 Gotherstrom A, Seabury CM, Praharani L, Babar ME, Correia De Almeida Regitano L,
458 Yildiz MA, Heaton MP, Liu WS, Lei CZ, Reecy JM, Saif-Ur-Rehman M, Schnabel RD,
459 Taylor JF. 2014. Worldwide patterns of ancestry, divergence, and admixture in
460 domesticated cattle. *PLoS Genettics* **10**(3).

461 Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform
462 population genetics analyses under Linux and Windows. *Molecular Ecology Resources*
463 **10**:564–7.

464 Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap.
465 *Evolution* **39**:783–791.

466 Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking
467 and background selection. *Genetics* **147**:915–925.

468 Gifford-Gonzalez D, Hanotte O. 2011. Domesticating animals in Africa: Implications
469 of genetic and archaeological findings. *Journal of World Prehistory* **24**:1–23.

470 Ginja C, Telo da Gama L, Penedo MCT. 2009. Y chromosome haplotype analysis in Portuguese
471 cattle breeds using SNPs and STRs. *Journal of Heredity* **100**(2):148–157.

472 Götherström A, Anderung C, Hellborg L, Elburg R, Smith C, Bradley DG, Ellegren H. 2005.
473 Cattle domestication in the Near East was followed by hybridization with aurochs bulls in
474 Europe. *Proceedings of the Royal Society B: Biological Sciences* **272**(1579):2345–2351.

475 Grigson C. 1991. An African Origin for African Cattle? Some Archaeological Evidence. *The*
476 *African Archaeological Review* **9**:119–144.

477 Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW, Rege JEO. 2002. African pastoralism:
478 genetic imprints of origins and migrations. *Science* **296**(5566):336–339.

479 Harpending HC. 1994. Signature of ancient population growth in a low-resolution mitochondrial
480 DNA mismatch distribution. *Human Biology* **66**:591– 600.

481 Hiendleder S, Lewalski H, Janke A. 2008. Complete mitochondrial genomes of Bos taurus and
482 Bos indicus provide new insights into intra-species variation, taxonomy and
483 domestication. *Cytogenetic and genome research* **120**(1-2):150–156.

484 Horsburgh KA, Prost S, Gosling A, Stanton JA, Rand C, Matisoo-Smith EA. 2013. The genetic
485 diversity of the Nguni breed of African Cattle (Bos spp.): complete mitochondrial genomes
486 of haplogroup T1. *PloS one* **8**(8).

487 Huang C, Zhang Q, Fu D, Basang W, Chu M, Yangla D, Wu X, Ma X, Guo X, Yan P, Liang C.
488 2020. The complete mitochondrial genome sequence and phylogenetic analysis of Sibu yak
489 (*Bos grunniens*). *Mitochondrial DNA Part B* **5**:46–47.

490 Ibeagha-Awemu EM, Jann OC, Weimann C, Erhardt G. 2004. Genetic diversity, introgression
491 and relationships among West/Central African cattle breeds. *Genetics Selection Evolution*
492 **36**(6):673–690.

493 Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary
494 Genetics Analysis across computing platforms. *Molecular Biology and Evolution* **35**:1547–
495 1549.

496 Lai SJ, Liu YP, Liu YX, Li XW, Yao YG. 2006. Genetic diversity and origin of Chinese cattle
497 revealed by mtDNA D-loop sequence variation. *Molecular Phylogenetics and*
498 *Evolution* **38**(1):146–154.

499 Lawson LJ, Hewit GM. 2002. Comparison of substitution rates in ZFX and ZFY introns of sheep
500 and goat related species supports the hypothesis of male-biased mutation rates. *Journal of*
501 *Molecular Evolution* **54**(1):54–61.

502 Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA
503 polymorphism data. *Bioinformatics* **25**:1451–2.

504 Loftus RT, MacHugh DE, Bradley DG, Sharp PM, Cunningham P. 1994a. Evidence for two
505 independent domestications of cattle. *Proceedings of the National Academy of Sciences of*
506 *the United States of America* **91**:2757–61.

507 Loftus RT, MacHugh DE, Ngere LO, Balain DS, Badi AM, Bradley DG, Cunningham EP.
508 1994b. Mitochondrial genetic variation in European, African and Indian cattle populations.
509 *Animal Genetics* **25**:265–71.

510 MacHugh DE, Shriver MD, Loftus RT, Cunningham P, Bradley DG. 1997. Microsatellite DNA
511 variation and the evolution, domestication and phylogeography of taurine and zebu cattle
512 (*Bos taurus* and *Bos indicus*). *Genetics* **146**:1071–86.

513 Mannen H, Kohno M, Nagata Y, Tsuji S, Bradley DG, Yeo JS, Nyamsamba D, Zagdsuren Y,
514 Yokohama M, Nomura K, Amano T. 2004. Independent mitochondrial origin and historical
515 genetic differentiation in North Eastern Asian cattle. *Molecular Phylogenetics and*
516 *Evolution* **32**:539–44.

517 Marshall F, Hildebrand E. 2002. Cattle before crops: the beginnings of Food production in
518 Africa. *Journal of World Prehistory* **16**:99–143.

519 Meghen C, MacHugh DE, Bradley DG. 1994. Genetic characteristics of West African cattle.
520 *World Animal Review* **78**:59–66.

521 Nijman IJ, Van Boxtel DC, Van Cann LM, Marnoch Y, Cuppen E, Lenstra JA. 2008. Phylogeny
522 of Y chromosomes from bovine species. *Cladistics* **24**(5):723–726.

523 Olivieri A, Gandini F, Achilli A, Fichera A, Rizzi E, Bonfiglio S, Battaglia V, Brandini S, De
524 Gaetano A, El-Beltagi A, Lancioni H, Agha S, Semino O, Ferretti L, Torroni A. 2015.
525 Mitogenomes from Egyptian Cattle Breeds: New Clues on the Origin of Haplogroup Q and
526 the Early Spread of Bos taurus from the Near East. *PLoS One* **10**(10)

527 Peng MS, Fan L, Shi NN, Ning T, Yao YG, Murphy RW, Wang WZ, Zhang YP. 2015.
528 DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals.
529 *Molecular Ecology Resources* **15:**1238–42.

530 Pérez-Pardal L, Sánchez-Gracia A, Álvarez I, Traoré A, Ferraz JBS, Fernández I, Patel A. 2018.
531 Legacies of domestication, trade and herder mobility shape extant male zebu cattle
532 diversity in South Asia and Africa. *Scientific reports* **8**(1):1–8.

533 Ramirez O, Ojeda A, Tomas A, Gallardo D, Huang LS, Folch JM, Galman-Omitogun O. 2009.
534 Integrating Y-chromosome, mitochondrial, and autosomal data to analyze the origin of pig
535 breeds. *Molecular biology and evolution* **26**(9):2061–2072.

536 Rege JEO. 1999. The state of African cattle genetic resources 1. Classification framework and
537 identification of threatened and extinct breeds. *Animal Genetic Resources* **25**:1–25.

538 Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise
539 genetic differences. *Molecular Biology and Evolution* **9**:552–69.

540 Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing
541 phylogenetic trees. *Molecular biology and evolution* **4**(4):406–425.

542 Sambrook J, Russell DW. 2001. Molecular cloning. In: C.S.H. Laboratory, ed. *A laboratory*
543 *manual*. New York: Cold Spring Harbor Laboratory Press.

544 Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA
545 polymorphism. *Genetics* **123**:585–95.

546 Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the
547 neighbor-joining method. *Proceedings of the National Academy of Sciences* **101**:11030–
548 11035.

549 Tarekegn GM, Tesfaye K, Mwai OA, Djikeng A, Dessie T, Birungi J, Tutah J. 2018.
550 Mitochondrial DNA variation reveals maternal origins and demographic dynamics of
551 Ethiopian indigenous goats. *Ecology and evolution* **8**(3):1543–1553.

552 Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of
553 progressive multiple sequence alignment through sequence weighting, position-specific gap
554 penalties and weight matrix choice. *Nucleic acids research* **22**(22):4673–4680.

555 Troy CS, MacHugh DE, Bailey JF, Magee DA, Loftus RT, Cunningham P, Chamberlain AT,

Sykes BC, Bradley DG. 2001. Genetic evidence for Near-Eastern origins of European cattle. *Nature* **410**:1088–91.

Verkaar EL, Nijman IJ, Beeke M, Hanekamp E, Lenstra JA. 2004. Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? *Molecular biology and evolution* **21**(7):1165–1170.

Wang GD, Xie HB, Peng MS, Irwin D, Zhang YP. 2014. Domestication genomics: evidence from animals. *Annual Review of Animal Biosciences* **2**:65–84.

Figure legend

Figure 1 Sampling locations of cattle in Nigeria and the network of 420 cattle samples based on 636 bp of the mtDNA D-loop region. (a) Map of cattle sampling locations from Nigeria (Zamfara, Kano, Katsina, Kaduna, and Sokoto from North West; Taraba and Plateau from North East; and Oyo from the West) and the haplogroup distribution based on mtDNA across Africa. (b) Median-joining network of 420 cattle samples constructed by using NETWORK v 4.6 (*Bandelt, Forster & Rohl, 1999*). Reference sequences used for haplotype network analysis included: Europe, n=76 (*Loftus et al., 1994a; Lai et al., 2006; Achilli et al., 2008; Hiendleder, Lewalski & Janke, 2008; Bonfiglio et al., 2012* and accession nos. AF034438-AF034446); West Asia, n=16 (*Achilli et al., 2008*); Egypt, n=31 (*Bonfiglio et al., 2012; Olivieri et al., 2015*); Ethiopia, n=126 (*Dadi et al., 2009; Bonfiglio et al., 2012*); Mozambique, n=16 (Accession nos. JQ684029-JQ684045), South Africa, n=34 (*Horsburgh et al., 2013*) and two additional Nigerian samples mined from GenBank (Accession no. L27731 and L27730). Sizes of the circles are proportional to haplotype frequencies. *m*, refers to number of mutation steps and

578 those not indicated are just one step mutation. Colours indicate the geographical distribution of
579 the sampling locations across Africa, Europe and West Asia as shown by the legend in (b).

580 **Table legend**

581 **Table 1** Genetic diversity of cattle in Africa based on mtDNA D-Loop

582 **Table 2** Mutations that describe the distinction of the three major haplogroups of cattle based on
583 ZFY intron 10 gene. The two genomic sites that separate each distinct haplogroup are in bold.
584 The seven mutations observed in Nigerian cattle (NY3) are italicized.

585 **Table 3** The frequency of haplotypes for Nigeria, Europe and Asia and their distribution based
586 on Y-chromosome at the ZFY intron 10 gene.

587

588

589

Figure 1

Sampling locations of cattle in Nigeria and the network of 420 cattle samples based on 636 bp of the mtDNA D-loop region

(a) Map of cattle sampling locations from Nigeria (Zamfara, Kano, Katsina, Kaduna, and Sokoto from North West; Taraba and Plateau from North East; and Oyo from the West) and the haplogroup distribution based on mtDNA across Africa . (b) Median-joining network of 420 cattle samples constructed by using NETWORK v 4.6 (*Bandelt, Forster & Rohl, 1999*) . Reference sequences used for haplotype network analysis included: Europe, n=76 (*Loftus et al., 1994a; Lai et al., 2006; Achilli et al., 2008; Hiendleder, Lewalski & Janke, 2008; Bonfiglio et al., 2012* and accession nos. AF034438-AF034446); West Asia, n=16 (*Achilli et al., 2008*); Egypt, n=31 (*Bonfiglio et al., 2012; Olivieri et al., 2015*); Ethiopia, n=126 (*Dadi et al., 2009; Bonfiglio et al., 2012*); Mozambique, n=16 (Accession nos. JQ684029-JQ684045), South Africa, n=34 (*Horsburgh et al., 2013*) and two additional Nigerian samples mined from Genbank (Accession no. L27731 and L27730). Sizes of the circles are proportional to haplotype frequencies. *m*, refers to number of mutation steps and those not indicated are just one step mutation. Colours indicate the geographical distribution of the sampling locations across Africa, Europe and West Asia as shown by the legend in (b).

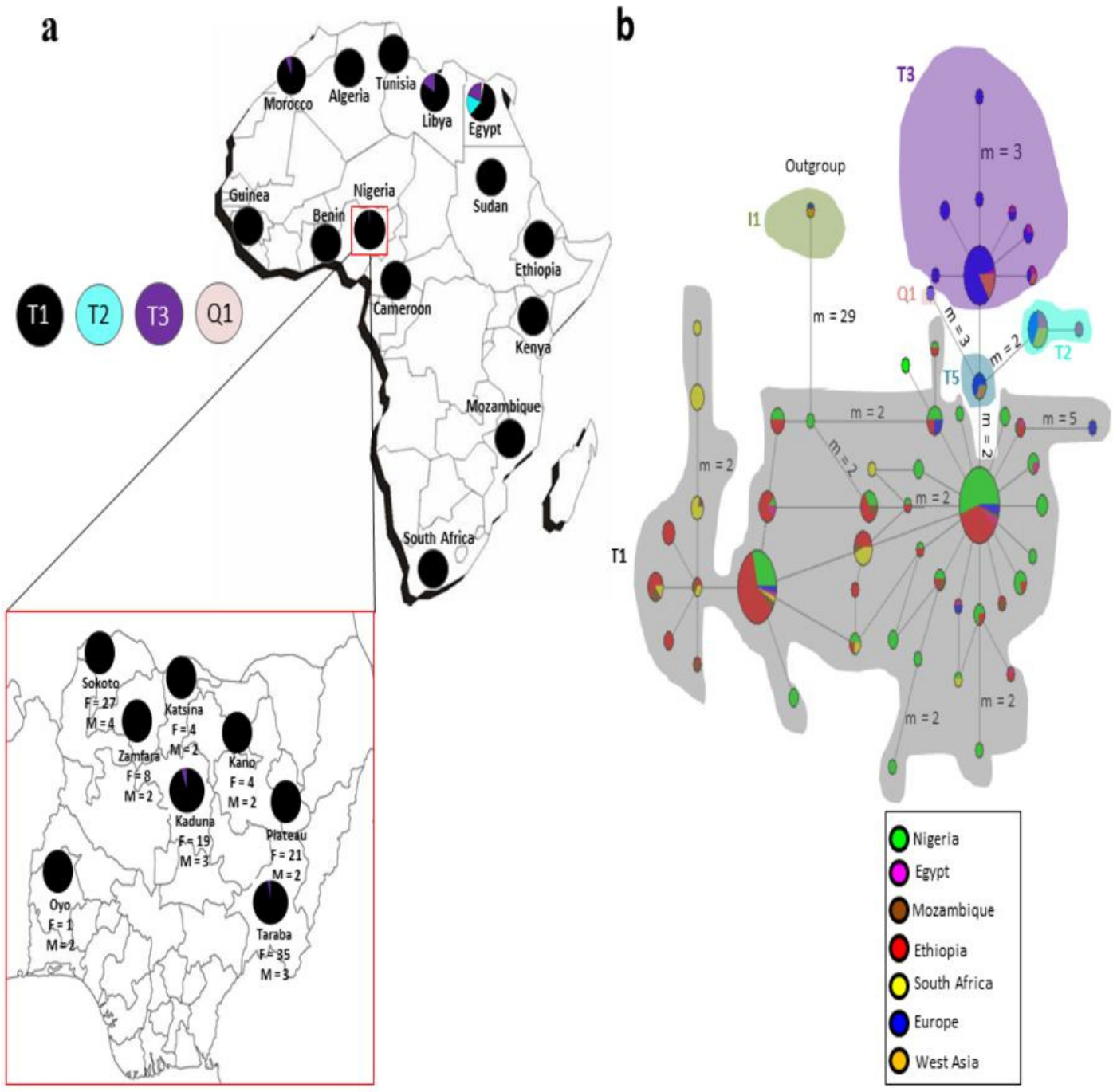
a
Morocco
Algeria
Tunisia
Libya
Egypt
Guinea
Nigeria
Benin
Sudan
Ethiopia
Cameroon
Kenya
Mozambique
South Africa
T1
T2
T3
Q1
Sokoto
F = 27
M = 4
Katsina
F = 4
M = 2
Zamfara
F = 8
M = 2
Kano
F = 4
M = 2
Kaduna
F = 19
M = 3
Plateau
F = 21
M = 2
Oyo
F = 1
M = 2
Taraba
F = 35
M = 3
b
T3
m = 3
Outgroup
I1
m = 29
Q1
m = 3
m = 2
T2
T5
m = 2
m = 5
m = 2
m = 2
m = 2
m = 2
m = 2
T1
Nigeria
Egypt
Mozambique
Ethiopia
South Africa
Europe
West Asia

Table 1(on next page)

Genetic diversity of cattle in Africa based on mtDNA D-Loop

Table 1. Genetic diversity of cattle in Africa based on mtDNA D-Loop

Population	S	PS	H	HD (SD)	π (SD)	Df
1. Nigeria	119	67	80	0.985(0.005)	0.051(0.029)	3.403
(a) North West[1]	62	44	44	0.984(0.007)	0.076(0.044)	3.338
(b) North East[2]	56	52	43	0.983(0.009)	0.067(0.038)	3.462
(c) West[3]	3*	4	3	1.000(0.272)	0.667(0.598)	2.667
2. Egypt	31	46	30	1.000(0.082)	0.010(0.005)	6.391
3. Mozambique	16	18	15	0.992(0.025)	0.180(0.109)	3.233
4. Ethiopia	126	70	83	0.969(0.009)	0.005(0.003)	3.365
5. South Africa	34	17	24	0.961(0.019)	0.006(0.003)	3.606

Note: S = sample size, PS = the number of polymorphic sites, H = the number of haplotypes, HD = haplotype diversity, π = nucleotide diversity, Df = the mean number of nucleotide differences, and SD = standard deviations. [1]Cattle samples from Zamfara, Kano, Katsina, Kaduna and Sokoto States; [2]Samples from Taraba and Jos, Plateau States; [3]Samples from Ibadan, Oyo State. *Two additional downloaded samples from GenBank (Accession no. L27731 and L27730). The estimation of haplotype and nucleotide diversity based on 636 bp mtDNA D-loop sequence was carried out by ARLEQUIN v. 3.5 (Excoffier & Lischer, 2010) software.

Table 2(on next page)

Mutations that describe the distinction of the three major haplogroups of cattle based on ZFY intron 10 gene. The two genomic sites that separate each distinct haplogroup are in bold. The seven mutations observed in Nigerian cattle (NY3) are italicized.

Table 2. Mutations that describe the distinction of the three major haplogroups of cattle based on ZFY intron 10 gene. The two genomic sites that separate each distinct haplogroup are in bold. The seven mutations observed in Nigerian cattle (NY3) are italicized.

	ZFY intron 10		
Haplogroup	**Reference**	**Alternative**	**Publication**
Y1	C GT	C **--**	
Y2	C GT	**C** **GT**	Gotherstrom *et al*. 2005 Ginja *et al*. 2009
Y3	C GT	**T** GT	
Y3	C GT *A* *A* *A* *T* *T* *T* *C*	T GT *G* *T* *C* *A* *G* *C* *T*	
g.784-g.805			This study

Table 3(on next page)

The frequency of haplotypes for Nigeria, Europe and Asia and their distribution based on Y-chromosome at the ZFY intron 10 gene.

Table 3. The frequency of haplotypes for Nigeria, Europe and Asia and their distribution based on Y-chromosome at the ZFY intron 10 gene.

Haplotype	Frequency	Nigeria	Europe	Asia
Hap_1Y3	2	2		
Hap_2Y3	2	2		
Hap_3Y3	1	1		
Hap_4Y3	1	1		
Hap_5Y3	1	1		
Hap_6Y3	1	1		
Hap_7Y3	15	12		3
Hap_8Y2	1		1	
Hap_9Y2	3		3	
Hap_10Y1	1		1	
Total	**28**	**20**	**5**	**3**

Genetic diversity	**H	20
	**PS	10
	**HD (SD)	1.000 (0.016)
	**Df (SD)	1.679 (1.027)

Note: ** = Genetic diversity of Nigerian cattle; H = number of haplotypes, HD = Haplotype diversity; PS = number of polymorphic sites, Df = Mean number of pairwise difference and SD = standard deviations.