GENETIC STRUCTURE AND DIFFERENTIATION AMONG AFRICAN *BOS TAURUS* CATTLE BREEDS

N.Z. Gebrehiwot[1], E.M. Strucken[1], H. Aliloo[1], K. Marshall[2] and J.P. Gibson[1]

[1]Centre for Genetic Analysis and Applications, School of Environmental and Rural Science, University of New England, Armidale, NSW, 2351 Australia
[2]International Livestock Research Institute and Centre for Tropical Livestock Genetics and Health, Nairobi, Kenya

SUMMARY

African taurine cattle populations are widely distributed in humid and sub-humid zones of West and Central Africa. We assessed the genetic structure and differentiation within and across 8 African *Bos taurus* populations: 4 N'Dama populations (N'Dama, N'Dama1, N'Dama2, N'Dama3), Lagunaire, Lagune, Somba, and Baoule. A total of 38k autosomal SNPs were used for principal component analyses (PCA), estimation of pairwise F_{ST} values and within population heterozygosity (F_{IS}), and neighbour-joining (NJ) tree construction. The first PC clearly differentiated Lagune and Lagunaire from N'Dama; PC2 separated Lagunaire, Lagune and one N'Dama population from the rest of taurine breeds; and PC3 separated N'Dama3 from Somba and Baoule. Estimates of pairwise F_{ST} values among the majority of populations ranged from 0.03 to 0.149, indicating low to moderate genetic differentiations, while a high genetic divergence between N'Dama3 and Lagune (F_{ST}=0.178), and N'Dama3 and Lagunaire (F_{ST}=0.168) was observed. No genetic subdivision was found between N'Dama and N'Dama1, and Lagune and Lagunaire. A higher heterozygosity (F_{IS} value from -0.011 to 0.025) was found in N'Dama, N'Dama1, N'Dama2, Lagune, Lagunaire, and Baoule breeds. The NJ tree clearly separated Lagune and Lagunaire as well as Somba and Baoule with a 100% and around 31% bootstrap value, respectively, from the other taurine populations. We highlighted that African taurine populations are diverse and genetic differences between sampling locations exists even within a breed. Therefore, choice of an African taurine breed to anchor African indigenous breeds should be carefully considered.

INTRODUCTION

Taurine cattle are known to have been first domesticated in the Near East and are believed to have been introduced to Africa through present day Egypt (Gifford-Gonzalez and Hanotte 2011). It is thought that the humpless taurine Hamitic Shorthorns and Longhorns arrived in sub-Saharan Africa around 4500 to 4000BP (Payne and Hodges 1995). Nowadays, African taurine cattle breeds have been categorized as: longhorn N'Dama of the far West forest savannah, and 14 humpless shorthorns such as Baoule, Somba and Lagune breeds which are widely distributed in the humid and sub-humid zones of West and Central Africa (Rege 1999). These zones are also known for their endemic trypanosomes, which may affect cattle distributions due to some breeds being tolerant of infection, whilst others are strongly affected by trypanosomes (Berthier *et al.* 2015). There is, however, no comprehensive study showing how genetically diverse African taurine breeds actually are. Up to now, only one African taurine breed, the longhorn N'Dama, has been largely studied and is often considered as a reference breed when other African indigenous breeds are analysed. The present study aimed at assessing genetic structure and differentiation within and among 8 African *Bos taurus* populations.

MATERIALS AND METHODS

SNP data for a total of 130 African *Bos taurus* individuals originating from different locations

of West Africa. These included N'Dama (separated into 4 populations: N'Dama (n=20), N'Dama1 (n=14), N'Dama2 (n=14), and N'Dama3 (n=17), sampled in Guinea, Cote d'Ivoire, South-Eastern Burkina Faso and South-Western Burkina Faso, respectively), Lagune (n=20) from Benin, Baoule (n=20) from Burkina Faso, Somba (n=20) from Togo, and Lagunaire (n=5) from "West Africa". Pooled *Bos indicus* (n=105) and European *Bos taurus* (n=100) were used as reference populations.

Animals and genotypes were sourced from the Bovine HapMap consortium (777k) and from Decker *et al.* (2014, 50k). After quality control, the genotypes of 777k SNPs and 50k SNPs datasets reduced to 735k SNPs and 45k SNPs, respectively, of which 38,556 SNPs were in common and used in this study.

Principal component analyses (PCA) were performed using a genomic relationship matrix (GRM) according to the first method of VanRaden (2008). The PCA were carried out with and without the reference breeds. To estimate population differentiation, pairwise F_{ST} was calculated according to Weir and Cockerham (1984). The degree of inbreeding was inferred from the F_{IS} coefficient calculated according to Nei (1977). Allele frequencies in the African taurine populations were used to construct a neighbour-joining tree. Bootstrapping (1000 replicates) was carried out to assess the strength of support for the internal nodes. All data analyses were performed using the R software (R Core Team, 2018).

RESULTS AND DISCUSSION

The first two PCs obtained from the analysis of African taurine and the reference populations explained 79% and 16.7% of the total variation in the GRM, respectively, and differentiated the African taurine, *Bos indicus*, European taurine from each other (Figure 1a). In comparison, the first 5 PCs obtained from the analysis of African taurine cattle populations explained a total of 65.22% of the variation in the GRM (Figure 1b and 1c).
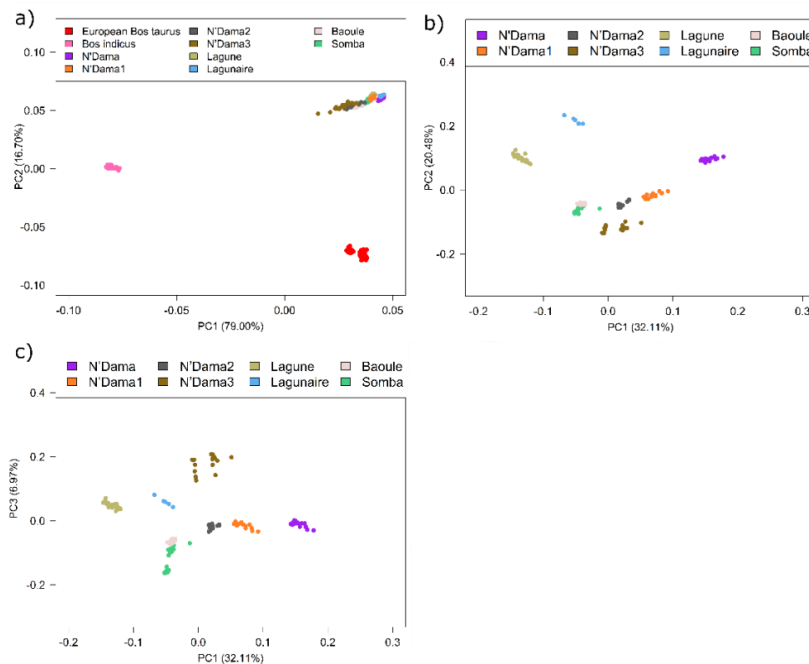

Figure 1. Plot PC1 vs PC2 for reference and African *Bos taurus* populations (a) Plot of PC1 vs PC2 (b) and PC1 vs PC3 (c) for only African *Bos taurus* populations

The first PC accounted for 32.11% of the total genetic variation and clearly differentiated Lagune and Lagunaire from N'Dama; whereas PC2 explained 20.48% of the total variation and separated Lagunaire, Lagune and N'Dama from the rest of the African taurine populations (Figure 1b). The third PC contributed 6.97% of the total variation and separated N'Dama3 from Somba and Baoule (Figure 1c). The majority of Somba and Baoule animals grouped tightly together, however, one Somba outlier was detected (Figure 1b and 1c). The N'Dama3 clusterd in two separate groups (Figure 1b and 1c), which might indicate that they were sampled from two different villages or sub-locations. Both Figure 1b and 1c showed that N'Dama1 and N'Dama2 clustered near each other, with N'Dama1 closer towards N'Dama.

Estimated F_{IS} values showed less inbreeding than expected under Hardy-Weinberg equilibrium for N'Dama1, N'Dama2, N'Dama3, and Baoule (Table 1). N'Dama3 showed the highest F_{IS}=-0.109 which confirms the separation of this breed into two clusters in the PC plots. The relatively high amount of within breed genetic variation indicated by F_{IS} values for all breed samples indicates a valuable reservoir of genetic diversity for future breeding endeavours and a viable target for conservation.

The pairwise F_{ST} estimates showed a high genetic divergence between N'Dama3 and Lagune (F_{ST} =0.178), and N'Dama3 and Lagunaire (F_{ST}=0.168), indicating a clear genetic difference among these breeds. Moderate pairwise F_{ST} values of 0.061 to 0.149 were observed between the majority of the African taurine populations (Table 1). The minimum possible genetic distance between populations is zero. Negative values of the estimate of genetic distance, F_{ST}, can arise by random sampling and can be interpreted as zero genetic distance between populations. Lagune and Lagunaire as well as N'Dama and N'Dama1 are inferred to have zero genetic distance.

Table 1. Estimated F_{IS} values (diagonal) within and pairwise F_{ST} (off-diagonals) between the African taurine cattle populations

	N'Dama	N'Dama1	N'Dama2	N'Dama3	Lagune	Lagunaire	Baoule	Somba
N'Dama	0.011	-0.017	0.030	0.099	0.145	0.133	0.085	0.067
N'Dama1		-0.005	0.030	0.100	0.149	0.135	0.086	0.067
N'Dama2			-0.011	0.083	0.127	0.109	0.061	0.044
N'Dama3				-0.109	0.178	0.168	0.112	0.095
Lagune					0.004	-0.006	0.123	0.100
Lagunaire						0.025	0.112	0.083
Baoule							-0.021	0.039
Somba								0.049

The neighbor-joining tree is consistent with the PCA and F_{ST} results (Figure 2). The African taurine populations were separated into two clades: Lagune and Lagunaire, which F_{ST} indicated can be regarded as one population, separated from the other breeds. Within the second clade, N'Dama3 separated on 100% of bootstraps from the other breeds, again confirming our previous findings. Somba and Baoule as well as the remaining N'Dama populations clustered together in 100% of bootstraps, with N'Dama and N'Dama1 clusters being consistent with the F_{ST}, which showed that they are effectively a single population. The fact that Somba and Baoule formed overlapping clusters in the PCA plots (Figure 1b and 1c) and had relatively low F_{ST} coupled with the geographical proximity of their sampling areas suggests that they might be considered sub-populations of the same breed.
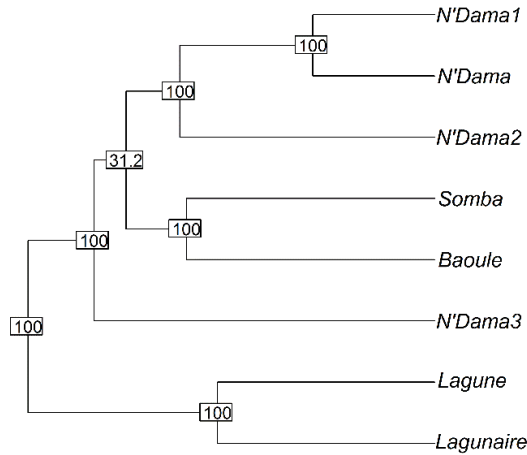
Figure 2. Phylogenetic tree constructed by the neighbour-joining method based on allele frequencies among the African *Bos taurus* cattle populations

The PCA (Figure 1a) showed that N'Dama2 and N'Dama3 spread towards the pooled *Bos indicus* reference breeds, showing that they are not pure African taurine breeds. In admixture analyses, results not shown here, where the African *Bos taurus* breeds were analysed along with many other African indigenous breeds, *Bos indicus* and European *Bos taurus* controls, N'Dama2 and particularly N'Dama3 appeared to have a small proportion of admixture with *Bos indicus* most likely coming from local zebu breeds (Gebrehiwot *et al*. in preparation). This likely explains N'Dama3 being clustered quite separately from other N'Dama samples.

CONCLUSIONS

Our study provides an insight into the genetics of the African *Bos taurus* breeds. The current research indicates that Lagune and Lagunaire, as well as N'Dama and N'Dama1 can be considered as single populations, respectively. The results presented are important for the design of conservation, improvement and breed management strategies of West African *Bos taurus* breeds.

REFERENCES

Berthier D., Peylhard M., Dayo G.-K., Flori L., Sylla S., Bolly S., Sakande H., Chantal I. and Thevenon S. (2015) *PloS One* **10**: e0126498.
Decker J.E., McKay S.D., Rolf M.M., Kim J., Alcalá A.M., Sonstegard T.S., Hanotte O., Götherström A., Seabury C.M. and Praharani L. (2014) *PLoS Genet.* **10**: e1004254.
Gebrehiwot N.Z., Strucken E.M., Aliloo H., Marshall K. and Gibson J.P. In preparation.
Gifford-Gonzalez D. and Hanotte O. (2011) *J. World Prehist*. **24**: 1.
Nei M. (1977) *Ann. Hum. Gene*t. **41**: 225.
Payne W. and Hodges. J. (1995) ' Tropical cattle: origins, breeds and breeding policy ' Blackwell Science, Oxford.
Rege J. (1999) *Anim. Genet. Res.* **25**: 1.
VanRaden P.M. (2008) *J. Dairy Sci.* **91**: 4414.
Weir B.S. and Cockerham C.C. (1984) *Evol.* **38**: 1358