

## Genetic diversity and relationships among indigenous Mozambican cattle breeds

I. Bessa<sup>1,2</sup>, I. Pinheiro<sup>1</sup>, M. Matola<sup>3</sup>, K. Dzama<sup>5#</sup>, A. Rocha<sup>4</sup> and P. Alexandrino<sup>1,2</sup>

<sup>1</sup> CIBIO, Research Centre in Biodiversity and Genetic Resources, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

<sup>2</sup> Department of Zoology and Anthropology, Faculty of Sciences, University of Porto, 4099-002 Porto, Portugal

<sup>3</sup> Faculty of Veterinary Science, University Eduardo Mondlane, CP 257, Maputo, Moçambique

<sup>4</sup> ICBAS, Institute of Biomedical Sciences Abel Salazar, University of Porto, 4050-Porto, Portugal

<sup>5</sup> Department of Animal Sciences, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

---

### Abstract

Three indigenous Mozambican cattle breeds, namely the Angone, Landim and Bovino de Tete were characterized using six proteins, 13 autosomal microsatellite loci and one Y-specific microsatellite locus (INRA124). The Mashona breed from Zimbabwe was also studied to elucidate the origin of the Bovino de Tete cattle. Expected mean heterozygosity ranged from 0.46 - 0.50 in the proteins and from 0.66 - 0.69 in the microsatellites. Population genetic variability was relatively high when compared to other African breeds. Only 4.5% of the total genetic variation could be attributed to the differences among the breeds.  $D_A$  genetic distances and principal component analysis suggest that Mozambican breeds occupy an intermediate position between Indian Zebu and African taurine cattle. The genetic contribution from Indian Zebu, estimated by mR and average percentage of Zebu diagnostic alleles, was highest in the Angone breed and lowest in the Landim breed. The indicine Y-specific allele was fixed in the Angone breed (classified as Zebu), was found in 62% of the Bovino de Tete breed and was absent in the Landim breed (classified as Sanga). The hybrid nature of these breeds was also revealed by using an admixture model to infer population structure. Cluster analysis correctly assigned individuals to their rightful populations with probabilities ranging from 0.96 to 0.98, using prior population information. The results support the hypothesis of the Bovino de Tete cattle being a result of crossbreeding between Sanga and Zebu breeds. This study presents the first extensive information on the genetic diversity and relationships among Mozambican cattle breeds and with other breeds from different continents.

---

**Keywords:** Proteins, microsatellite, Angone, Landim, Bovino de Tete, introgression

# Corresponding author. Email: kdzama@sun.ac.za

### Introduction

Almost 70% of the Mozambican human population is dependent on agriculture as a source of income and domesticated cattle are tightly interwoven in the cultural and economic structure of African life. The classification of African cattle breeds has been complicated by historical crossbreeding that has taken place on a large scale between the original domesticated Taurine cattle of Africa and invading herds of Zebu cattle from Asia (Epstein, 1971; Marshall, 2000). Hybridization between the two subspecies occurred in varying degrees over most of Africa and nowadays the continent represents a biological mosaic of the two morphologies. Mozambican cattle breeds are an example of this process. Three indigenous breeds exist in this country: Angone, Landim and Bovino de Tete. Based on their morphology, the Angone is classified as a Zebu type, while the Landim is classified as a Sanga breed (Epstein, 1971). The Bovino de Tete has been generally considered a Sanga breed but its origin is unclear. Some authors consider that this breed was derived from crossbreeding between the Landim and the Angone breeds (Rege & Tawah, 1999) whilst others suggest that it was derived from the Mashona breed (Sanga type), which inhabits the border of Zimbabwe with Mozambique (Morgado, 2000). Bovino de Tete cattle are found in low numbers within a narrow stretch of Mozambique along the western border up to the southwest in the Tete Province (Alberro, 1983; Morgado, 2004).

Mozambican cattle breeds are managed through traditional husbandry and have multipurpose uses (Rocha *et al.*, 1991; Regé & Tawah, 1999; Morgado, 2000). Contrary to the highly selected commercial

domestic cattle, the Mozambican breeds have been bred since their origin under extensive conditions, with low levels of artificial selection. As a consequence, these breeds have become adapted to local environmental conditions and are resistant to a number of endemic subtropical diseases (Van Rensburg, 1953; Dionisio & Syrstad, 1990; Scholtz *et al.*, 1991; Carvalheira *et al.*, 1995; Hanks & Pereira, 1998). However, extensive crossbreeding, replacement with exotic breeds, and social and environmental disasters have placed these indigenous bovines at risk of extinction. The re-introduction of cattle has been taking place in Mozambique on a large scale, after the huge loss of the national herd during the years of civil war between 1975 and 1994 (Hanks & Pereira, 1998).

There is little information on the genetic diversity of these populations or on the extent to which their descendent gene pools have been perturbed by introgression and migration. Therefore, the evaluation of the population genetic structure of Mozambican cattle is of major interest to obtain the necessary elements for supporting breed conservation and the improvement of breeding programmes. The objectives of this study were: i) to assess the genetic diversity of Mozambican cattle breeds; ii) to evaluate the levels of crossbreeding and introgression by Zebu cattle, iii) to determine the genetic relationships among Mozambican breeds and in comparison with other European, African and Asian cattle breeds and iv) to elucidate the origin of Bovino de Tete cattle.

The resultant information should provide a basis for future conservation and management of the Mozambican cattle breeds.

## Materials and Methods

Blood samples from 166 presumably unrelated individuals of Mozambican cattle breeds, Angone, Landim and Bovino de Tete, were collected from several representative herds and breeding stations, located in the regions of Angonia, Maputo and localities surrounding Tete city, respectively. Additionally, 40 samples from the Mashona breed were collected in Zimbabwe near the Mozambique border.

The samples were separated into red blood cells and serum. Six polymorphic proteins were surveyed: albumin (*ALB*), transferrin (*TF*), carbonic anhydrase II (*CAII*), vitamin D binding protein (*GC*), haemoglobin B (*HBB*) and peptidase B (*PEPB*). The later enzyme was run on starch gels, whilst the other proteins were analysed by isoelectric focusing on acrylamide gels (Leal, 1998). No samples for protein analysis were available for the Mashona population.

DNA was extracted from peripheral blood using Proteinase K digestion followed by extraction with standard Phenol-chloroform or high-salt procedures (Sambrook *et al.*, 1995). Thirteen dinucleotide microsatellites (*INRA5*, *INRA35*, *INRA37*, *INRA23*, *HEL1*, *HEL5*, *HEL9*, *ILSTS5*, *ILSTS1*, *ETH225*, *ETH3*, *ETH152* and *BM2113*) were amplified by PCR. The selection of markers was based on the FAO panel (<http://www.projects.roslin.ac.uk/>). A Y-indicine/taurine diagnostic microsatellite marker (*INRA124*) (Vaiman *et al.*, 1994) was also amplified in 13 males of each Mozambican breed and 20 of the Mashona breed. For all microsatellites, approximately 50 ng of genomic DNA was included with 200  $\mu$ M dNTP, 1.5  $\mu$ M MgCl<sub>2</sub>, 0.5 U *Taq* polymerase, 5 pmol of each primer and 1x reaction buffer (10 mM Tris-HCl pH 8.3) in a 25  $\mu$ L reaction volume. For each reaction, an initial 4 min denaturation at 94 °C, was followed by 35 cycles of 30 s at 94 °C, 30 s at 54 - 60 °C, 45 s at 72 °C and a final extension step at 72 °C for 5 min. Products were checked in 2% agarose gel in TBE buffer. PCR products were separated by electrophoresis in 8% denaturing polyacrylamide gels, followed by silver staining (Bassam *et al.*, 1991). Allele identification and counting were made directly from the gels. To avoid misidentification of alleles, eight reference samples with known genotypes were used in the gels. These samples were provided by D. MacHugh (College of Life Sciences, University College Dublin, Ireland) and were previously reported in MacHugh *et al.* (1997).

Allele frequencies were determined by direct allele counting. Genetic variability for the microsatellite and protein loci was measured by estimating observed ( $H_o$ ) and unbiased expected ( $H_e$ ) heterozygosities (Nei, 1978) as well as the mean number of alleles per locus (MNA) and the number of private alleles ( $N_a$ , alleles found in only one breed) with the GENEPOP program v. 3.1 (Raymond & Rousset, 1995). Deviations from Hardy-Weinberg equilibrium (HWE) for each locus, in each population, were investigated using an exact test based on Markov chain iteration as implemented in GENEPOP v. 3.1 (Raymond & Rousset, 1995). Critical significance levels for multiple testing were corrected following the sequential Bonferroni procedure (Rice, 1989).

Unbiased  $F$ -statistics (Wright, 1951; Weir & Cockerham, 1984) and associated significance levels were calculated for allozymes and microsatellite data using the program FSTAT 2.9.3.1 (Goudet, 2002). The significance of  $F$ ,  $\theta$ ,  $f$  estimates was determined by permuting alleles within the total population, genotypes within the total population and alleles within samples, respectively, with 5000 permutations (Goudet, 2002). The  $F_{ST}$  values among pairs of breeds were also calculated and the indicative P-values were adjusted by a Bonferroni procedure (Rice, 1989).

For the microsatellite markers, the programme STRUCTURE was used to identify the  $K$  clusters of the genetic origin among individuals and to verify correct assignment of individuals to their rightful breeds (Pritchard *et al.*, 2000; Falush *et al.*, 2003). The programme implements a Bayesian Markov Chain Monte Carlo (MCMC) approach that uses the model-based clustering method for inferring population structure using multilocus genotypes. The admixture model was used in cluster analysis and the model with prior population information in assignment verification. Posterior probability values for  $K$  (log likelihood; lnL) were estimated assigning priors from  $K = 2$  to  $K = 4$ , the latter corresponding to the number of breeds included in the analysis, at 100 000 MCMC repetitions and a 100 000 sample burn-in period. For each value of  $K$ , five independent replicates of MCMC were conducted in order to verify the consistency of the estimates across runs. We chose the value of  $K = 2$ , which showed the highest lnL value.

The estimator mR was used to analyse the Zebu introgression in the Mozambican and Mashona populations with the programme ADMIX 1.0 (Bertorelle & Excoffier, 1998). This coefficient was chosen since it presented a lower variance compared to the other two estimators (mY, mC) provided by the programme. Three Zebu cattle breeds from the Indian subcontinent were considered as the putative Asian *B. indicus* parental population (Hariana, Sahiwal, Tharparker). The N'Dama cattle from central and western Africa were considered as the putative *B. taurus* source population. Allele frequencies of the presumed populations were taken from MacHugh *et al.* (1997). Only seven microsatellites were found in common between this work and ours (*HEL1*, *ILSTS5*, *HEL5*, *ETH225*, *ETH152*, *BM2113* and *ILSTS1*). However, these microsatellites demonstrated to be powerful tools for breed differentiation (MacHugh *et al.*, 1997; Jordana *et al.*, 2003; Egito *et al.*, 2007) and, therefore, should allow a reliable estimate of the Zebu introgression in the Mozambican and Mashona breeds. Additionally, the average percentage of Zebu diagnostic alleles per locus was calculated in each breed. These alleles were present in six microsatellites (*INRA5*, *HEL1*, *ILSTS5*, *ETH225*, *ETH152*, *ILSTS1*) in agreement with the data obtained by MacHugh *et al.* (1997) and Cymbron *et al.* (2005). Thirteen diagnostic alleles of Zebu ancestry were defined in these six microsatellites by their presence at a high frequency in Asian Zebu populations and their absence in the N'Dama and European breeds.

The  $D_A$  genetic distance (Nei *et al.*, 1983) was computed separately for microsatellite data (12 loci), proteins (5 loci) and for a combined data set including all 17 loci using the DISPAN computer package (Ota, 1993). In addition, data from the Mashona breed and from published literature were used to investigate the genetic relationships between the Mozambican populations and other breeds from Africa, Europe and Asia using allele frequencies from seven commonly employed microsatellite loci (*ETH225*, *ETH152*, *HEL1*, *HEL5*, *ILSTS1*, *BM2113* and *ILSTS5*) (MacHugh *et al.*, 1997). A Neighbor-joining (NJ) tree (Saitou & Nei, 1987) based on  $D_A$  distances was constructed using the algorithm implemented in the PHYLIP package (Felsenstein, 2004) and the reliability of each node was estimated by bootstrapping (after 1000 replicates) over loci. Genetic relationships within Mozambican cattle and other breeds were also explored by principal components analysis (PCA) using the STATISTICA programme (StatSoft, 1995).

## Results

A summary of the amount of genetic variation found within and between population samples at allozymes and microsatellite loci is presented in Table 1. As expected, much greater allelic variation was detected at the microsatellite loci in comparison to the allozymes. A total of 120 alleles were found across six protein markers (20 alleles) and 13 microsatellite markers (100 alleles). In these breeds, an intermediate pattern of allele size distribution between African taurine and Zebu breeds was observed, all presenting alleles from the two source populations. The mean number of alleles (MNA) per locus for protein data was 3 in all breeds and for microsatellite loci varied between 5.9 and 6.4. Mean  $H_o$  and  $H_e$  values at protein loci were similar for Mozambican breeds –  $H_o$  values ranged from 0.469 to 0.475 and  $H_e$  from 0.458 to 0.472. At microsatellites, the  $H_o$  ranged from 0.587 to 0.631 and  $H_e$  from 0.643 to 0.688.

**Table 1** Levels of genetic diversity calculated for allozymes and microsatellite loci in Mozambican cattle breeds

Population	N <sup>1</sup>	Allozymes					Microsatellites				
		MNA <sup>2</sup>	Ho <sup>3</sup>	He <sup>4</sup>	F <sub>IS</sub> <sup>5</sup>	Na <sup>6</sup>	MNA	Ho	He	F <sub>IS</sub>	Na
Angone	51	3.0	0.469 (0.180)	0.498 (0.192)	0.029	1	6.3	0.616 (0.161)	0.688 (0.126)	0.107	8
Landim	66	3.0	0.476 (0.127)	0.472 (0.160)	0.045	1	6.0	0.587 (0.207)	0.661 (0.207)	0.114	1
Bovino de Tete	49	3.0	0.475 (0.180)	0.458 (0.163)	-0.043	1	6.4	0.604 (0.171)	0.673 (0.164)	0.105	7
Mashona	40	---	---	---	---	---	5.9	0.631 (0.192)	0.643 (0.175)	0.020	5

<sup>1</sup> Number of individuals sampled; <sup>2</sup> Mean number of alleles per locus; <sup>3</sup> Observed heterozygosity; <sup>4</sup> Unbiased heterozygosity; <sup>5</sup> Within-population inbreeding estimate; <sup>6</sup> Number of private alleles within analysed breeds. Standard deviations are given in parenthesis.

Several private alleles were found. However, with the exception of allele 91bp at the *ILSTS1* locus in the Landim breed (10%), all the other private alleles had frequencies of lower than 2%. The Angone breed showed the highest number of private alleles (n = 8).

All populations were in Hardy-Weinberg equilibrium, after sequential Bonferroni correction, except for the Angone (P = 0.004,  $\alpha$  = 0.05, k = 13) and the Mashona (P = 0.004,  $\alpha$  = 0.05, k = 12) for locus *HEL5* and the Landim (P = 0.004,  $\alpha$  = 0.05, k = 13) for locus *ILSTS1* that showed a deficiency of heterozygotes.

Results of *F*-statistics for each of the 19 loci analysed in the studied breeds are shown in Table 2. For microsatellite markers, we observed on average a deficit of heterozygotes (*f*) of 10.5% (P < 0.001) for each one of the analysed breeds and 14.6% (P < 0.001) in the whole population. Eight of these loci contributed significantly to the deficit of heterozygotes within populations (F<sub>IS</sub>), while all microsatellites, except *INRA23* and *ETH225*, affected significantly the global heterozygote deficit (F<sub>IT</sub>). For protein data, only *HBB* contributed significantly (P < 0.05) to the observed deficit. Levels of differentiation per locus ( $\theta$ ) were relatively low and ranged from 0.007 to 0.131 in microsatellite loci and from 0.004 to 0.094 in protein loci. Given all loci, approximately 4.5% of the total genetic variation could be explained by differences between breeds, and the remaining 95.5% to differences among individuals. When the Mashona breed was included in the analysis, values of *f*,  $\theta$  and *F* were similar to corresponding values observed in Mozambican cattle only (Table 2).

Significant pairwise differentiation (F<sub>ST</sub>) occurred between the Mozambican breeds after Bonferroni corrections (P < 0.01) (Table 3). However, when comparing these breeds with Mashona breed only the pair Mashona/Landim presented significant genetic differentiation (P < 0.01).

Bayesian analysis using STRUCTURE (Pritchard *et al.*, 2000) showed that the samples had the highest probability of representing two clusters (K = 2). Based on an assignment of individuals without previous population information, K = 2 gave a split of the Angone/Bovino de Tete breeds to one cluster and Landim/Mashona to the other cluster. This pattern corresponds effectively to a zebu (Angone and Bovino de Tete) and a taurine cluster (Landim and Mashona) (Figure 1).

**Table 2** *F*-statistics analysis for each locus in Mozambican cattle breeds (six protein loci and 13 microsatellites)

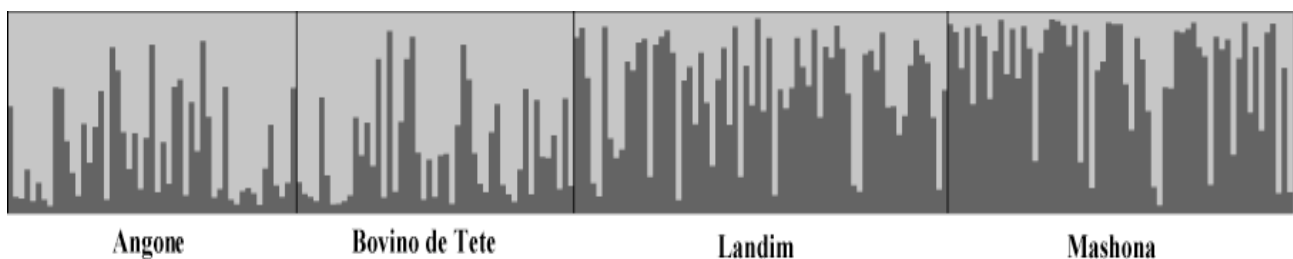
Allozymes	Locus	$F_{IS}$ ( $f^1$ )	$F_{ST}$ ( $\theta^2$ )	$F_{IT}$ ( $F^3$ )
	<i>HBB</i>	0,086*	0,094*	0,172**
	<i>CAII</i>	-0,154	-0,005	-0,161
	<i>ALB</i>	0,026	0,004	0,030
	<i>TF</i>	0,028	0,016*	0,044
	<i>GC</i>	0,021	0,041*	0,062
	<i>PEPB</i>	-0,030	0,050*	0,022
	Mean	0,009	0,036*	0,045*
Microsatellites				
	<i>INRA5</i>	0,080 (0,065) <sup>4</sup>	0,131* (0,086)*	0,200** (0,146)*
	<i>INRA35</i>	0,071 (0,043)	0,060 (0,046)	0,127* (0,087)
	<i>INRA37</i>	0,087* (0,066)	0,055* (0,054)*	0,137* (0,116)**
	<i>INRA23</i>	-0,014 (-0,017)	0,034 (0,020)	0,021 (0,003)
	<i>ETH3</i>	0,140* (0,131)*	0,025 (0,020)*	0,161* (0,149)*
	<i>ILSTS5</i>	0,123* (0,060)	0,066 (0,050)	0,181* (0,107)*
	<i>HEL1</i>	0,132* (0,154)*	0,041* (0,071)*	0,168* (0,214)**
	<i>HEL5</i>	0,274** (0,291)**	0,011 (0,024)	0,282** (0,308)**
	<i>ETH225</i>	0,149* (0,120)*	0,108* (0,096)*	0,240** (0,204)**
	<i>ETH152</i>	0,007 (-0,060)	0,010* (0,061)*	0,017 (0,005)
	<i>ILSTS1</i>	0,251* (0,138)*	0,062 (0,039)	0,297** (0,172)*
	<i>BM2113</i>	-0,012 (-0,034)	0,020 (0,017)	0,009* (-0,016)
	<i>HEL9</i>	0,118*	0,007	0,124**
	Mean	0,105** (0,080)**	0,047** (0,048)**	0,146** (0,124)**

<sup>1</sup> Within-population inbreeding estimate; <sup>2</sup> Measure of population differentiation; <sup>3</sup> Total inbreeding estimate. Statistical significance from permutation tests: \*\* P < 0.001, \* P < 0.05.

<sup>4</sup> Values in parenthesis are F statistics for Mozambican and Mashona breeds combined as one population (calculated for 12 microsatellites).

**Table 3** Pairwise estimates of genetic differentiation (measured as  $F_{ST}$ ) between Mozambican and Mashona cattle breeds using 12 microsatellites (bellow the diagonal) and 5 allozymes (above the diagonal)

	Angone	Landim	Bovino de Tete	Mashona
Angone	---	0.033*	0.047*	---
Landim	0.048*	---	0.034*	---
Bovino de Tete	0.038*	0.033*	---	---
Mashona	0.054 <sup>NS</sup>	0.043*	0.075 <sup>NS</sup>	---

\*  $P < 0.01$ , NS – not significant**Figure 1** Clustering assignment of Mozambican and Mashona cattle breeds obtained by STRUCTURE analysis at  $K = 2$ . Each individual is represented by a thin vertical line that is divided into segments whose size and colour correspond to the relative proportion of the animal genome corresponding to a particular cluster. Breeds are separated by thin black lines.

Estimated membership probabilities of the Angone and Bovino de Tete breeds to cluster 1 were 0.674 and 0.679, respectively. On the other hand, the Landim and Mashona breeds showed estimated membership probabilities to cluster 2 of 0.603 and 0.691, respectively. With the aid of prior population information, the majority of individuals were assigned to their rightful populations with posterior probabilities from 0.960 to 0.978 (Table 4).

**Table 4** Estimated membership probabilities of three Mozambican cattle breeds and the Mashona breed from Zimbabwe to different genetic clusters and assignment of individuals to their rightful breeds

Population	Main genetic clusters ( $K = 2$ ) <sup>a</sup>		Probability of assignment of individuals to <sup>b</sup>	
	Cluster 1	Cluster 2	Own breed	Other breeds
Angone	0.674	0.326	0.969	0.031
Landim	0.397	0.603	0.969	0.031
Bovino de Tete	0.679	0.321	0.978	0.022
Mashona	0.309	0.691	0.960	0.040

<sup>a</sup> Used admixture model.<sup>b</sup> Used model supporting prior population information.

Within Mozambican breeds the estimated contribution of the Indian Zebu was highest in the Angone cattle breed according to the estimator mR and the mean percentage of Zebu diagnostic alleles per locus (Table 5). On the other hand, the lowest contribution of Indian Zebu genes was exhibited in the Landim breed. In accordance with the data obtained with the estimator mR, the Mashona breed showed a lower contribution of Indian Zebu genes than the Mozambican breeds. The analysis of the microsatellite *INRA124* located on the Y-chromosome revealed that the indicine specific allele was fixed in the Angone, it was found in 62% of the Bovino de Tete cattle, and was absent in the Landim and Mashona breeds (Table 5).

**Table 5** Estimated percentage of Indian zebu contribution\* to the Mozambican cattle breeds and the Mashona breed of Zimbabwe

Population	Autosomal Microsatellites	Autosomal Microsatellites	Chromosome Y
	Estimated admixed coefficient mR <sup>a</sup> (%)	Mean of diagnostic zebu alleles/locus (%)	Zebu specific allele (%)
Angone	66	33.8	100
Landim	49	26.3	0
Bovino de Tete	59	31.0	62
Mashona	46	21.5	0

\* Zebu contribution to the admixed Mozambique population computed as an average after 1000 random bootstrap samples.

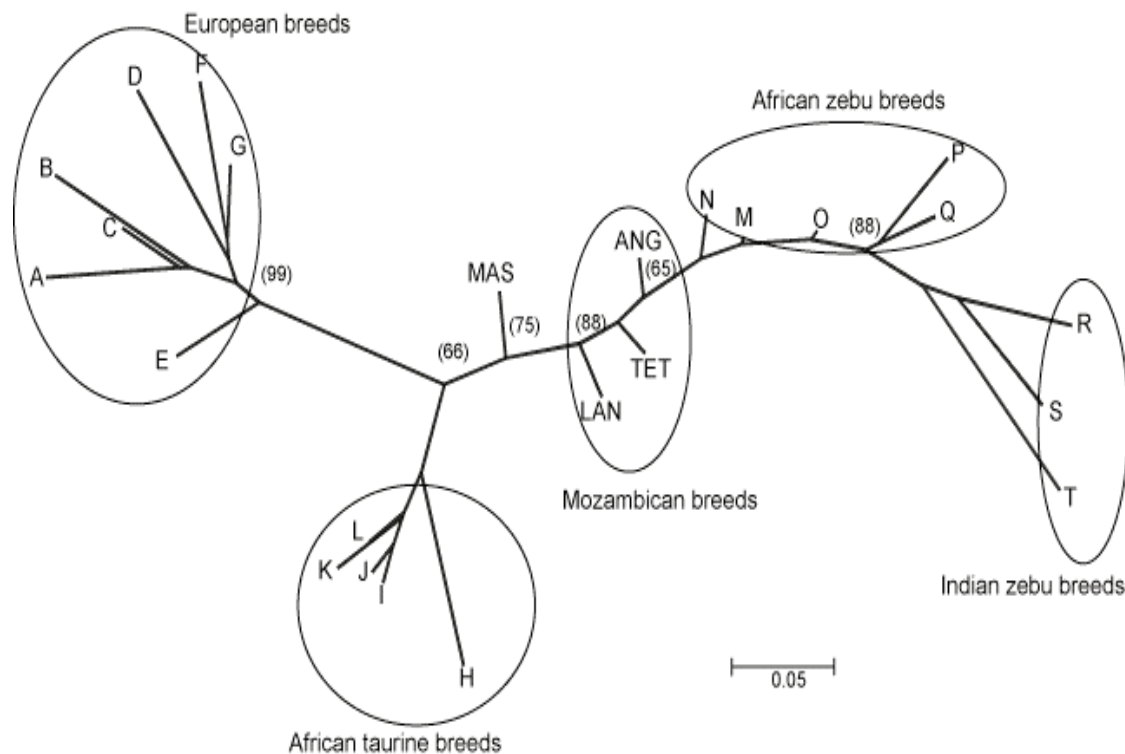
Both genetic distance matrixes ( $D_A$ ) obtained from the five allozymes and 12 microsatellite loci showed a low differentiation between breeds (Table 6). The smallest genetic distance was found between the Landim and Mashona breeds and the largest value between the Landim and Angone breeds.

**Table 6**  $D_A$  genetic distances between Mozambican and Mashona cattle breeds estimated from 12 microsatellites (bellow the diagonal) and 5 allozymes (above the diagonal)

	Angone	Landim	Bovino de Tete	Mashona
Angone	---	0.028	0.028	---
Landim	0.112	---	0.029	---
Bovino de Tete	0.070	0.079	---	---
Mashona	0.085	0.065	0.082	---

An unrooted NJ tree (Saitou & Nei, 1987) was constructed from  $D_A$  distances based on seven microsatellites using a total of 24 populations from Europe, Asia and Africa (Figure 2). The phenogram obtained revealed strong support nodes separating the breeds from the three continents. The Mozambican breeds occupied an intermediate position between Indicine and African taurine populations. The Mashona breed was located nearer to the African taurine breeds, being closer to this group than to the Landim breed. On the other hand, the Angone breed was placed nearest to the Zebu cattle. This pattern was also evident in

the PCA graph (Figure 3). The first and the second PCs explained 23.1% and 14.6%, respectively, of the variation in the data. PC1 separated the European from the Indian breeds and the African breeds occupied intermediate positions. In the PC2 the largest differentiation was observed between the African and the other breeds.



**Figure 2** Phylogenetic relationships (unrooted NJ tree) among Indian, African and European cattle populations constructed from  $D_A$  genetic distances using seven microsatellite loci.

Bootstrap values superior to 50% are shown besides each node. Data from European, African and Indian cattle were assembled from MacHugh *et al.* (1997).

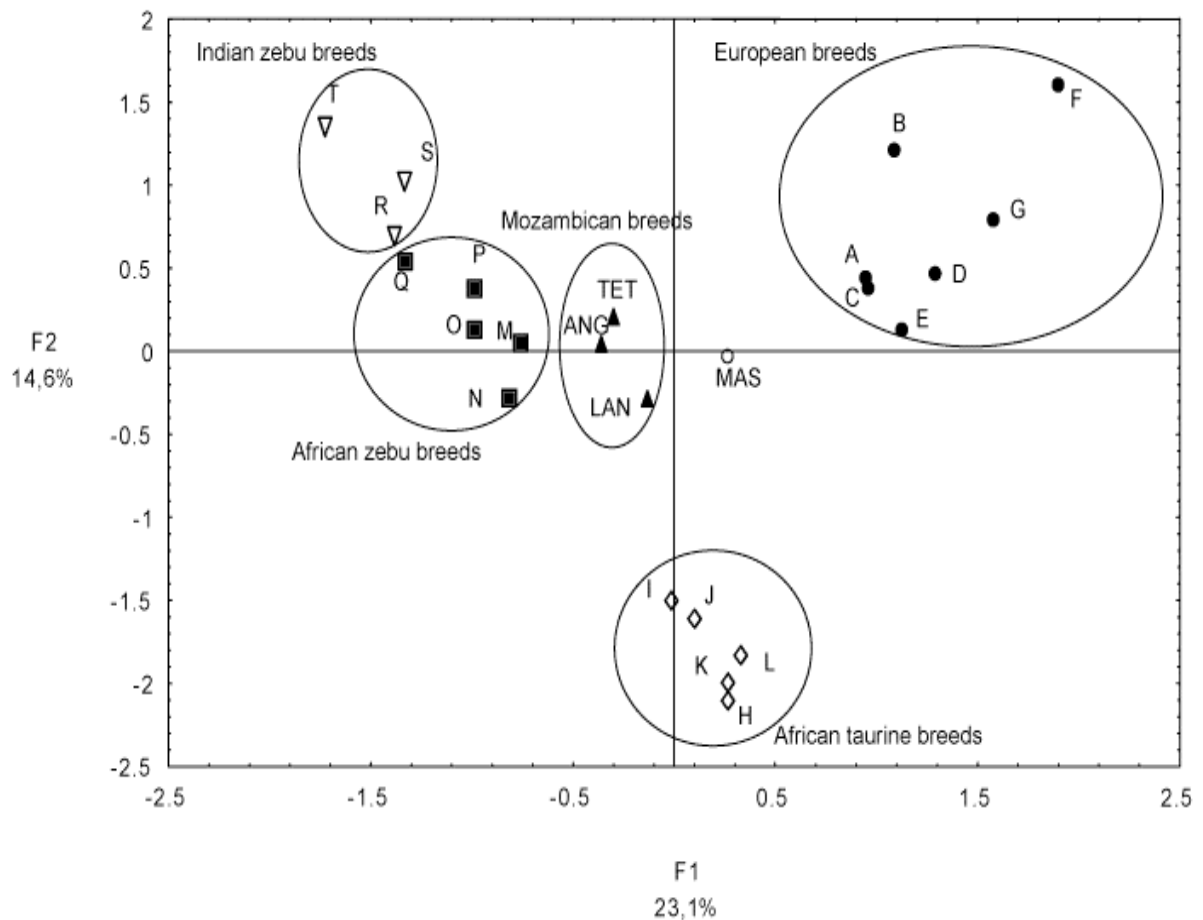
Key: ANG - Angone, TET - Bovino de Tete, LAN - Landim, MAS - Mashona, A - Angus, B - Hereford, C - Jersey, D - Kerry, E - Charolais, F - Friesian, G - Simmental, H - N'Dama (Gambia), I - N'Dama (Guinea), J - N'Dama (Guinea Bissau), K - N'Dama (Nigeria), L - N'Dama (Senegal), M - Gobra, N - Maure, O - White Fulani, P - Butana, Q - Kenana, R - Hariana, S - Sahiwal, T - Tharparker.

## Discussion

Patterns of genetic variants assayed from paternally and biparentally inherited genetic systems revealed that extensive hybridization between the two existent bovine subspecies is part of the ancestry of Mozambican cattle breeds. This is evident from the distribution of group specific alleles of Indian Zebu and African Taurine origin in these breeds.

The average number of alleles and levels of gene diversity were similar for all breeds, suggesting that there are no appreciable differences in the level of genetic variability among Mozambican breeds. The amount of genetic diversity in these breeds was comparable to or higher than those reported for other breeds (MacHugh *et al.*, 1997; Loftus *et al.*, 1999; Cañon *et al.*, 2001; Hanotte *et al.*, 2002; Freeman *et al.*, 2004; Ibeagha-Awemu & Erhardt, 2005; Freeman *et al.*, 2006; Ulloa-Arvizu *et al.*, 2008). These high diversity indices can be explained by the presence of genes from two genetically differentiated groups namely Taurine and Zebu. However, it could also be due to extensive gene flow among them and with imported breeds from neighbouring countries, especially after the substantial loss of effective population size during the years of civil war.





**Figure 3** Bi-dimensional plot of the PCA scores (Factor 1, 2) derived from the analysis of allele frequencies of seven microsatellite loci among 24 breeds:

Key: ANG - Angone, TET – Bovino de Tete, LAN - Landim, MAS - Mashona, A – A. Angus, B - Hereford, C - Jersey, D - Kerry, E - Charolais, F - Friesian, G - Simmental, H – N’Dama (Gambia), I - N’Dama (Guinea), J – N’Dama (Guinea Bissau), K - N’Dama (Nigeria), L - N’Dama (Senegal), M - Gobra, N – Maure, O – White Fulani, P – Butana, Q – Kenana, R – Hariana, S – Sahiwal, T – Tharparker. (Source: MacHugh *et al.*, 1997).

The discrepancy obtained between mitochondrial DNA and nuclear markers in previous works suggests a Zebu male-biased introgression in Africa (Bradley *et al.*, 1996; MacHugh *et al.*, 1997; Hanotte *et al.*, 2000). The results obtained in this study of Mozambican breeds, with both autosomal markers and the Y-chromosome microsatellite, are congruent with this scenario. The highest percentage of Zebu introgression was observed in the Angone breed, which is in agreement with morphological data and with the observed higher level of diversity. Due to its location, the Angone was more exposed to crossbreeding with the Zebu than the other breeds analysed. On the other hand, the Landim breed apparently includes few zebu alleles, while Bovino de Tete exhibit intermediate values. Moreover, the results obtained with the Mashona breed are in conformity with the conventional idea that the Sanga breeds from southern Africa were exposed to a lower Zebu introgression than the Sanga cattle from northeast Africa, which could have allowed the maintenance of the typical genetic structure of the Sanga cattle breeds (Hanotte *et al.*, 2002; Freeman *et al.*, 2006).

The hybrid nature of these breeds is also expressed by the absence of real distinct clusters corresponding to predefined breeds as shown by the model-based clustering method of Pritchard *et al.* (2000). The higher values of estimated membership probabilities of Angone and Bovino de Tete breeds to

cluster 1 and Landim and Mashona breeds to cluster 2, highlight that these clusters most likely correspond to the two bovine lineages, Zebu (cluster 1) and Taurine (cluster 2).

The  $F_{ST}$  and the  $D_A$  genetic distance values suggested a low genetic differentiation among the three Mozambican breeds. This result seems to be a consequence of management practices that promote some gene flow and breed substructure. The absence of herd books and the presence of indiscriminate crossbreeding encouraged a high level of gene flow between the breeds, thus promoting a low level of differentiation. However, a significant differentiation ( $P < 0.01$ ) between the Mozambican breeds is still perceptible. Moreover, the model-based clustering analysis of the breeds with prior population information (Pritchard *et al.*, 2000) assigned individuals to their rightful populations, indicating that the Mozambican breeds still conserve some of their genetic identity.

The NJ tree representing breeds from three continents placed the Bovino de Tete breed in an intermediate position between the Angone and Landim, while the Mashona breed was placed nearest to African taurine breeds than the other Sanga cattle analysed. However, using phylogenetic methods it is very difficult to separate the effect of admixture or introgression from that of common ancestry (Slatkin & Maddison, 1990). Multivariate analysis of microsatellite allele frequencies has been a powerful tool to reveal the underlying evolutionary history and admixture among distantly related Taurine and Zebu populations from Europe, Asia and Africa (MacHugh *et al.*, 1997; Freeman *et al.*, 2006). Figure 3 illustrates a PCA where factors 1 and 2 are presented. This analysis clearly reveals a cline of Zebu introgression. The Mozambican cattle, the Mashona breed and the other African Sanga breeds form a gradient between the Zebu and the African Taurine breeds. The Angone breed was undoubtedly more influenced by Zebu genes than other breeds analysed in this study, but this breed seems to be less influenced than other breeds from Northeast Africa. On the other hand, Landim and Mashona cattle had a lower influence from Zebu genes than other Sanga breeds that are situated in the north of the continent, which is in agreement with a widespread movement of Zebu cattle from east to south (MacHugh *et al.*, 1997; Hanotte *et al.*, 2002; Freeman *et al.*, 2006).

Relative to the origins of the Bovino de Tete, the data reported here support the hypothesis of a crossbreeding between Zebu and Sanga breeds, as suggested by Regé & Tawah (1999). The Bovino de Tete breed had a higher percentage of Zebu genes than the Sanga breeds analysed, which points towards a Angone derivation. However, the results do not show whether Bovino de Tete was derived from crossbreeding between Angone and Landim or between Angone and Mashona. The later hypothesis seems more reasonable given the geographic proximity between the parental breeds.

This work is the first extensive study about the genetic structure of Mozambican cattle breeds. Although the genetic diversity of these breeds seems to have been affected by gene flow and, therefore, little differentiation exists among them, it appears that they retain some of their genetic identity. These breeds are an important economic resource for Mozambican people and their high variability makes them suitable candidates for conservation and improvement. Controlling gene flow between breeds, by adopting effective breeding and management practices, will facilitate the conservation of historical breeds and preserve special characteristics of each breed.

### Acknowledgements

We thank breeders' associations for their help in collecting samples, and D. MacHugh for providing reference samples from African cattle. We thank J. Harris and F. Muchadeyi for their comments on the draft paper. This work was supported by ICCTI-UEM project funding. I. Bessa was supported by a PhD grant from FCT (PRAXISXII/BD/15578/98).

### References

- Alberro, M., 1983. The indigenous cattle of Mozambique. *World Anim. Rev.* 48, 12-17.
- Bassam, B.J., Caetano-Anolles, G. & Gresshoff, P.M., 1991. Fast and sensitive silver staining of DNA in polyacrilamide gels. *Anal. Biochem.* 196, 80-83.
- Bertorelle, G. & Excoffier, L., 1998. Inferring admixture proportions from molecular data. *Mol. Biol. Evol.* 15, 1298-1311.
- Bradley, D.G., MacHugh, D.E., Cunningham, P. & Loftus, R.T., 1996. Mitochondrial diversity and the origins of African and European cattle. *Proc. Natl. Acad. Sci. USA.* 93, 5131-5135.

- Cañon, J., Alexandrino, P., Bessa, I., Carleos, C., Carretero, Y., Dunner, S., Ferrand, N., Garcia, D., Jordana, J., Laloe, D., Pereira, A., Sanchez, A. & Moazami-Goudarzi, K., 2001. Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genet. Sel. Evol.* 33, 311-332.
- Carvalho, J.G.V., Blake, R.W., Pollak, E.J. & Van Soest, P.J., 1995. Comparison of Landim and Africander cattle in southern Mozambique: Female fertility, reproduction and beef offtake. *J. Anim. Sci.* 73, 3527-3533.
- Cymbron, T., Freeman, A.R., Loftus, R.T., Malheiro, M.I. & Bradley, D.G., 2005. Microsatellite diversity suggests different origins for Mediterranean and Northern European cattle populations. *Proc. Biol. Sci.* 272, 1837-1843.
- Dionisio, A.C. & Syrstad, O., 1990. Productivity of Nguni and Africander cattle in Mozambique. *Livest. Prod. Sci.* 24, 29-36.
- Egito, A.A., Paiva, S.R., Albuquerque, M.S., Mariante, A.S., Almeida, L.D., Castro, S.R. & Grattapaglia, D., 2007. Microsatellite based genetic diversity and relationships among ten Creole and commercial cattle breeds raised in Brazil. *BMC Genetics* 8, 83.
- Epstein, H., 1971. *The Origin of the Domestic Animals of Africa*, Vol. I, Africana Publishing Corporation, New York, USA.
- Falush, D., Stephens, M. & Pritchard, J.K., 2003. Inference of population structure: Extensions to linked loci and correlated allele frequencies. *Genetics* 164, 1564-1587.
- Felsenstein, J., 2004. PHYLIP: Phylogeny Inference Package, version 3.6b, <http://evolution.gs.washington.edu/phylip.html> [consulted: 28 May 2007].
- Freeman, A.R., Meghen, M., MacHugh, D.E., Loftus, R.T., Achukwi, M.D., Bado, A., Sauveroché, B. & Bradley, D.G., 2004. Admixture and diversity in West African cattle populations. *Mol. Ecol.* 13, 3477-3487.
- Freeman, A.R., Bradley, D.G., Nagda, S., Gibson, J.P. & Hanotte, O., 2006. Combination of multiple microsatellite data sets to investigate genetic diversity and admixture of domestic cattle. *Anim. Genet.* 37, 1-9.
- Goudet, J., 2002. FSTAT: A program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3.2). URL: <http://www2.unilch/izea/software/fstat.html>
- Hanks, J. & Pereira, C.L., 1998. *The Development of a Restocking Decision Support Aid for Mozambique*, PAN Livestock Services Limited, pp.31.
- Hanotte, O., Bradley, D.G., Ochieng, J.W., Verjee, Y., Hill, E.W. & Regé, J.E., 2002. African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336-339.
- Hanotte, O., Tawah, C.L., Bradley, D., Okomo, M., Verjee, Y., Ochieng, J. & Regé, J.E.O., 2000. Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-Saharan African cattle breeds. *Mol. Ecol.* 9, 387-396.
- Ibeagha-Awemu, E.M. & Erhardt, G., 2005. Genetic structure and differentiation of 12 African *Bos indicus* and *Bos taurus* cattle breeds, inferred from protein and microsatellite polymorphisms. *J. Anim. Breed. Genet.* 122, 12-20.
- Jordana, J., Alexandrino, P., Beja-Pereira, A., Bessa, I., Cañon, J., Carretero, Y., Dunner, S., Laloë, D., Moazami-Goudarzi, K., Sanchez, A. & Ferrand, N., 2003. Genetic structure of eighteen local south European beef cattle breeds by comparative *F*-statistics analysis. *J. Anim. Breed. Genet.* 120, 73-87.
- Leal, C., 1998. Genetic characterization of Barrosã cattle breed. MSc. thesis, Faculty of Sciences, University of Porto, Portugal. (in Portuguese).
- Loftus, R.T., Ertugrul, O., Harba, A.H., El-Barody, M.A.A., MacHugh, D.E., Park, S.D. & Bradley, D.G., 1999. A microsatellite survey of cattle from a centre of origin: the Near East. *Mol. Ecol.* 8, 2015-2022.
- MacHugh, D.E., Shriver, M.D., Loftus, R.T., Cunningham, P. & Bradley, D.G., 1997. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and Zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics* 146, 1071-1086.
- Marshall, F., 2000. The origins of domesticated cattle in Eastern Africa. In: *The Origins and Development of African Livestock: Archaeology, Genetics, Linguistics and Ethnography*. Eds Blench, R.M. & MacDonald, K.C., University College London Press, London. pp. 191-221.
- Morgado, F.P., 2000. A Pecuaría no Norte de Moçambique, VEGA. pp. 357. (in Portuguese).
- Morgado, F.P., 2004. A Pecuaría no Centro de Moçambique, VEGA. pp. 470. (in Portuguese).

- Nei, M., 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583-590.
- Nei, M., Tajima, F. & Tatenno, Y., 1983. Accuracy of estimated phylogenetic trees from molecular data II. Gene frequency data. *J. Mol. Evol.* 19, 153-170.
- Ota, T., 1993. *DISPAN: Genetic Distances and Phylogenetic Analysis*. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, Pennsylvania.
- Pritchard, J.K., Stephens, M. & Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945-959.
- Raymond, M. & Rousset, F., 1995. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism, *J. Hered.* 86, 248-249.
- Regé, J.E.O. & Tawah, C.L., 1999. The state of African cattle genetic resources. II. Geographical distribution, characteristics and uses of present-day breeds and strains. *Anim. Genet. Resour. Inform.* 26, 1-25.
- Rice, W.R., 1989. Analysing tables of statistical tests. *Evolution* 43, 223-225.
- Rocha, A., Starkey, P. & Dionisio, A.C., 1991. Cattle production and utilisation in smallholder farming systems in southern Mozambique. *Agric. Syst.* 37, 55-75.
- Saitou, N. & Nei, M., 1987. The Neighbor-Joining Method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406-425.
- Sambrook, J., Fritsch, E.F. & Maniatis, T., 1995. *Molecular Cloning: a Laboratory Manual*, 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press, New York.
- Scholtz, M.M, Spickett, A.M., Lombard, P.E. & Enslin, C.B., 1991. The effect of tick infestation on the productivity of cows of three breeds. *Onderstepoort J. Vet. Res.* 58, 71-74.
- Slatkin, M. & Maddison, W.P., 1990. Detecting isolation by distance using phylogenies of genes. *Genetics* 126, 249-260.
- StatSoft, 1995. Inc. STATISTICA for Windows [Computer program manual]. Tulsa, O.K., USA. <http://www.statsoft.com>
- Ulloa-Arvizu, R., Gayosso-Vázquez, A., Ramos-Kuri, M., Estrada, F.J., Montaña, M. & Alonso, R.A., 2008. Genetic analysis of Mexican Criollo cattle populations. *J. Anim. Breed. Genet.* 125, 351-359.
- Vaiman, D., Mercier, D., Moazami-Goudarzi, K., Eggen, A., Ciampolini, R., Lépingle, A., Velmata, R., Kaukinen, J., Varvio, S-L., Martin, P., Levéziel, H. & Guérin, G., 1994. A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mamm. Genome* 5, 288-297.
- Van Rensburg, S.W.J., 1953. Bovine sterility caused by infectious diseases in South Africa. *Br. Vet. J.* 109, 226-233.
- Weir, B.S. & Cockerham, C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.
- Wright, S., 1951. The genetical structure of populations. *Ann. Eugenics* 15, 323-54.