



HAL
open science

Genetic and morphological characterisation of the Ankole Longhorn cattle in the African Great Lakes region

Deo B. Ndumu, Roswitha Baumung, Olivier Hanotte, Maria Wurzinger, Mwai A. Okeyo, Han Jianlin, Harrison Kibogo, Johann Sölkner

► **To cite this version:**

Deo B. Ndumu, Roswitha Baumung, Olivier Hanotte, Maria Wurzinger, Mwai A. Okeyo, et al.. Genetic and morphological characterisation of the Ankole Longhorn cattle in the African Great Lakes region. *Genetics Selection Evolution*, 2008, 40 (5), pp.467-490. hal-00894632

HAL Id: hal-00894632

<https://hal.science/hal-00894632>

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Genetic and morphological characterisation of the Ankole Longhorn cattle in the African Great Lakes region

Deo B. NDUMU^{1,2,3}, Roswitha BAUMUNG^{1*}, Olivier HANOTTE³,
Maria WURZINGER¹, Mwai A. OKEYO³, Han JIANLIN^{3,4},
Harrison KIBOGO³, Johann SÖLKNER¹

¹Department of Sustainable Agricultural Systems, BOKU-University of Natural Resources
and Applied Life Sciences, Vienna, Austria

²Ministry of Agriculture, Animal Industry and Fisheries, Directorate of Animal Resources,
P.O. Box 513, Entebbe, Uganda

³International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi 00100, Kenya

⁴CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources,
Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS),
Beijing 100094, China

(Received 17 July 2007; accepted 30 April 2008)

Abstract – The study investigated the population structure, diversity and differentiation of almost all of the ecotypes representing the African Ankole Longhorn cattle breed on the basis of morphometric (shape and size), genotypic and spatial distance data. Twenty-one morphometric measurements were used to describe the morphology of 439 individuals from 11 sub-populations located in five countries around the Great Lakes region of central and eastern Africa. Additionally, 472 individuals were genotyped using 15 DNA microsatellites. Femoral length, horn length, horn circumference, rump height, body length and fore-limb circumference showed the largest differences between regions. An overall F_{ST} index indicated that 2.7% of the total genetic variation was present among sub-populations. The least differentiation was observed between the two sub-populations of Mbarara south and Luwero in Uganda, while the highest level of differentiation was observed between the Mugamba in Burundi and Malagarasi in Tanzania. An estimated membership of four for the inferred clusters from a model-based Bayesian approach was obtained. Both analyses on distance-based and model-based methods consistently isolated the Mugamba sub-population in Burundi from the others.

Ankole Longhorn cattle / microsatellite / geometric morphometric / genetic distance / spatial distance

*Corresponding author: roswitha.baumung@boku.ac.at

1. INTRODUCTION

The progenitors of the present-day African Ankole Longhorn cattle can be traced back by archaeological findings to the Nile delta, to about 7000 BC, from where along with human migration, groups of humpless Longhorns are thought to have left the Lower Nile for Abyssinia towards the end of the third pre-Christian millennium [8]. They are also thought to have interbred with the Lateral-horned Zebus to produce the various Sanga cattle, which later migrated south of the Sahara towards the Great Lakes and beyond [15,21]. Previous studies by Freeman *et al.* [9] and by Hanotte *et al.* [14] have indicated minimal recent male-mediated indicine gene introgression into the Ankole cattle populations, either through the Zenga or *Bos indicus* populations. Many pre-colonial kingdoms in the area are also associated with the Longhorn cattle. Various tribes, most of them from these former kingdoms, have since kept the Longhorns, albeit under different production systems and using different indigenous selection criteria. The different Longhorn cattle races mainly go by the same tribal names as their owners, and they include the Bahima cattle found in south-western half of the Cattle Corridor of Uganda, the Kigezi cattle from the south-western Ugandan highland, the Ntuuku cattle from the Lake Albert region of the Albertine Rift Valley in Uganda, the Watusi and Inkuku cattle from Rwanda, the Inyaruguru and Inyambu cattle from Burundi, the Enyambu cattle kept by the Banyambu people of north-western Tanzania, the Malagarasi Ankole ecotype kept by the pastoralists of Tutsi descent in the Malagarasi river valley of western Tanzania and the Bashi cattle kept by the Bashi people of north Kivu in DR Congo. The production systems under which the cattle were kept are described elsewhere [38], while the ecological descriptions have been reported by Grimaud *et al.* and Okello *et al.* [12,25] and are therefore not covered here. Traditional production systems require multipurpose animals, capable of providing a wide range of products and services. The Ankole Longhorn cattle of the Great Lakes region provide employment and are a source of income; the cattle are a form of insurance and accumulation of wealth; they have an important social and cultural role such as dowry payment as well as other intangible values [38].

While the Ankole Longhorn cattle are multipurpose, they are also adapted to the environmental rigours of the region. They are tolerant against ticks [24] and possess a demonstrable level of resistance to theileriosis [27]. The cattle, like other indigenous breeds, can withstand severe droughts, survive on low-quality feeds and tolerate helminths to some degree [3,11].

Like other breeds in the region, the genetic diversity of Ankole Longhorn cattle is under threat from indiscriminate crossbreeding, breed substitution,

accelerated admixture from other local breeds, epizootics, famine and civil strife, as well as from a lack of systematic breed development programmes [38].

The aim of this study was to evaluate the morphological and genotypic differences between and among the Ankole Longhorn cattle populations, to investigate the relationship of such differences to their spatial geographic distances, and to broadly relate them to the breed's genetic variation and the breeding goals (selection preferences) of their owners, all of which have significant implications for utilisation strategies and the breed's sustainable conservation.

2. MATERIALS AND METHODS

2.1. Sampling

Morphometric body measurements, blood samples for microsatellite DNA analysis and satellite geographic positions were taken from 11 sub-populations of the Ankole Longhorn cattle present in five countries in the African Great Lakes region. The countries include Uganda, Rwanda, Burundi, Tanzania and the DR Congo (Fig. 1). Morphometric measurements were taken from 439 females. Blood samples were collected from 472 animals, of which 33 were male and 439 were female. An average of four animals was sampled per herd within each sub-population, with due care taken to avoid sampling of closely related individuals. No reference animals were genotyped in this study.

An Etrex[®] global positioning system (GPS) device employing a satellite navigation system was used for the definition of the particular geographic location of the different herds that were sampled. Data were downloaded from the device and read using the ArcView and MapSource software. Data for the 66 geographic positions determined for each sampled herd included latitudes and longitudes of the locations.

2.1.1. Morphometric measurements

Twenty-one morphometric body measurements were taken of each animal at predefined anatomical points on the horns, head, dewlap, forequarter, barrel, hindquarters, including horn tip interval, horn base circumference, horn length, horn lower interval, head length, head width, muzzle circumference, dewlap distances, heart chest girth, height at withers, fore-arm length, fore-limb circumference (smallest circumference around the metacarpus), fore-quarter length, body length, hip width, pins width, rump height, rump length, lumbosacral angle, rump angle and femoral length. The measurements included distances (in centimetres), circumferences (in centimetres), angles (in degrees) as well as

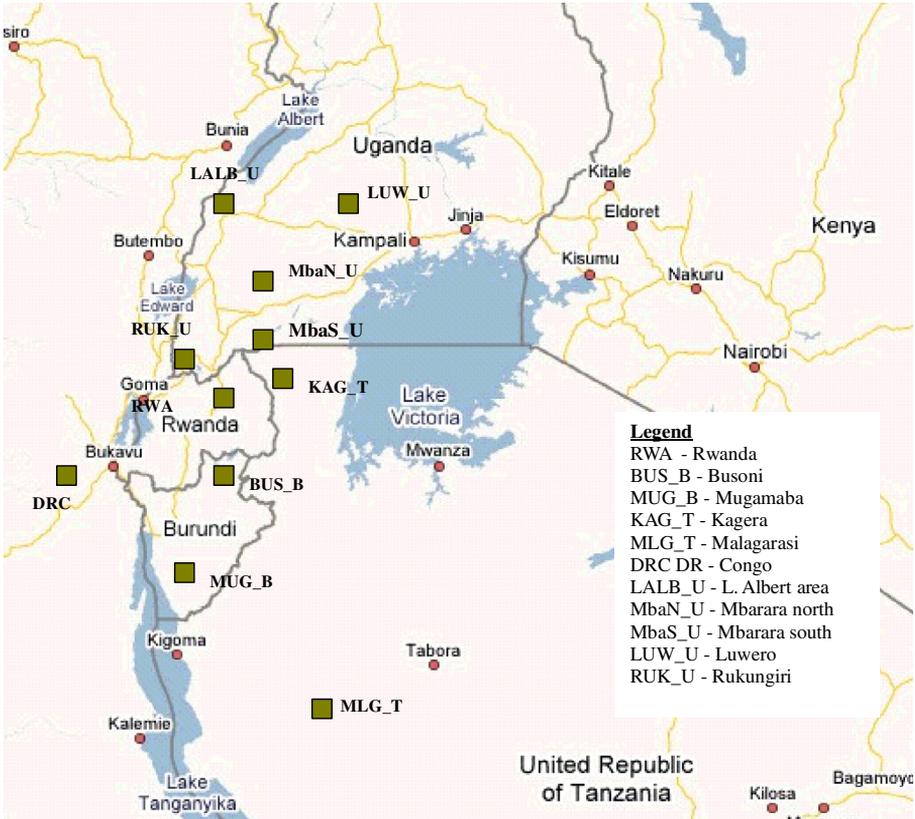


Figure 1. Map indicating the sampling locations in the African Great Lakes region.

the description of coat colour and pattern and colour of horns. The instruments used were a measuring stick (hippometer), chest band, measuring tape, an outside calliper and a digital spirit level (inclinometer).

2.1.2. Genotyping

Blood samples were transferred to the molecular laboratory of ILRI for genotyping. DNA was extracted following a modified phenol-chloroform extraction and ethanol precipitation [33]. Fifteen microsatellite DNA markers (*ILSTS006*, *INRA032*, *MGTG4B*, *TGLA122*, *AGLA293*, *ETH225*, *HEL001*, *ILSTS023*, *BM2113*, *ETH152*, *ILSTS050*, *INRA035*, *CSSM66*, *ILSTS005*, *INRA005*) drawn from the FAO/ISAG recommended list [17] were employed in this study (Tab. I in Appendix II). Fragment amplification was accomplished by polymerase chain reactions (PCRs) using the GeneAmp® PCR System 9700 thermocycler

Table I. Genetic diversity in the 11 Ankole Longhorn sub-populations based on 15 microsatellite markers; gene diversity (unbiased H_z), H_o , allelic richness (based on minimum sample size of 26 diploid individuals – 52 genes), MNA and Wright's F_{IS} .

Sub-population	Sample size	He	H_o	Allelic richness	MNA	F_{IS}
RWA	48	0.74 ± 0.02	0.72 ± 0.02	7.20	8.20 ± 3.67	0.037*
BUS_B	42	0.72 ± 0.03	0.67 ± 0.02	6.80	7.27 ± 2.15	0.073***
MUG_B	46	0.74 ± 0.02	0.67 ± 0.02	7.28	8.20 ± 1.52	0.098***
KAG_T	42	0.79 ± 0.02	0.72 ± 0.02	8.08	8.73 ± 3.13	0.085***
MLG_T	40	0.73 ± 0.03	0.67 ± 0.02	6.82	7.27 ± 1.87	0.080***
DRC	41	0.73 ± 0.03	0.72 ± 0.02	6.46	6.93 ± 1.79	0.018 ^{NS}
LALB_U	39	0.75 ± 0.03	0.72 ± 0.02	7.60	8.27 ± 2.12	0.036*
MbaN_U	39	0.74 ± 0.02	0.74 ± 0.02	6.46	6.80 ± 1.42	0.009 ^{NS}
MbaS_U	42	0.74 ± 0.02	0.69 ± 0.02	6.66	7.13 ± 2.10	0.065**
LUW_U	44	0.74 ± 0.02	0.68 ± 0.02	6.32	6.87 ± 1.41	0.082***
RUK_U	49	0.73 ± 0.02	0.69 ± 0.02	6.43	7.27 ± 1.87	0.044*

NS: non-significance; *significance at $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

on either the basic or touch-down programs. Genotyping was done by capillary electrophoresis on the Applied Biosystems 3730 DNA Analyzer instrument. Genotypes were analysed using the GeneMapper (version 3.7) software while employing the advanced peak detection algorithm and the third order least squares (LS) method under the Microsatellite Default. Allele sizes were conveniently scored using the BINS system.

2.2. Statistical analysis

2.2.1. Morphological description of size

The morphological description of the variation in the traits measured among the 11 sub-populations was done using the SAS[®] general linear models (GLM) procedure [34]. Models were kept relatively simple to avoid over-parameterisation. The LS means were computed for the traits measured and a test of significance between different sub-populations was done using the Tukey-Kramer multiple comparisons method. Multivariate analyses [34] were used to investigate the morphological structure and quantify differences among the sub-populations. Stepwise discriminant analysis [39] was applied to gain information about traits particularly important in the separation of sub-populations. Additionally, canonical discriminant analysis was employed to obtain the function of all traits necessary for the separation of sub-populations. Results from the latter

analysis were represented by squared distances between standardised class means according to Mahalanobis. This enabled a pairwise comparison of morphological structures between the different sub-populations. A plot derived from the multidimensional scaling (MDS) procedure [36] on the squared distance matrix was used to visually portray association between the least and/or the most differentiated sub-populations.

2.2.2. Geometric morphometric description

In the analysis of head and body shape with methods of geometric morphometrics, distance, angular and circumference measurements were converted into a set of two-dimensional Cartesian coordinates applying simple geometric functions (sine, cosine, Pythagoras' theorem, calculation of diameters of circles from their circumference).

Geometric morphometrics, developed by Rohlf and Marcus [30], Bookstein [6] and Adams *et al.* [1], provide a set of tools to deal with the shape of specimens, while multivariate statistics on measures of distance [39] tend to distinguish sub-populations different in size.

The Procrustes analysis applied here follows several steps. First, all the landmark configurations are scaled by standardising the size to a unit centroid size, the centroid size corresponding to the square root of the sum of the squared distances between the centroid (*i.e.* centre of gravity of the landmarks) and each of the configured landmarks. Then, the centroids of all the landmark configurations are superimposed and translated to the origin. The landmark configurations are rotated against a consensus configuration so that the sum of the squares of the residual distances between corresponding landmarks is a minimum. Finally, the aligned landmarks undergo a series of transformations, maintaining the characteristics of shape while reducing the number of dimensions. The resulting $2n - 4$ relative warp scores (n being the number of landmarks) are the dependent variables in conventional multivariate statistics. MDS and cluster analysis were applied to the distance matrices, and the results of MDS proved to be more instructive in graphical presentation.

Figures 2a and 3a indicate the six landmarks of the head and the eight landmarks of the body. Two angles in the front part of the body were approximated, as they were not measured, muzzle and horn circumference were assumed to form circles; while the chest depth necessary to define landmark 8 of the body was calculated according to the literature data relating it to chest circumference [35]. Figure 2b shows the mean unaligned (raw) coordinates of six landmarks of the head, while Figure 3b presents rescaled and aligned coordinates of the eight landmarks of the body.

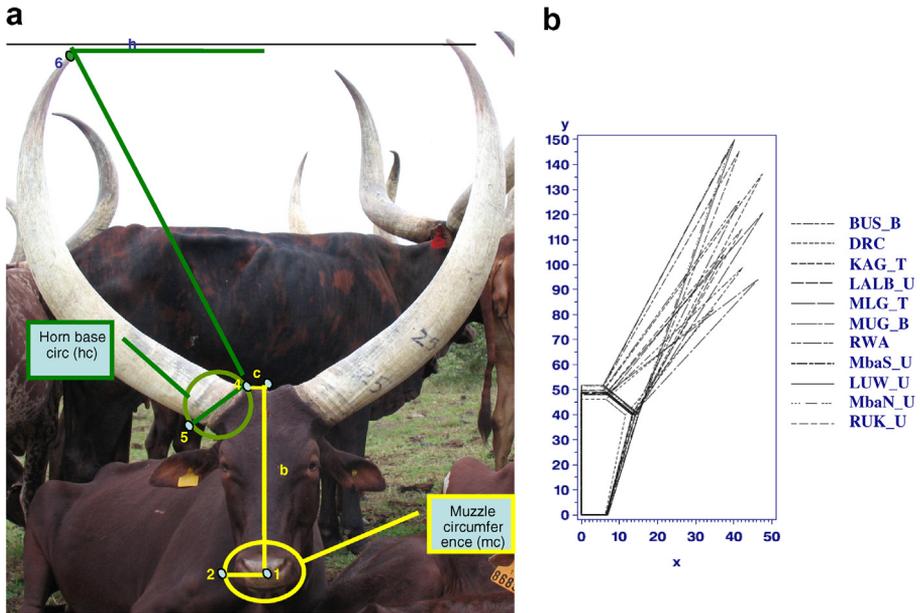


Figure 2. (a) Landmarks defining the shape of the head: landmark 1: (0, 0) is the reference point; landmark 2 $(-mc/2\pi, 0)$; landmark 3 is (0, b); landmark 4 is $(-c, b)$; landmark 5 $(-c - hc/\pi\sqrt{2}, b - hc/\pi\sqrt{2})$; landmark 6 $(-h, b + \sqrt{e^2 - (h - c)^2})$ where b, c, e and h are as in the graph, mc is muzzle circumference and hc is horn base circumference. (b) Mean unaligned (raw) head coordinates for the 11 regions.

2.2.3. Genetic characterisation

A total of 6893 successful genotypes from 15 loci and 472 individuals from 11 sub-populations were used to investigate and describe the genetic diversity of the sub-populations. Allele frequencies and number of alleles, across loci and sub-populations as well as the mean number of alleles (MNA) and allelic richness across sub-populations were estimated using the FSTAT software [10].

Observed heterozygosity (H_o) and gene diversity were also calculated across loci and sub-populations using the Excel Microsatellite Toolkit. Tests for deviation from the Hardy-Weinberg equilibrium across loci and populations as well as the estimation of the unbiased P -value using the Markov chain Monte Carlo (MCMC) algorithm according to Guo and Thompson [13] were computed with the GENEPOP program [29]. Wright’s F_{IS} index values [37] were computed to assess the closeness of each sub-population to random breeding conditions, and tests of significance at 5% indicative adjusted nominal level were done using the FSTAT program [10] with 165 000 randomisations.

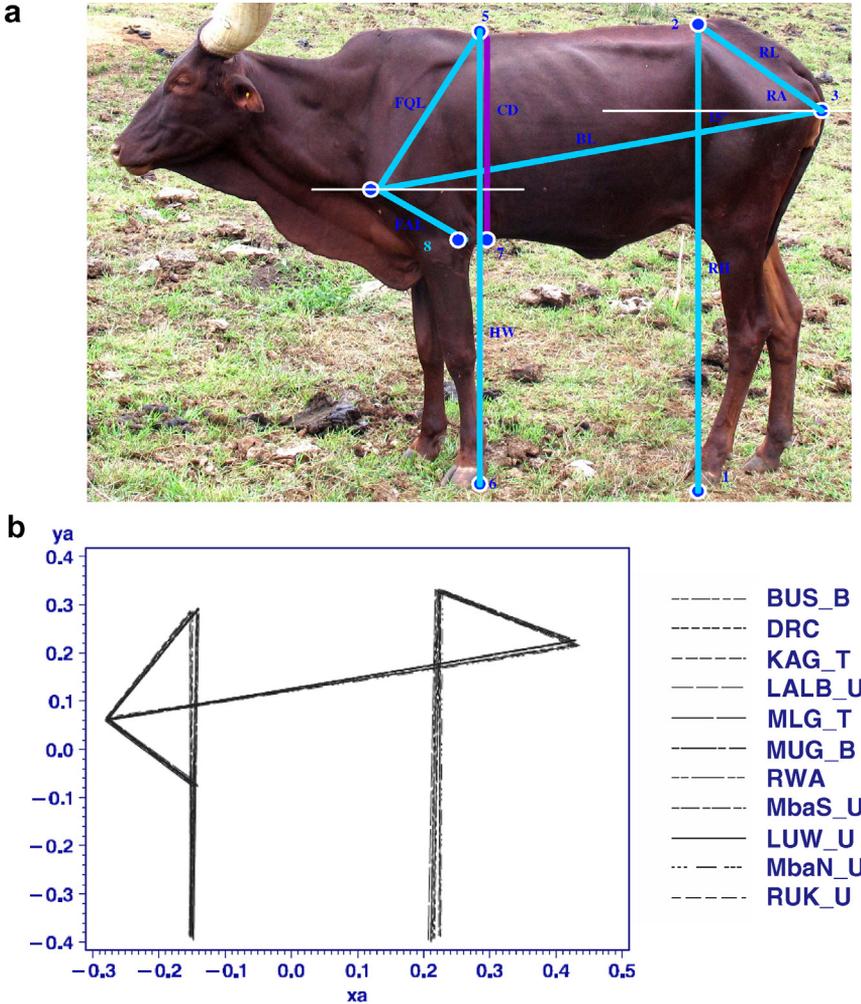


Figure 3. (a) Landmarks used to define the shape of the body: landmark 1: (0, 0) is the reference point; landmark 2 (0, RH); landmark 3 ($RL \cdot \cos(RA)$, $RH - RL \cdot \sin(RA)$); landmark 4 ($RL \cdot \cos(RA) - BL \cdot \cos(15^\circ)$, $HW - FQL \cdot \sin(60^\circ)$); landmark 5 ($RL \cdot \cos(RA) - BL \cdot \cos(15^\circ) + FQL \cdot \cos(60^\circ)$, HW); landmark 6 ($RL \cdot \cos(RA) - BL \cdot \cos(15^\circ) + FQL \cdot \cos(60^\circ)$, 0); landmark 7 ($RL \cdot \cos(RA) - BL \cdot \cos(15^\circ) + FQL \cdot \cos(60^\circ)$, $HW - CD/2.6442$); landmark 8 ($RL \cdot \cos(RA) - BL \cdot \cos(15^\circ) + FAL \cdot \cos(45^\circ)$, $HW - FQL \cdot \sin(60^\circ) - FAL \cdot \sin(45^\circ)$). RH = rump height, RL = rump length, RA = rump angle, BL = body length, HW = height at withers, CD = chest depth, FQL = fore-quarter length, FAL = fore-arm length. (b) Mean unaligned (raw) body coordinates for the 11 regions.

The molecular genetic relationship was explored by way of pairwise comparisons of Nei's DA distances [23] between sub-populations estimated using the Dispan program [26]. Furthermore, gene differentiation (F_{ST} index) among the sub-populations and pairwise F_{ST} between the sub-populations were investigated following Wright's method [37] using GENETIX [4] and FSTAT software [10]. The significance of pairwise F_{ST} estimates was tested at 5% indicative adjusted nominal level using the FSTAT program [10] with 55 000 permutations. On the basis of Nei's DA distance matrix, a dendrogram derived from the Neighbour-Joining algorithm [32] was constructed in the Dispan program [26].

To infer population structure, individual animals were probabilistically assigned to sub-populations using Structure 2.0 [28], which employs a model-based Bayesian clustering approach. For ancestry, we assumed the admixture model, while for allele frequencies, we assumed a model for correlated frequencies. By these assumptions and from a pre-assigned number of clusters (K), the program, using the MCMC algorithm, computed the estimate of the natural logarithm of the posterior probability of the clusters K in the population given the observed genotypic composition G ($\text{Ln Pr}(K/G)$). The latter is directly proportional to the estimated natural logarithm of the probability (Pr) of the observed genotype composition (G) given a pre-assigned number of clusters (K) in the structure program data set – $\text{Ln Pr}(G/K)$. To estimate the number of clusters in our data, we set K between 2 and 11 with 10 independent runs of the Gibbs sampler for each value of K , including a burn-in period of 10^6 iterations followed by 10^6 MCMC iterations. We used default settings in all runs, that is, an admixture model with correlated frequencies and the parameter of individual admixture alpha set to be the same for all clusters and with a uniform prior. The graphical display of the population structure was done using DISTRUCT [31].

2.2.4. Geographic distances

Geographic distances were described by combining coordinate data comprising latitudes and longitudes of the individual herds sampled together with the microsatellite data set. The geographic data were converted into a spatial distance matrix, whereby all individuals of the same sub-population shared the same average spatial location. The individuals of each sub-population were treated as dependent, and the regression analysis in SPAGeDi [16] took into account pairwise comparisons between groups of individuals of a sub-population rather than individuals themselves.

2.2.5. Mantel tests

The emerging evidence of the resolution capacity of the geometric morphometrics in the study of the variation of anatomical structures [6] provides an impetus to validating patterns of geographic variation in cattle populations. This can also be done in conjunction with genetic analyses. Consequently, we performed canonical discriminant analyses to arrive at sets of Mahalanobis distances, which were then included in a series of Mantel tests [20] comparing genetic, morphometric and geographic distances using the *zt* program [5].

3. RESULTS

3.1. Morphological description

The results obtained by SAS[®] GLM for coefficients of determination are shown in Table II in Appendix II. The 10 most important traits (horn length, thigh length, rump height, dewlaps, horn base size, fore-limb circumference, horn tip interval, heart chest girth, horn lower interval and muzzle circumference) separating sub-populations according to a stepwise discriminant analysis are presented in Table III in Appendix II, in their order of level of contribution to the discrimination of the sub-populations. Results of the canonical discriminant analysis are illustrated in Figure 4. The first canonical variate separates three sub-populations of Uganda, namely the Mbarara north, Mbarara south and Luwero, from the rest of the sub-populations. The second variate further separates the Rukungiri sub-population of Uganda from the remaining sub-populations. The Malagarasi sub-populations of Tanzania and the Rwandan sub-population are close to each other, and the two sub-populations in Burundi are close to the DR Congo sub-population.

The Mahalanobis squared distances between sub-populations are significant ($P < 0.05$), except for those of three pairs between the Ugandan sub-populations of Mbarara north, Mbarara south and Luwero (Tabs. IV and V in Appendix II).

A plot of the results of MDS procedure [34] performed on the partial warps scores matrix for body shape and head shape among the sub-populations is presented in Figures 5a and 5b, respectively.

3.2. Genetic characteristics

The characteristics of the 15 microsatellites used for this analysis are shown in Table I in Appendix II. A total of 207 alleles were observed in 472 individuals from the 11 sub-populations, while the average number of samples typed per locus was 459.5. The highest number of alleles observed, per locus, was 25

Table II. Pairwise comparison of $F_{ST} - \theta$ values between sub-populations.

	BUS_B	MUG_B	KAG_T	MLG_T	DRC	LALB_U	MbaN_U	MbaS_U	LUW_U	RUK_U
RWA	0.010 ^{NS}	0.047***	0.026***	0.046***	0.036***	0.024***	0.036***	0.028***	0.032***	0.038***
BUS_B		0.047***	0.038***	0.042***	0.036***	0.035***	0.051***	0.033***	0.039***	0.046***
MUG_B			0.026***	0.058***	0.051***	0.049***	0.051***	0.035***	0.040***	0.049***
KAG_T				0.027***	0.027***	0.021***	0.020***	0.015***	0.016***	0.020***
MLG_T					0.008***	0.024***	0.025***	0.014***	0.016***	0.028***
DRC						0.018***	0.014***	0.006***	0.003 ^{NS}	0.016***
LALB_U							0.021***	0.021***	0.023***	0.018***
MbaN_U								0.002 ^{NS}	0.002 ^{NS}	0.012**
MbaS_U									0.000 ^{NS}	0.008***
LUW_U										0.011***

NS: non-significance; **significance at $P < 0.01$ and *** $P < 0.001$.

Table III. Mantel tests from matrix comparisons of distance measurements.

Distance 1	Distance 2	Correlation coefficient (<i>R</i>)	Level of significance
F_{st}	Body size (Mahalanobis)	-0.15	NS
F_{st}	Body shape (Mahalanobis)	0.10	NS
F_{st}	Head shape (Mahalanobis)	-0.09	NS
F_{st}	Spatial distance	0.04	NS
Body shape (Mahalanobis)	Spatial distance	0.37	**
Body size (Mahalanobis)	Spatial distance	0.21	NS
All size traits (Mahalanobis)	Spatial distance	0.26	*
Head shape (Mahalanobis)	Spatial distance	0.22	NS
Body shape (Mahalanobis)	Body size (Mahalanobis)	0.76	***

NS: non-significance; *significance at $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.

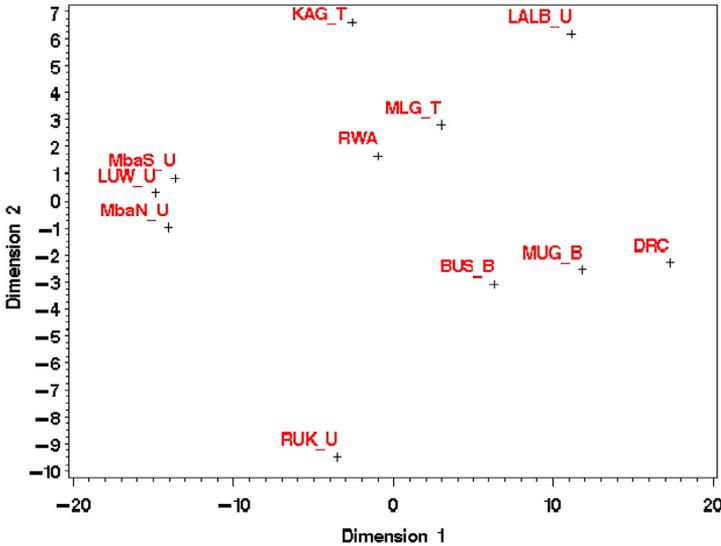


Figure 4. A plot of the results of the MDS procedure performed on the Mahalanobis squared distance matrix for body size.

at *TGLA122*, while the lowest number was 7 at *ILSTS005*. H_o ranged between 0.41 for *ILSTS023* and 0.80 for *MGTG4B*, while the expected heterozygosity (H_e) range was between 0.56 for the *ILSTS023* and 0.81 for the *INRA032*.

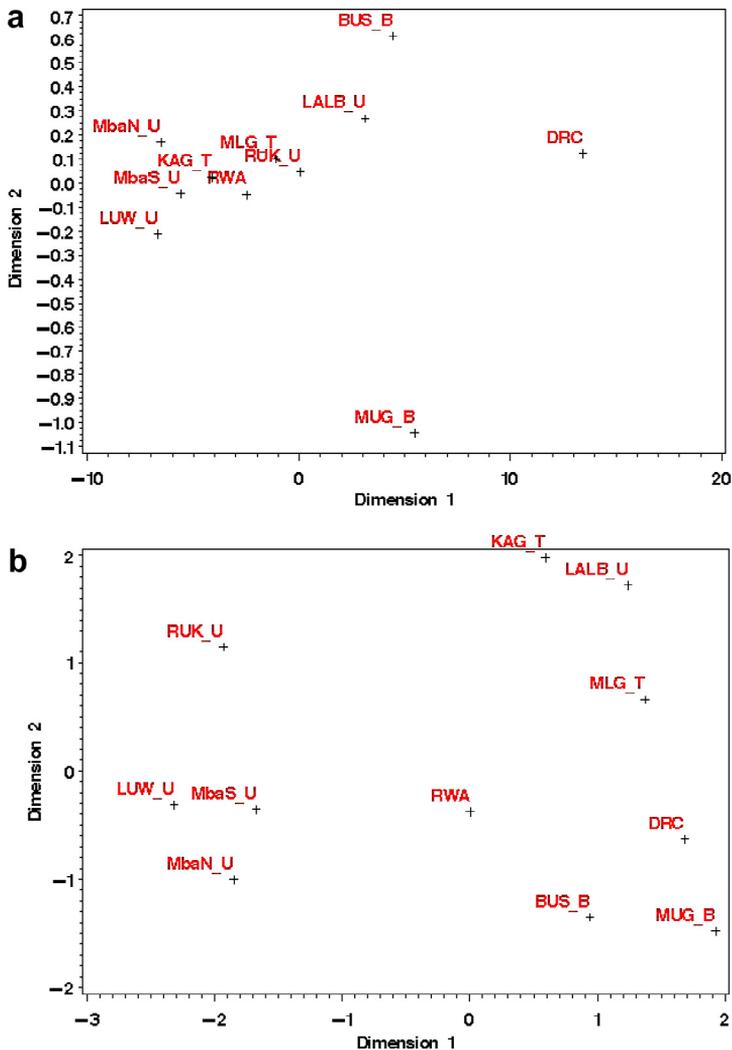


Figure 5. A plot of the results of the MDS procedure performed on the (a) head shape matrix and (b) body shape matrix.

The relative magnitude of gene differentiation F_{ST} estimate of 0.027 among all sub-populations was obtained, showing that genetic variation is mainly present within the sub-populations. The highest allelic richness, based on 26 individuals per sub-population, MNA and He, averaged over loci, were observed in the Kagera sub-population. The lowest values for the corresponding parameters were found in the sub-populations from Luwero, Mbarara north and Busoni.

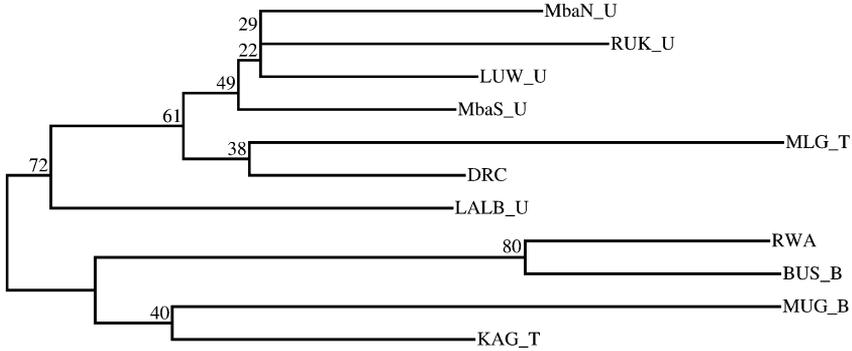


Figure 6. A Neighbour-Joining dendrogram illustrating the genetic divergence between the 11 cattle sub-populations based on Nei’s DA distance.

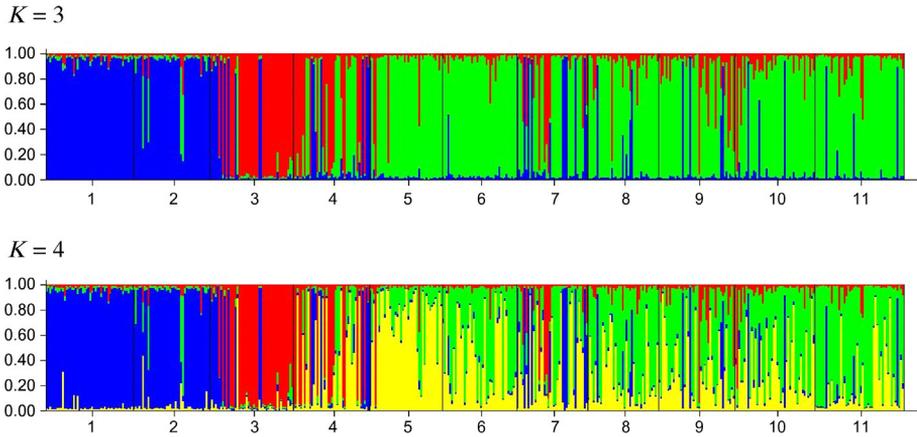


Figure 7. Estimated membership of the inferred clusters at the maximum value for $\ln \Pr(G/K)$ within each of the 10 runs for $K = 3$ and $K = 4$, respectively, the pre-assigned sub-populations 1 through to 11 are from the regions of RWA(1), BUS_B(2), MUG_B(3), KAG_T(4), MLG_T(5), DRC(6), LALB_U(7), MbaN_U(8), MbaS_U(9), LUW_U(10) and RUK_U(11).

The range of the H_o was between 0.67 for Busoni and 0.74 for Mbarara north, while that of the H_e was between 0.72 for Busoni and 0.79 for Kagera. Nine out of the 11 sub-populations showed significant positive F_{IS} estimates ($P < 0.05$), an indication of inbreeding within the herds (Tab. I).

The relationship between the sub-populations is illustrated in Figure 6. The pairwise F_{ST} estimates between the sub-populations are presented in Table II.

Furthermore, nearly all the sub-populations differentiated significantly ($P < 0.01$), with the exception of three pairs between the three sub-populations of Uganda, namely the Mbarara north, Mbarara south and Luwero, the pair between the Rwandan sub-population and the Busoni sub-population in Burundi, and also the pair between the Luwero sub-population in Uganda and the DRC cattle ($P < 0.05$).

To correctly determine the most probable structure/composition of clusters (K) in the data set, a plot of the natural logarithm of the posterior probability of the observed genotype (G) given the assigned clusters K ($\text{Ln Pr}(G/K)$) in each of 10 runs was plotted against the assigned cluster K , as illustrated by the graph shown in Appendix I. The highest value of $\text{Ln Pr}(G/K)$ for a given K is represented by a black dot in the graph. The maximisation of $\text{Ln Pr}(G/K)$ yields the most probable structure K . It should be noted that the plot and all points form a peak at $K = 4$, indicating a maximum $\text{Ln Pr}(G/K)$. Furthermore, in Figure 7, where each individual is graphically represented by a vertical line divided into K , colour segments may be well represented by $K = 4$, although $K = 3$ might represent the sub-populations in light of other considerations as discussed below.

3.3. Spatial distances

A spatial distance matrix of pairwise comparison between the sub-populations is presented in Table VI in Appendix II. It should be noted that the sub-populations were more dispersed between latitudes than between longitudes (Fig. 1).

3.4. Comparison between genetic, morphometric (size and shape) and spatial distances

Genetic, geographic, linear morphometric and geometric morphometric comparisons are presented in Table III. Spatial distances were not significantly related to the pairwise F_{ST} distances. A very close relationship was observed between body shape and body size.

4. DISCUSSION

The proceeding discussion (Sects. 4.1–4.3) attempts to describe the observed results of quantitative morphological analyses in relation to other patterns of variations such as spatial, genotypic and even geometric morphometric variations. Furthermore, these variation patterns are related to current and

pre-current production environments of the Ankole Longhorn cattle, as well as to implications for future use, development and conservation of this resource.

4.1. Morphological description

4.1.1. Size morphology

The morphological variation in size traits, among the 11 sub-populations, shows that the three sub-populations from Uganda of Mbarara north, Mbarara south and Luwero had significantly larger and longer horns than cattle from other regions. The cattle from the two Mbarara sub-populations and also the Luwero sub-population were significantly the tallest at the withers and had the longest body length. They also had significantly longer rumps and femoral length. They were, however, not significantly different from the cattle of Kagera in north-western Tanzania in terms of head length, heart chest girth, hip width and rump height. These ecotypes (from central Uganda and north-western Tanzania) graze the relatively large expanses of the wooded savannahs in the lowlands of the rift valley system.

Conversely, the cattle from DR Congo had the smallest heart chest girth and the shortest horns and were also significantly shorter in both rump and body length compared to almost all other populations. However, they were not significantly different from those of Mugamba in Burundi and Rukungiri in Uganda. These ecotypes are reared in the highland areas, where severe feed scarcity has been occurring increasingly over time, due to stiff competition for land with crop production as a result of rising human population densities in these regions. Under such circumstances, the larger-bodied animals, with relatively higher feed requirements, would be the first to suffer and consequently could thus have probably been progressively and systematically bred out over time in these highland production environments.

4.1.2. Stepwise discriminant analysis

The 10 most important traits separating regions include five each of the size and appearance measurements (*i.e.* horn length, horn circumference, horn tip interval, dewlap size, horn lower interval). This can, partly, be explained by the indigenous selection criteria of the different but related ethnic groups who keep the different races of the Ankole Longhorn cattle. Among the different ethnic groups in the study areas, morphological size traits and those of aesthetic value play different but equally important roles in the selection decisions made by herd owners. Indeed, aesthetic traits were highly ranked as selection criteria

by the herd owners from the three central Ugandan sub-populations, as documented by Ndumu *et al.* [22] and Wurzinger *et al.* [38].

4.1.3. Canonical discriminant analysis

The scatter plot from the MDS procedure (Fig. 4), graphically illustrating the Mahalanobis squared distances, places the 11 sub-populations in the expected groupings according to their variation in morphological size. The three central Ugandan sub-populations, which showed the smallest distances in the matrix and were significantly larger in size than the other sub-populations, are scaled away from the rest of the sub-populations and grouped together. On the other hand, the smaller-sized cattle from the highlands of DR Congo, as well as Mugamba and Busoni in Burundi, are scaled away to form their own group. The intermediate-sized cattle belonging to the sub-populations of Kagera and Malagarasi in Tanzania, Rwanda and the Ugandan Lake Albert area fall into neither of the two extreme groups.

4.1.4. Geometric morphometric description

Body size is considered as important as body shape when describing the differentiation of sub-populations. While centroid size from geometric morphometrics is an obvious choice, we decided to perform canonical discriminant analysis and individual analyses of variance for a set of distance measures (*e.g.* height at withers and body length). The results from discriminant analysis were expected to be similar [2].

Head shape represents those traits of appearance related to horn circumference, horn length and horn tip interval. The Mahalanobis head shape matrix (Tab. V in Appendix II) shows that the largest relative warp score distances are between the three central Ugandan sub-populations and the DRC cattle. In terms of magnitude, the next large distances are observed between the three central Ugandan populations and those from Busoni and Mugamba in Burundi. These pairwise distances are followed (in magnitude) by those between the three central Ugandan populations and the sub-population from the Lake Albert area in Uganda. This is in agreement with the multivariate analysis of the linear size morphometric measurements, where the three central Ugandan populations were found to be significantly different from the other populations from Burundi, Rwanda, DR Congo and also the Ugandan sub-populations from Rukungiri and the Lake Albert area.

The MDS scatter plot in Figure 5a illustrates that the Ugandan, Tanzanian and Rwandan sub-populations are separated from the rest. These differences can be further accounted for by the particular indigenous selection criteria of the herd

owners from central Uganda and north-western Tanzania, who tend to emphasise aesthetic traits more than what the cattle owners of the other sub-populations do. These findings are also consistent with the observations made by Wurzinger *et al.* [38], who found that appearance-related features such as coat colour and horns were ranked highly as selection criteria mainly in the Uganda Central and Kagera areas.

Like head shape and linear morphometric analyses, the largest body shape distances were found between four out of the five Ugandan sub-populations (except Lake Albert sub-population) and the other sub-populations, *i.e.* the Mugamba sub-population in Burundian highlands, followed by the DR Congo cattle. It is observed that aside from the Lake Albert sub-population, the rest of the sub-populations are grouped by their similarity in selection criteria and production systems within their respective countries, as observed in the study of Wurzinger *et al.* [38].

4.2. Genetic diversity

Gene differentiation analysis (F_{ST} index) shows that only 2.7% of the total genetic variation is explained by the differences among sub-populations. This value is much lower than that observed for differences among other cattle breeds, for instance the moderate differentiation levels of multilocus F_{ST} values of 6.0% observed among 12 African *Bos indicus* and *Bos taurus* cattle breeds [18] and 6.8% among 18 south European beef breeds [19]. A higher differentiation of 8.9% was, however, observed among beef cattle in traceability studies conducted in Italy [7], while a much higher F_{ST} value of 11.0% was found in the red Kandhari and Deoni cattle breeds in western India [36]. These Ankole Long-horn cattle sub-populations studied may have been separated for only a fewer number of generations within the African Great Lakes region, albeit with considerable migration involving exchange of genes between some sub-populations, part of which is still going on today.

The sub-populations of Luwero and Mbarara north from Uganda exhibited a lower genetic diversity in terms of MNA per locus. It is also observed that most of the sub-populations showed significant inbreeding within the herds.

Three out of the five Ugandan sub-populations had the least differentiation in terms of their insignificant pairwise F_{ST} estimates, but they all showed a rather large divergence from the sub-population of Mugamba in Burundi. This contrasts with the morphometric analyses of body and horn shapes and sizes, where larger distances were noted between three Ugandan sub-populations and the DR Congo cattle. Moreover, the Mugamba sub-population in Burundi had an equally large genetic differentiation from the DR Congo and Rwandan sub-populations and

Busoni sub-population in Burundi as well. It has to be pointed out that the reliability of the dendrogram (Fig. 6) is quite low. However, several observations from the dendrogram are consistent with a 're-assignment', through the Bayesian approach in the structure program, of the DR Congo sub-population from its cluster with the Malagarasi sub-population from Tanzania in the dendrogram (Fig. 6) to the Ugandan sub-populations, with which it has a similar genotypic admixture (Fig. 7). In the DR Congo, there is another Ankole sub-population belonging to the "Banyamulenge" herders (that was not sampled due to logistical limitations). The Banyamulenge sub-population is similar in morphology to the Ugandan Ankole Longhorn cattle. Probable gene flow between this population and the one sampled in the DR Congo might explain the similar structural assignment of the DR Congo cattle (sampled) with the Ugandan sub-populations.

The largest differentiation was present between the Mugamba sub-population in Burundi highlands and the sub-population in Malagarasi valley in Tanzania. Indeed, individuals from the two sub-populations fall into two separate clusters by posterior assignment, according to the structure program (Fig. 7). This might be explained by the location of the Mugamba sub-population on the rift valley escarpment, which is relatively difficult to access. This and socio-political upheavals prevailing in the country since colonial times could have played a large role in restricting this sub-population to this area.

All five sub-populations of Uganda, along with the north-western Tanzanian sub-population, form a single cluster with a minimum F_{ST} value of 0 between Mbarara south and Luwero sub-populations and a maximum F_{ST} value of 0.023 between the Luwero and Lake Albert sub-populations in the Western Rift Valley, implying that these six sub-populations are genetically closely related. This is confirmed by the Bayesian clustering method in the structure program, albeit with the exclusion of the Kagera sub-population in Tanzania and inclusion of the DR Congo cattle.

The low genetic differentiation between the Ugandan cattle sub-populations observed from the F_{ST} estimates and the results from the structure program can be attributed mainly to the nomadic lifestyle of the herd owners as well as to their traditional and cultural practices. Before administrative sedentarisation resulting from a ranching scheme in 1994, there was relatively unrestricted movement along the Cattle Corridor stretching from the south-western to the north-eastern part of the country.

After sedentarisation, gene flow occurs mainly as a result of cultural exchange of cattle, trade and other traditional practices of the herd owners, as well as of sporadic, semi-nomadic practices dictated by spells of severe periodic drought. Although the Lake Albert and Rukungiri sub-populations are located in the western part of the Rift Valley and highlands within Uganda, respectively, their

genetic proximity to Cattle Corridor sub-populations is due to an observable gene flow between these sub-populations, which can be accounted for by past and ongoing cattle trade and socio-cultural exchanges.

The political boundaries separating the otherwise spatially close cattle sub-populations in Tanzania, Uganda and Rwanda might explain the different genetic clustering observed in the results of the structure program. Notably, the north-western Tanzanian sub-population of Kagera comprises an admixture of individuals of similar major assignment as those from the sub-populations of Malagarasi also in Tanzania and that of Mugamba in Burundi.

The structuring and relatedness of the cattle of the sub-populations of Rwanda and that of Busoni in Burundi into a single cluster can be attributed to their spatial proximity. This group also represents a small- to medium-sized type of Ankole Longhorn cattle, often referred to in Rwanda as Inkuku or in Burundi as Inyaruguru (shorter and smaller-horned), which was herded in pre-colonial times by commoners and subjects of the kingdoms. And these are as distinguished from the almost legendary large-sized Ankole Longhorn cattle of the same time referred to as the Inyambu/Enyambu, which belonged mainly to royalty and to the rulers of the same kingdoms.

4.3. Morphology, genotype and geography

The genetic relationship results are not in accordance with findings from the phenotypic, morphometric and spatial analyses, as shown by the results from Mantel tests presented in Table III. This can largely be attributed to the location of the microsatellites used in this study at selectively neutral loci. However, there is a correlation between all morphometric distances of shape and size with spatial distances, partly explained by the production environments and selection criteria as discussed above.

5. CONCLUSION

The 11 Ankole Longhorn cattle sub-populations can be considered to consist of four main groups with slight genetic distinctions. All analyses, morphometric, genetic and geographic, consistently show that the smallest distances are between the central Ugandan sub-populations of Mbarara north, Mbarara south and Luwero, indicating that these three sub-populations are homogenous and form a single genetic entity quite close to the sub-populations of the Lake Albert area and Rukungiri, also in Uganda, as well as the cattle from DR Congo. The latter sub-populations are structurally different from Kagera cattle in Tanzania in that they are comprised mainly of individuals of similar structure as those from

the sub-populations of Malagarasi, also in Tanzania, and that of Mugamba in Burundi. The Busoni sub-population in Burundi and the Rwandan cattle form a separate cluster. The fourth structure was less overt throughout all analyses, but was revealed by the model-based Bayesian method in structure. It consists mainly of cattle from the Malagarasi river valley and, to a lesser extent, contains individuals of an admixture of the Ugandan and DR Congo sub-populations.

ONLINE MATERIAL

The supplementary file (Appendixes I and II) supplied by the authors is available at: <http://www.gse-journal.org>.

Appendixes I and II: Figure and Tables I –VI.

ACKNOWLEDGEMENTS

We wish to express our sincere gratitude to the Austrian Government for financial support. We would also thank the International Livestock Research Institute (ILRI) for hosting and coordinating project activities, as well as for providing laboratory logistical support for molecular genetic analyses. We are most grateful to the NARS in the collaborating countries of Burundi, DR Congo, Rwanda, Uganda and Tanzania. Our particular thanks go to the DG – ISAR, Dr. M. Bagabe, the DG ISABU, Dr. Deo Manirakiza, Dr. Katunga-Musale, Dr. Rutagwenda, Dr. D. Rwemalika, Mr. J. Mutabazi, Dr. C. Rutebalika, Dr. Rusita, Dr. Deo Mwesigye, Dr. M. Kasawe, Mr. Klaus Wolgenannt, Mr. Nathan Kabahigi, Mr. Apuuli Bomera, Mr. Nubuhoro, Dr. Deo Luvakule, Dr. M.L. Orotta. We would also like to thank Dr. M. Agaba, Mr. J. Mwakaya and Mr. J. Audho for their kind assistance during molecular laboratory investigations. And above all, we gratefully acknowledge the cooperation of the cattle owners.

REFERENCES

- [1] Adams D.C., Rohlf F.J., Slice D.E., Geometric morphometrics: ten years of progress following the ‘revolution’, *Ital. J. Zool.* 71 (2004) 15–16.
- [2] Alibert P., Moureau B., Dommergues J.-L., David B., Differentiation at a microgeographical scale within two species of ground beetles, *Carabus auronitens* and *C. nemoralis* (Coleoptera, Carabidea): a geometrical morphometric approach, *Zool. Scr.* 30 (2001) 299–311.
- [3] Baker R.L., Rege J.E.O., Genetic resistance to disease and other stresses in improvement of ruminant livestock in the tropics, in: *Proceedings of the 5th*

- World Congress on Genetics Applied to Livestock Production, 7–12 August 1994, Vol. 20, University of Guelph, Guelph, pp. 405–412.
- [4] Belkhir K., Borsa P., Chikhi L., Raufaste N., Bonhomme F., Genetix version 4.05, logiciel sous Windows TM pour la génétique des populations, Laboratoire Génome, Population, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France (2004) <http://univ-montp2.fr/~genetix/genetix/intro.htm> [consulted: 20 March 2008].
- [5] Bonnet E., van de Peer Y., zt: a software tool for simple and partial Mantel tests, *J. Stat. Soft.* 7 (2002) 1–12. <http://www.psb.ugent.be/~erbon/mantel/> [consulted: 20 March 2008].
- [6] Bookstein F.L., Biometrics, biomathematics and the morphometric synthesis, *Bull. Math. Biol.* 58 (1996) 313–365.
- [7] Ciampolini R., Cetica V., Ciani E., Mazzanti E., Fosella X., Marroni F., Biagetti M., Sebastiani C., Papa P., Filippini G., Cianci D., Presciuttini S., Statistical analysis of individual assignment tests among four cattle breeds using fifteen STR loci, *J. Anim. Sci.* 84 (2006) 11–19.
- [8] Faulkner D.E., Epstein H., The indigenous cattle of British dependent territories in Africa, Publication of the Colonial Advisory Council for Agriculture, Animal Health and Forestry, No. 5, HMSO, London, 1957.
- [9] Freeman A.R., Hoggart C.J., Hanotte O., Bradley D.G., Assessing the relative ages of admixture in the bovine hybrid zones of Africa and the Near East using X chromosome haplotype mosaicism, *Genetics* 173 (2006) 1503–1510.
- [10] Goudet J., FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices, Version 2.9.3 (2001) <http://www2.unil.ch/popgen/softwares/fstat.htm> [consulted: 20 March 2008].
- [11] Gregory K.E., Trail J.C., Marples H.J., Kakonge J., Characterisation of breeds of *Bos indicus* and *Bos taurus* cattle for maternal and individual traits, *J. Anim. Sci.* 60 (1985) 1165–1174.
- [12] Grimaud P., Mpairwe D., Chalimbaud J., Messad S., Faye B., The place of Sanga cattle in dairy production in Uganda, *Trop. Anim. Health Prod.* 39 (2007) 217–227.
- [13] Guo S.W., Thompson E.A., Performing the exact test of the Hardy-Weinberg proportion for multiple alleles, *Biometrics* 48 (1992) 361–372.
- [14] Hanotte O., Tawah C.L., Bradley D.G., Okomo M., Verjee Y., Ochieng J., Rege J.E., Geographic distribution and frequency of a taurine *Bos taurus* and indicine *Bos indicus* Y specific allele amongst sub-saharan African cattle breeds, *Mol. Ecol.* 9 (2000) 387–396.
- [15] Hanotte O., Bradley D.G., Ochieng J.W., Verjee Y., Hill E.W., Rege J.E., African pastoralism: genetic imprints of origins and migrations, *Science* 296 (2002) 336–339.
- [16] Hardy O.J., Vekemans X., SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels, *Mol. Ecol. Notes* 2 (2002) 618–620 <http://www.ulb.ac.be/sciences/ecoevol/spagedi.html> [consulted: 20 March 2008].

- [17] Hoffmann I., Marsan P.A., Barker J.S.F., Cothran E.G., Hanotte O., Lenstra J.A., Milan D., Wiegand S., Simianer H., New MoDAD marker sets to be used in diversity studies for the major farm animal species: recommendations of a joint ISAG/FAO working group, in: Proceedings of the 29th International Conference on Animal Genetics, 11–16 September 2004, Meiji University, Tokyo, Japan, p. 123 (abstr).
- [18] Ibeagha-Awemu E.M., Erhard G., 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 (2005) 12–20.
- [19] Jordana J., Alexandrino P., Beja-Pereira A., Bessa I., Canon J., Carretero Y., Dunner S., Laloë D., Moazami-Goudarzi K., Sanchez A., Ferrand N., Genetic structure of eighteen local south European beef cattle breeds by comparative *F*-statistics analysis, *J. Anim. Breed. Genet.* 120 (2003) 73–87.
- [20] Mantel N., The detection of disease clustering and a generalized regression approach, *Cancer Res.* 27 (1967) 209–220.
- [21] Marshall F., The origins and spread of domestic animals in East Africa, in: Blench R.M., MacDonald K.C. (Eds.), *The Origins and Development of African livestock: Archeology, Genetics, Linguistics and Ethnography*, UCL Press, London, 2000, pp. 191–221.
- [22] Ndumu D.B., Baumung R., Wurzinger M., Drucker A.G., Okeyo A.M., Semambo D., Sölkner J., Performance and fitness traits versus phenotypic appearance in the African Ankole Longhorn cattle: a novel approach to identify selection criteria for indigenous breeds, *J. Livest. Sci.* 113 (2008) 234–242.
- [23] Nei M., Tajima F., Tateno Y., Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data, *J. Mol. Evol.* 19 (1983) 153–170.
- [24] Norval R.A.I., Host susceptibility to infestation with *Amblyomma hebraeum*, *Insect Sci. Appl.* 13 (1992) 489–494.
- [25] Okello S., Sabiiti E.N., Milk production of indigenous Ankole cattle on Uganda as influenced by seasonal variations in temperature, rainfall and feed quality, *Makerere Univ. Res. J.* 1 (2006) 73–92.
- [26] Ota T., DISPAN – Genetic distance and phylogenetic analysis, Pennsylvania State University, Pennsylvania, USA, 1993.
- [27] Palling R.W., Mpangala C., Littikhuizen B., Sibomana G., Exposure of Ankole cattle and crossbreds to theileriosis in Rwanda, *Trop. Anim. Health Prod.* 23 (1991) 203–213.
- [28] Pritchard J.K., Stephens M., Donnelly P., Inference of population structure using multilocus genotype data, *Genetics* 155 (2000) 945–959.
- [29] Raymond M., Rousset F., GENEPOP, population genetics software for exact tests and ecumenicism, Version 3.4, *J. Hered.* 86 (1995) 248–249. <http://genepop.curtin.edu.au/> [consulted: 20 March 2008].
- [30] Rohlf F.J., Marcus L.F., A revolution in morphometrics, *Trends Ecol. Evol.* 8 (1993) 129–132.
- [31] Rosenberg N.A., DISTRUCT: a program for the graphical display of population structure, *Mol. Ecol. Notes* 4 (2004) 137–138.

- [32] Saitou N., Nei M., The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4 (1987) 406–425.
- [33] Sambrook J., Fritsch E.F., Maniatis T., *Molecular cloning: a laboratory manual*, 2nd edn., Cold Spring Harbour Laboratory Press, 1989.
- [34] SAS[®] for Windows, Cary, NC: SAS[®] Institute Inc., SAS[®]/STAT Software Release 9.1.3, 2003.
- [35] Schumann H., *Richtig messen – richtig züchten*, Bayer, Landwirtschaftsverlag, München, Basel, Wien, 1968.
- [36] Sodhi M., Mukesh M., Mishra B.P., Mitkari K.R., Prakesh B., Ahlawat S.P., Evaluation of genetic differentiation in *Bos indicus* cattle breeds from Marathwada region of India using microsatellite polymorphism, *Anim. Biotechnol.* 16 (2005) 127–137.
- [37] Weir B.S., Cockerham C.C., Estimating *F*-statistics for the analysis of population structure, *Evolution* 38 (1984) 1358–1370.
- [38] Wurzinger M., Ndumu D., Baumung R., Drucker A., Okeyo A.M., Semambo D.K., Byamungu N., Sölkner J., Comparison of production systems and selection criteria of Ankole cattle by breeders in Burundi, Rwanda, Tanzania and Uganda, *Trop. Anim. Health Prod.* 38 (2006) 571–581.
- [39] Zechner P., Zohman F., Sölkner J., Bodo I., Habe F., Marti E., Brem G., Morphological description of the Lipizzan horse population, *Livest. Prod. Sci.* 69 (2001) 163–177.