Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

Accuracy of pastoralists' memory-based kinship assignment of Ankole cattle: a microsatellite DNA analysis

D.R. Kugonza[1,2], G.H. Kiwuwa[1,†], D. Mpairwe[1], H. Jianlin[2,3], M. Nabasirye[1], A.M. Okeyo[2] & O. Hanotte[2,‡]

1 Department of Agricultural Production, School of Agricultural Sciences, College of Agricultural and Environmental Sciences, Makerere University, Kampala, Uganda
2 International Livestock Research Institute (ILRI) Nairobi, Kenya
3 CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing, China

Keywords
Cattle dyads; indigenous knowledge; memorized records; parentage; Relatedness.

Correspondence
D.R. Kugonza, Department of Agricultural Production, School of Agricultural Sciences, College of Agricultural and Environmental Sciences, Makerere University, PO Box 7062, Kampala, Uganda. Tel: +256 414 532269; Fax: +256 414 541641; E-mail: donkugonza@agric.mak.ac.ug
†Deceased (2009).
‡Present address: School of Biology, Centre for Genetics and Genomics, University of Nottingham, Nottingham NG7 2RD, UK.

Received: 20 September 2010;
accepted: 4 June 2011

Summary

This study aimed to estimate the level of relatedness within Ankole cattle herds using autosomal microsatellite markers and to assess the accuracy of relationship assignment based on farmers' memory. Eight cattle populations (four from each of two counties in Mbarara district in Uganda) were studied. Cattle in each population shared varying degrees of relatedness (first-, second- and third-degree relatives and unrelated individuals). Only memory-based kinship assignments which farmers knew with some confidence were tested in this experiment. DNA isolated from the blood of a subsample of 304 animals was analysed using 19 microsatellite markers. Average within population relatedness coefficients ranged from 0.010 ± 0.005 (Nshaara) to 0.067 ± 0.004 (Tayebwa). An exclusion probability of 99.9% was observed for both sire–offspring and dam–offspring relationships using the entire panel of 19 markers. Confidence from likelihood tests performed on 292 dyads showed that first-degree relatives were more easily correctly assigned by farmers than second-degree ones ($p < 0.01$), which were also easier to assign than third-degree relatives ($p < 0.01$). Accuracy of kinship assignment by the farmers was $91.9\% \pm 5.0$ for dam–offspring dyads, $85.5\% \pm 3.4$ for sire–offspring dyads, $75.6\% \pm 12.3$ for half-sib and $60.0\% \pm 5.0$ for grand dam–grand offspring dyads. Herd size, number of dyads assigned and length of time spent by the herder with their cattle population did not correlate with error in memorizing relationships. However, herd size strongly correlated with number of dyads assigned by the herder ($r = 0.967$, $p < 0.001$). Overall, we conclude that memorized records of pastoralists can be used to trace relationships and for pedigree reconstruction within Ankole cattle populations, but with the awareness that herd size constrains the number of kinship assignments remembered by the farmer.

Introduction

Ankole cattle are indigenous to the eastern region of Africa where they are the main cattle breed of Uganda. They are also found in Burundi, Democratic Republic of Congo, Rwanda and Tanzania. They are mainly owned by Bahima pastoral communities to whom they are important for milk, ghee and meat production and where they fulfil many cultural functions (Infield 2002; Nakimbugwe & Muchunguzi

  doi:10.1111/j.1439-0388.2011.00937.x

2003). Bahima people are well known for their great love for cattle and are breeders who mainly select for one colour and horn type (Kugonza *et al.* 2011). However, traditional Ankole cattle keepers do not keep written records of their herd's history, and it is believed that to achieve their breeding objectives, they rely on memory-based information systems (FAO, 2003, 2004).

Identification of kinship relationships between animals plays a central part in animal breeding, quantitative genetics, conservation biology and ecology. Interest in investigating livestock genetic diversity and inference of family relationships using molecular markers at DNA level has been growing consistently over the past two decades. Microsatellites are still recognized as very effective genetic markers for this purpose because they are commonly found across the genome, typically display many alleles per locus and are co-dominant markers (Barbará *et al.* 2007). Parentage and kinship studies using microsatellite markers have been conducted in several species. For instance, microsatellite markers have been used in progeny tests to assess the paternity of probable offspring of Gir dairy sires (Baron *et al.* 2002), routine parentage testing in Portuguese autochthonous horse breeds (Luis *et al.* 2002), parentage control in pigs (Putnová *et al.* 2003) and to prove kin selection in turkeys (Krakauer 2005). Attempts to use these markers to classify animals by relatedness showed that it is possible to discriminate unrelated pairs from half-sibs, full-sibs and parent–progeny pairs in mice (Blouin *et al.* 1996) and in baboons (Van Horn *et al.* 2008).

Relationship coefficient or coefficient of relatedness or relatedness or R is defined as the probability that any two individuals share a given gene by virtue of being descended from a common ancestor, and this coefficient calculates the proportion of genes that two individuals have in common as a result of their genetic relationship (http://www.genetic-genealogy.co.uk). The formula for computing R is:

$$(R_{XY}) = \Sigma\ (1/2)^n \text{ (Falconer 1989)}$$

where R_{XY} is the coefficient of relationship between the two relatives X and Y and n is the number of connecting links or paths separating them. Σ refers to the fact that if there is more than one connecting path, the paths are computed separately and their coefficients are then added together.

Microsatellite parentage assignment is very effective when exclusion probabilities (EP) are calculated (Heyen *et al.* 1997), as these estimate the probability that an animal, e.g. sire, is correctly excluded as parent of a specific offspring. The EP value is influenced by the number of microsatellite loci typed and their heterozygosities (Blouin *et al.* 1996); the degree of relatedness of the candidate parent to the true parent of an offspring; and the pool size of candidate parents (Sherman *et al.* 2004). Usha *et al.* (1995) and Heyen *et al.* (1997) suggested the use of at least five microsatellite markers to achieve 0.99 probability of exclusion of an incorrect sire. A panel of 15 microsatellites yielded an EP of 99.998% during horse paternity testing (Tozaki *et al.* 2001). Besides genetic relatedness estimated using genetic markers, other factors such as age, sex and physical proximity of dyads are also important information when assessing parentage probability.

While DNA marker–based pedigree assignment has been proved to be a feasible option for commercial livestock producers in the developed world (Dodds *et al.* 2005; Gomez-Raya *et al.* 2007; Van Eenennaam *et al.* 2007), the same may not be true in developing countries where access to relevant molecular biology infrastructure and cost of molecular genetic analysis might be an issue. In such context, parentage information based on human memory of relationships may represent an important alternative for acquiring pedigree information, such as required in breeding improvement programmes. The objectives of this study therefore were to determine the degree of relatedness within Ankole cattle populations using 19 microsatellite markers and to assess the accuracy level of relationship assignment of selected cattle dyads based on the farmers' memory.

Materials and methods

Study sites and experiment design

This study used a two-stage random sampling technique. Two neighbouring counties of Mbarara district (Kazo and Nyabushozi, 0°4′–0°12′N; 30°47′–30°49′E), which have the highest concentration of Ankole cattle in Uganda, were purposively selected. The counties are located in Western Uganda, a savannah grassland–predominated area with elevations of 760–900 m above mean sea level. The area receives an annual rainfall of 875–1200 mm, and it is occasionally hit by intense drought. Within each county, Bahima farmers having an experience of keeping Ankole cattle of at least 20 years and owning semiclosed pure Ankole herds numbering at least 50 breeding animals (49 cows and one bull) were identified. Four farms were then selected from each

of the two counties (Kasiisi, Kituuha, Nasasira and Rwokusooka in Kazo County; Kaibanda, Museveni, Mwesigye and Tayebwa in Nyabushozi County), avoiding farms in the same parish, in case they may have shared sires. An additional cattle population at a government-owned Nshaara ranch was incorporated as a control population for comparative purposes, because the ranch keeps written records on parentage of the cattle.

At each farm, the respondent farmer was interviewed about the following: (i) their age, (ii) number of years spent with the cattle population prior to the study, (iii) presence in the herd of cattle that share genetic relationships and (iv) herd population size. The questions were specifically asked to determine whether these factors influence the level of accuracy of the respondent in assigning relationships between animals of a particular population. The respondent was then asked to identify all the animals in the population sharing a genetic relationship, relying only on memory. Exception was at Nshaara ranch, where written records were used.

From each of the cattle populations, the respondent farmer identified animal pairs (dyads) sharing a specific genetic relationship category. The categories were parent–offspring (i.e. dam–daughter, dam–son, sire–daughter or sire–son) and full-sibs (calves having same parents) collectively grouped as first-degree (1°) relatives [a theoretical mean 50% of alleles shared expected to be identical by descent (IBD)]; grandparent–grand offspring (i.e. grand dam–granddaughter, grand dam–grandson) and paternal half-sibs grouped as second-degree (2°) relatives (average 25% of alleles shared IBD); great grandparent–great grand offspring (i.e. great grand dam–great granddaughter, great grand dam–great grandson) grouped as third-degree (3°) relatives (12.5% of alleles shared IBD). The numbers of dyads sharing the different relationship categories are presented in Table 1. We opted to assess only the accuracy of memory-based assignments that the farmer felt confident to tell us rather than to assess the relationships between all dyads within the herds. While the latter could have provided us with the information on which proportion of first-, second- or third-degree dyads remain recorded in the memory of the farmers, we were constrained by logistic as well as financial challenges. Also, it may be argued that a breed improvement programme based on farmers' memory will in any case only use information on known dyads relationships.

The number of dams and breeding bulls in each herd is also indicated in Table 1, as well as the number of animals analysed in this study. At the time of this study, only one herd (Tayebwa) had more than one bull known to be actively breeding (Table 1), although we cannot exclude that following herd mixing, e.g. in a communal pastoral area, calves may be sired by non-resident bulls. For each herd in this study, the resident bull was always sampled.

Table 1 Dyad relationships in Ankole cattle populations and accuracy of parentage and kinship assignment of dyads by farmers

				Number of dyads assigned by farmer[a]					
Population sampled	Number of dams[b]	Number of breeding bulls	Number of animals sampled	Dam–offspring[c]	Sire–offspring[c]	Half-sibs[d]	Grand dam–Grand offspring[d]	G/grand dam–G/grand offspring[e]	Total
Kaibanda	120	1	39	22 (22)	15 (13)	5 (5)	3 (2)	2(2)	47 (44)
Kasiisi	80	1	28	7 (6)	8 (7)	1 (0)	7 (3)	–	23 (16)
Kituuha	98	1	26	7 (7)	8 (7)	10 (8)	5 (3)	1(1)	31 (26)
Mwesigye	101	1	32	17 (10)	19 (16)	7 (4)	2 (1)	2(2)	47 (33)
Nasasira[f]	120	1	44	21 (20)	–	1 (1)	4 (2)	1(1)	27 (24)
Nshaara	250	1	41	20 (20)	3 (2)	–	–	1(1)	24 (24)
Rwokusooka	100	1	52	28 (26)	14 (12)	16 (12)	7 (6)	–	65 (56)
Tayebwa	109	3	42	13 (13)	3 (3)	5 (4)	7 (4)	–	28 (24)
All populations	978	10	304	135 (124)	70 (60)	45 (34)	35 (21)	7(7)	292 (246)

[a]Number of correctly assigned dyads in parentheses.

[b]Kituuha owned an additional (proximal) herd of 108 dams, Mwesigye had two of 104 and 95 while Tayebwa also had two of 88 and 107 dams, respectively.

[c]Dyads sharing a first-degree genetic relationship.

[d]Dyads sharing a second-degree genetic relationship.

[e]Great grand dam–great grand offspring dyads sharing a third-degree genetic relationship.

[f]The resident bull was recently acquired and had no known progeny in the herd.

Sample collection and DNA extraction

Blood samples were collected from animals in each of the selected populations by jugular puncture, using 10-ml disposable syringes, and were spread on FTA Whatman® filter paper (Whatman® Bioscience, Maidstone, UK) to dry in open air. The filter papers were thereafter labelled, packed and taken for laboratory analysis at the International Livestock Research Institute, Nairobi, Kenya. FTA purification reagent (Whatman® Bioscience) and Tris–EDTA (TE) buffer, pH 7.6, were then used to prepare DNA.

Microsatellite markers and genotyping

Laboratory work was performed at the International Livestock Research Institute (ILRI), Nairobi, Kenya. Twenty bovine microsatellite markers recommended for the measurement of domestic animal diversity (FAO, 2002) were used in DNA genotyping. The markers were as follows: *AGLA293*, *BM1824*, *BM2113*, *ETH152*, *ETH225*, *ILSTS005*, *ILSTS006*, *ILSTS008*, *ILSTS013*, *ILSTS023*, *ILSTS028*, *ILSTS033*, *ILSTS036*, *ILSTS50*, *ILSTS103*, *MGTG4B*, *TGLA53*, *TGLA122*, *TGLA126* and *TGLA227*. PCR amplification was performed in a total reaction volume of 10 μl on a GeneAmp® thermocycler 9700 (*PE* Applied Biosystems, Foster City CA, USA). Each PCR contained 20–50 ng template DNA, 1 μl of 2 mM $MgCl_2$, 0.5 μl of 0.125 mM of dNTPs, 0.1 μl of each of the forward and reverse primers, and 0.5 units of enzyme *Taq* DNA polymerase (Promega, Madison WI, USA). All amplifications included an initial denaturing step of 5 min at 95°C, followed by 35 cycles of 30 s at 95°C, 1 min at the primer annealing temperature (50–65°C) and 1-min extension at 72°C. Final extension step was for 10 min at 72°C followed by storage at 4°C.

PCR products were loaded and separated on a 4.5% denaturing polyacrylamide gel using an automated ABI377 DNA Sequencer (Applied Biosystems) and internal lane size standard Genescan-350 TAMRA (Applied Biosystems). Microsatellite fragments were analysed using GENESCAN™, version 3.1.2 software, and allele sizes were determined with the GENOTYPER™, version 2.0 software (Applied Biosystems). The third-order least squares was used for allele size calling.

Statistical analyses and pedigree reconstruction

Statistical analysis was performed on data generated from nineteen markers only, as genotyping raw data from marker *ILSTS008* resulted in many artefact peaks and the locus was not considered further. Relatedness between dyads was calculated using the KINSHIP version 1.3.1 software (Goodnight & Queller 1999). Relatedness ranged from −1 to +1 with positive values, indicating that two individuals share more alleles than expected by chance and negative values indicating that two individuals share fewer alleles than expected by chance. Background relatedness in each population was determined as the mean relationship between cows in the particular population (with calves excluded to minimize the inclusion of genetically related but unassigned dyads). The dyads to which the farmers assigned a relationship could have been (i) related at the expected recorded first, second or third degree, (ii) related but wrongly assigned a degree relationship and (iii) unrelated but assigned as related.

An allele-sharing matrix for each population was generated using MICROSATELLITE TOOLKIT, version 3.1 (http://animalgenomics.ucd.ie/sdpark/ms-toolkit/). Memory-based kinship assignments assessment was made at two levels. Sire–offspring and dam–offspring dyads were examined using the likelihood approach (Garcia *et al.* 2002). The average EP for non-exclusion of candidate parents was calculated using the program CERVUS version 3.0 (Kalinowski *et al.* 2007) (http://www.fieldgenetics.com/). Here, the power of a set of genetic markers to exclude candidate parents was systematically computed into an EP (Garcia *et al.* 2002). This exclusion technique is based on Mendelian rules of inheritance and uses incompatibilities between parents and offspring to reject particular parent–offspring hypotheses (Jones & Arden 2003). In the current study, using CERVUS software, a total of 10 000 simulations were conducted to assess the significance of the difference between the candidate parent and an arbitrary randomly chosen individual. LOD scores (logarithm of the likelihood ratio to base *e*) and Delta statistics were generated to assess the reliability of assigning parentage.

Significant tests for half-sib and grand dam–grand offspring assignments were carried out using the KINSHIP software. The expected coefficients of identity by descent (IBD) were 0.5 for dyads with a first-degree relative relationship, 0.25 for dyads with a second-degree relative relationship and 0.125 for dyads with a third-degree relative relationship (Blouin 2003). There was no need for adjusting the IBD coefficients because the threshold/mean background relatedness (0.034) within populations was smaller than the lowest IBD coefficient (0.125) expected

between dyad members before exclusion. Assigned dyads were related at the expected level (first-, second- or third-degree relatives). Other dyads were either related but with a wrongly assigned degree of relationship or unrelated but assigned as related.

To evaluate possible factors that may influence the accuracy of parentage association, correlation analysis were conducted between level of error in dyad relationship assignment with herd size; length of time spent by farmer with the cattle population; and number of assigned dyads. Validated dyads in respective populations were then used to reconstruct population pedigrees. This was done with PEDIGRAPH^{TM} version 2.2 (Garbe & Da 2005).

Results

Variability and statistical power of the selected loci for pedigree analysis

In total, 304 individuals were genotyped, including 183 candidate offspring, 106 candidate dams, 10 candidate sires, 47 candidate half-sibs, 22 candidate grand dams and seven candidate great grand dams. Animals were used in one or more categories. First-degree relative relationships included 270 animals in 205 dyads, second-degree relatives included 135 animals in 80 dyads, and third-degree relatives included 14 animals in seven dyads (Table 1). Summary statistics for the nineteen markers used in the analysis of the cattle (n = 304) are given in Table 2. The number of samples for which markers were informative ranged between 266 (*ILSTS033*) and 302 (*ILSTS006*). Genotype information was not considered at five samples for four markers, 15 samples for three markers, 34 samples for two markers and 104 samples for one marker following non-amplification or weak signal, preventing unambiguous allele size calling. Complete microsatellite genotyping information with the 19 markers was obtained for 146 animals.

The number of alleles per locus detected for all the animals ranged from six to nineteen with an average of 10.53, while the mean expected heterozygosity (He) was 0.727 (range: 0.465–0.839). Locus *ILSTS033* had the lowest number of heterozygote animals (n = 117), while *ILSTS006* had the highest (n = 257). The mean polymorphic information content (PIC) was 0.688 (range: 0.403–0.817). Probabilities of exclusion per locus ranged from 0.492 to 0.890 when only information of one parent was

Table 2 Number of individuals (n) genotyped, number of samples that amplified (na), number of alleles (k), observed heterozygosity (Ho), expected heterozygosity (He), polymorphic information content (PIC), probabilities of exclusion for one parent known (Excl 1) or both parents known (Excl 2) and loci conformance to Hardy–Weinberg expectations for 304 Ankole cattle at nineteen microsatellite loci

Locus	n	na	k	Ho	He	PIC	Excl 1	Excl 2	HW	Null frequency
ILSTS005	304	290	9	0.655	0.657	0.608	0.753	0.584	NS	−0.0036
ILSTS006	304	302	10	0.851	0.821	0.795	0.530	0.356	NS	−0.0203
ILSTS013	304	300	6	0.480	0.465	0.403	0.890	0.770	NS	−0.0143
ILSTS023	304	290	7	0.531	0.666	0.599	0.768	0.612	***	0.1137
ILSTS028	304	299	16	0.736	0.722	0.677	0.680	0.507	NS	−0.0178
ILSTS033	304	266	7	0.440	0.573	0.542	0.809	0.633	***	0.1388
ILSTS036	304	273	17	0.806	0.839	0.817	0.492	0.323	NS	0.0182
ILSTS050	304	301	10	0.791	0.796	0.764	0.584	0.405	NS	0.003
TGLA53	304	295	19	0.715	0.814	0.797	0.515	0.341	NS	0.0647
ILSTS103	304	281	8	0.751	0.754	0.710	0.654	0.477	NS	−0.0002
TGLA122	304	301	11	0.738	0.807	0.779	0.559	0.382	NS	0.0429
TGLA126	304	273	6	0.714	0.759	0.716	0.649	0.472	NS	0.0283
ETH152	304	292	6	0.661	0.631	0.566	0.790	0.637	NS	−0.0257
ETH225	304	293	11	0.747	0.800	0.770	0.571	0.394	NS	0.0324
TGLA227	304	289	14	0.654	0.761	0.724	0.629	0.451	**	0.0717
AGLA293	304	299	13	0.853	0.805	0.781	0.546	0.370	**	0.0387
BM1824	304	301	6	0.568	0.568	0.523	0.824	0.662	NS	−0.0032
BM2113	304	301	13	0.811	0.784	0.749	0.602	0.423	NS	−0.0175
MGTG4B	304	292	11	0.760	0.790	0.758	0.588	0.410	NS	0.0179
All		304	200				0.9998	0.9999	NS	0.0205
Mean		291.5	10.53	0.698	0.727	0.688				

NS = $p > 0.05$.
**$p < 0.01$.
***$p < 0.001$.

available (Excl 1) and from 0.323 to 0.770 when information on second parent was included (Excl 2). The total exclusion power predicted by simulation was 99.98% for first parent and 99.99% for second parent. As the number of alleles per locus increased beyond ten, PIC, Ho, He, Excl 1 and Excl 2 improved (Figure 1). However, the trend of these parameters for less than ten alleles was not discernable. The Hardy–Weinberg analysis indicated that four (*ILSTS023*, *ILSTS033*, *TGLA227* and *AGLA293*) of the nineteen markers studied were not in equilibrium (data not shown). This is not surprising given that genetic substructuring is likely with animals within herd more related to each other than between herds. Also, given the selection practice of the farmers, none of the populations may be considered as panmictic. The four markers were therefore not excluded in the analysis given their informativeness for pedigree analysis. Null-allele frequency estimates were between 0 and 13.8%, while the mean for all loci assessed in this study was 2.1%. The highest proportion of alleles shared (0.42 ± 0.04) was between Kaibanda and Tayebwa populations, while the lowest was (0.38 ± 0.04) between Kituuha and Nasasira; for more information, see Supporting information (Table S1).

Kinship analysis and pedigree reconstruction

The values and the distribution of relatedness coefficients for dyads in all the herds are presented in Table 3 and Figure 2a,b. The marker-based mean relatedness between all the dyads ranged between

Table 3 Mean coefficient of relatedness (r_{xy}) within Ankole cattle populations

	Coefficient of relatedness +SE[1]	
Population	All dyads (n)	Excluding related dyads[2] (n)
Kaibanda	$0.070^{a} \pm 0.006$ (741)	$0.028^{a} \pm 0.004$ (113)
Kasiisi	$0.050^{b} \pm 0.009$ (378)	$0.026^{a} \pm 0.001$ (85)
Kituuha	$0.149^{c} \pm 0.010$ (325)	$0.025^{a} \pm 0.002$ (65)
Mwesigye	$0.061^{b} \pm 0.007$ (496)	$0.045^{b} \pm 0.005$ (118)
Nasasira	$0.055^{b} \pm 0.006$ (946)	$0.027^{a} \pm 0.003$ (146)
Nshaara	$0.014^{d} \pm 0.005$ (820)	$0.010^{c} \pm 0.005$ (144)
Rwokusooka	$0.066^{a} \pm 0.004$ (1326)	$0.044^{b} \pm 0.004$ (120)
Tayebwa	$0.083^{e} \pm 0.005$ (861)	$0.067^{d} \pm 0.004$ (125)
All populations	0.063 ± 0.002 (5893)	0.034 ± 0.005 (916)

[1]Difference in mean coefficient of relatedness between all dyads and excluding the related dyad categories was significant at $p < 0.05$ for all populations except Kaibanda and Rwokusooka.
[2]Farmers memory information, only cows were used, heifers and calves were excluded.
Numbers of all possible dyads in parentheses as calculated using Kinship software (Goodnight & Queller 1999).

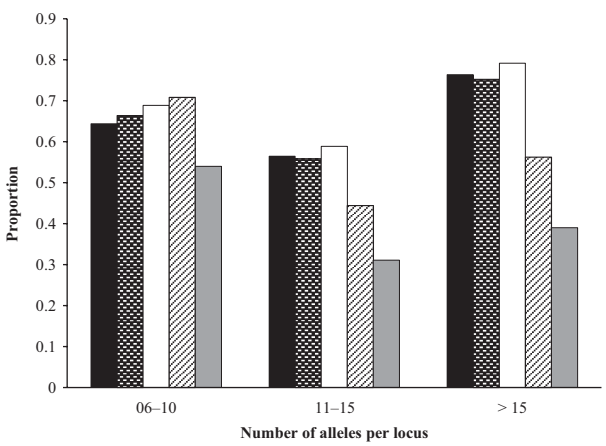

Figure 1 Variation of polymorphic information content (solid bars), observed heterozygosity (dotted bars), expected heterozygosity (open bars), exclusion probability of parent 1 (stripped bars) and exclusion probability of parent 2 (stippled bars) with number of alleles.

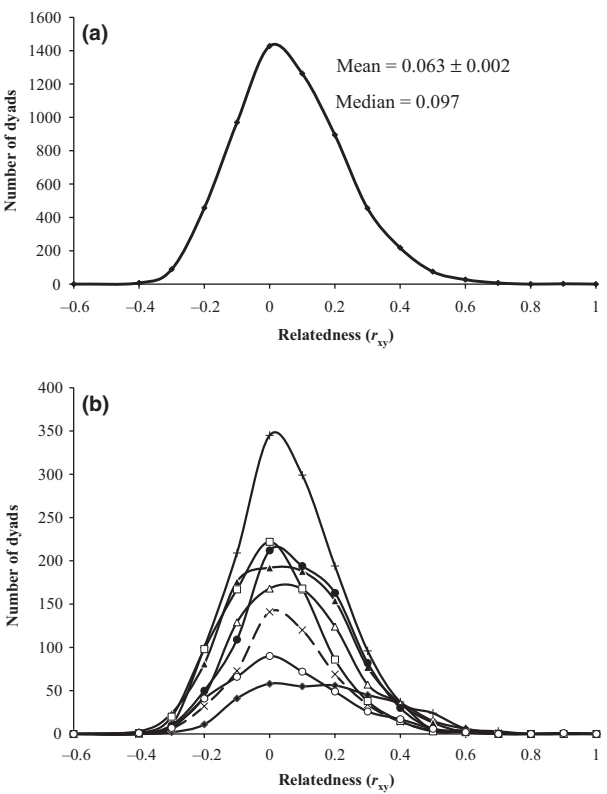

Figure 2 Observed distributions of relatedness for (a) all individuals and (b) pairs of each of the following populations: Kituuha (♦); Kasiisi (○); Nasasira (▲); Rwokusooka (+); Kaibanda (△); Tayebwa (•); Mwesigye (×); Nshaara (□).

0.014 ± 0.005 in Nshaara and 0.149 ± 0.010 in Kituuha, and across all the eight populations, the mean was 0.063 ± 0.002. Excluding related dyads, the level of relatedness within herd dropped considerably, ranging from 0.010 ± 0.005 in Nshaara herd to 0.067 ± 0.004 in Tayebwa herd. Across the eight populations, the mean coefficient of relatedness excluding related dyads was 0.034 ± 0.005, well below the expected theoretical value of 0.125 for third-degree relatives. The level of relatedness between unrelated dyads did not significantly differ between Kaibanda, Kasiisi, Kituuha and Nasasira herds (Kazo County cluster) but differed significantly ($p < 0.05$) from the levels observed for Mwesigye, Nshaara, Rwokusooka and Tayebwa herds (Nyabushozi County cluster) (Table 3). Also, the mean value for Kazo was 0.081 ± 0.023, while for Nyabushozi, it was 0.056 ± 0.014.

A right skew of relatedness values was observed when all the dyads across all populations were amalgamated (Figure 2a). With exception of Nasasira and Kituuha, the plots for relatedness exhibited a distribution that tended towards normality (Figure 2b). Most dyads in the population showed a positive value of relatedness, but only Kasiisi and Rwokusooka had dyads with $r_{xy} > 0.7$ (Figure 2b). The distribution of relatedness values for the four classes of genetic relationship is presented in Figure 3. The highest level of relatedness between dyads was observed between dam and offspring ($r_{xy} = 0.8$). The plots show non-normal distribution, especially for the grand dam–grand offspring and half-sib relationships.

The accuracy of parentage assignment by farmers was high for each of the eight cattle populations (Table S2). The mean percentages (±SE) of accurately assigned dyads across all herds were 91.9% ± 5.01 for dam–offspring dyads, 85.5% ± 3.46 for sire–offspring dyads, 75.6% ± 12.3 for half-sib dyads and 60.0% ± 5.00 for grand dam–grand offspring dyads. Excluding Nshaara, governmental herd where written pedigree records were used, these percentages were 90.4% ± 5.26 for dam–offspring dyads and 88.6% ± 2.02 for sire–offspring dyads. By farm, the efficiency of kinship assignment ranged between 58.8 and 100% for dam–offspring dyads (Table S2). In three of the eight herds, there was maximum accuracy (100%), while the rest scored over 85%, except Mwesigye who was a clear outlier at 58.8%. For sire–offspring dyads, all memory-based cases had an accuracy of 84% or above, while Nshara ranch which uses written records only had three reported sire–offspring dyads one of which wrongly assigned (Table 1). The number of farmer-assigned half-sib dyads was small but with a mean accuracy of 75.6%. With exception of the herds of Mwesigye and Kasiisi, all the farmers assigned accurately over 80% of all the four relationships dyads. In Mwesigye and Kasiisi, approximately 69% of all possible dyads were correctly assigned. First-degree relatives were easier to assign compared with second-degree relatives (90% success rate compared with 69%). While all the third-degree relationship dyads reported (n = 7) were correctly assigned, this category included only 2.4% of the total number of dyads reported.

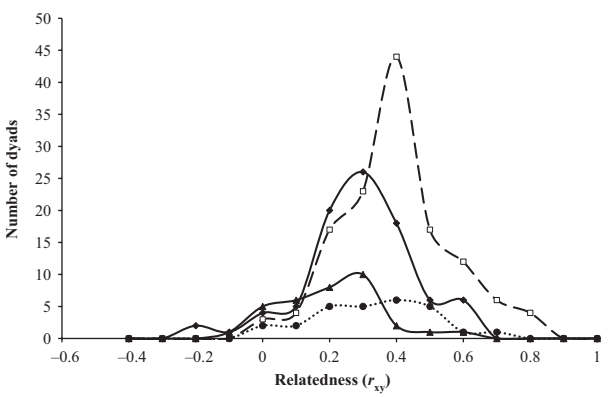

Figure 3 Observed distributions of relatedness for dyads possessing each of the following relationship categories: dam–offspring (□); sire–offspring (♦); grand dam–grand offspring (▲); half-sib (•).

Forty-nine offspring of 184, which had been sampled for relationships other than paternity, and for which the farmers had not assigned paternity, were successfully matched with the resident sires in their respective herds. In Tayebwa population for instance, there were three mature breeding bulls, and analysis assigned eleven, three and one offspring to them.

Herd size, represented by the number of dams, ranged from 80 dams (Kasiisi) to 120 dams (Nasasira and Kaibanda) excluding Nshaara herd, which included 250 dams. The number of years spent by herders in managing their respective herd ranged from 20 to 58. Pearson's correlation analysis showed a non-significant relationship between level of error in dyads assignment and the following: herd size ($r = -0.002$, $p = 0.997$, $n = 8$), length of time spent by the herder with the cattle population ($r = 0.306$, $p = 0.504$, $n = 8$) and the number of dyads assigned ($r = 0.017$, $p = 0.966$, $n = 8$). However, a strong correlation was found between herd size and number of dyads assigned ($r = 0.967$, $p < 0.001$, $n = 8$).

All the farmer-assigned dyads that were validated by genetic analysis were then used to reconstruct

pedigrees for each of the eight cattle populations studied (Figure S1). Kaibanda, Rwokusooka and Tayebwa memorized pedigrees appear deeper (having more generations) but narrower, as compared to the rest of the pedigrees that are shallow, but wider in scope of the herd.

Discussion

In this study, we address the issue of the accuracy of parentage and kinship assignment based on the human memory, for Ankole cattle, an indigenous African cattle breed. While a few studies have documented so far the use of human memory records in place of written records to understand how pastoralist communities manage their livestock in relation to the agro-ecosystems (e.g. Waters-Bayer *et al.* 2003; Bharwada & Mahajan 2006; MacOpiyo *et al.* 2006), the evaluation of the accuracy of memory-based information for pedigree recording in cattle herds in pastoral areas in Africa had not been carried out so far to the best of our knowledge.

We assessed the accuracy of the memory of the herders through the use of microsatellite markers. We found a high number of alleles per locus and high number of heterozygote alleles, for example eleven loci (*ILSTS006*, *ILSTS028*, *ILSTS036*, *ILSTS050*, *TGLA53*, *TGLA122*, *ETH225*, *TGLA227*, *AGLA293*, *BM2113* and *MGTG4B*) had ten and more alleles each and with mostly heterozygote animals. Moreover, the level of relatedness observed within the herds, considering only supposedly unrelated cows, was very low (0.034), approximately one-quarter of what would be expected for third-degree relatives. Therefore, we can be very confident that any wrongly assigned dyads were correctly identified. From our analysis, a smaller panel of five markers that amplify well, have high parent exclusion power and high heterozygosity, and hence, useful for kinship analysis is made up of *ILSTS006*, *ILSTS036*, *TGLA53*, *ETH225* and *AGLA293*.

The relative numbers of dyads in the respective kinship categories were within expected proportions for a normal breeding population where dam–offspring dyads are expected to be the most frequent, while grand dam–grand offspring dyads are expected to be the least frequent. Errors in kinship assignment were generally low when a first-degree relationship was involved, compared with second-degree ones (Tables 1 and S2). Logically, we would expect a weakening in human memory with time and more particularly with the number of cattle breeding generations. Nevertheless, remarkably, even for a third-degree relationship, the accuracy of the memory pedigree assignment was 60%. It indicates that even after five years of generation time, the herders still remember such relationships within the herds. The criteria used by Ankole pastoralists to remember relationships between their animals have been documented (Kugonza 2008) and include names given to the animals, coat colour and pattern, body shape, horn curvature and the shape of the back, all part of the indigenous knowledge of the Bahima people.

Generally, we did observe that the relatedness genetic value was in agreement with the degree of relationships (Figure 3). They were however two cases in which dyads identified as sire–offspring by the farmer showed a level of relatedness higher than 0.7. This was likely attributed to 'line breeding', which is commonly practiced in Ankole cattle (Nakimbugwe & Muchunguzi 2003; Kugonza *et al.* 2011). Memory of particular cattle lines (matrilineal or patrilineal) may be attributed to their sociocultural importance. For instance, a line that was inherited by the cattle keeper from his/her parents or ancestors or a line that originated from a prized animal from a famous Ankole cattle breeder (Museveni 1997) may be better remembered. For sire–offspring relationship line, breeding involves selection of Ankole bulls for breeding of a particular phenotype, typically dark red coat colour (*bihogo*), even if they may be closely related to other animals in the herd. A paternity misidentification rate of 14.3% was observed among the paternal–offspring dyads analysed. Paternity misidentification in our study could be attributed to memory loss and to the presence of more than one breeding bull. Ankole cattle are usually herded in groups averaging one hundred cows and one breeding bull, and a herd is locally called *igana ly'ente* (meaning one hundred cows) (Kugonza 2008). Farmers may own one, two or more herds in close proximity (Table 1), but these are not allowed to mix even if they are on the same farm. It is nevertheless possible that some of the proximal bulls were responsible for siring progeny of the misidentified dyads in our study.

Forty-nine individuals across the populations, originally sampled for non-paternity dyad analysis, were successfully nested with sires within the respective herds. Farmers may have forgotten to assign these offspring or were highly uncertain about their paternity and therefore did not want to speculate about them. The high percentage of correctly assigned sire–offspring dyads suggests that it was the latter rather than the former. This is further supported by the results observed in the Tayebwa herds where there

were three mature breeding bulls. Microsatellite sire analysis assigned eleven offspring to one bull, four offspring to another bull and only one calf to the third bull. The herder was right in identifying three sire–offspring dyads, but only did for a fifth of the possible cases. Although all the three bulls were mature, there was dominance by one bull. The other bulls possibly contribute to successful breeding by the dominant bull through subordinate roles such as competitive courtship. Differences in age class and sperm quality may also explain the observed results.

Pedigree reconstruction enables a better understanding of the dyad relationships that are easier for Ankole cattle keepers to memorize. We observe that there are two main forms of memorizing information among farmers, namely a vertical and a horizontal plane. While Kaibanda, Rwokusooka and Tayebwa population pedigrees had dyads up to four generations apart (vertical relationship memory), other cattle owners found it easier or more important to remember as many animals in a population as possible (horizontal relationship memory). This was at the expense of remembering more of the relationships shared between them. The latter practice enhances prevention of animal losses for instance through theft, but it is of limited value to selective improvement and control of inbreeding in the population.

Interestingly, Nshaara ranch herd with written pedigree documentary information only recorded three sire–offspring dyads, one of these was wrongly assigned, and also, no half-sibs and grand dam–grand offspring dyads had been recorded. It suggests that despite memorized information, Ankole cattle herders are providing very good pedigree information comparable to and it could be argued even better than could be obtained with a written recording system. The latter depending of course on how good the written records of relationships are maintained and updated. Remarkably, all the herders questioned here achieved a high level of accuracy in dyad assignment (68% and above). It should be noted that the overall level of wrong pedigree assignment (16.1%) in our study was much lower than the 36% reported in Gir cattle (Baron *et al.* 2002) and was within range of 4 and 23% reported in dairy cattle (Geldermann *et al.* 1986; Ron *et al.* 1996), where parentage information are recorded in writings. One of the criteria in selecting herds for our study was that they are being taken care of by herders having spent at least 20 years with those herds. So, all herders were very experienced, and in such a context, the results might not be so surprising. Moreover, herder must be praised for their honesty. It is credible that herders may have deliberately declined to assign relationships when they were unsure about the pedigree relationships. This is supported by the strong correlation that we found between herd size and number of dyads assigned by the herders, while correlation analysis also showed that accuracy of relationship assignment is independent of factors such as herd size, number of dyads present and the length of time spent by a herder with a particular cattle herd.

This study shows that the capability to correctly assign kinship in a population is quite established among Ankole cattle keepers, with all herders scoring 68% or above of correctly assigned dyads. Hence, the farmers' criteria for assigning relationships between dyads are not haphazard but largely depend on indigenous knowledge shared across the Bahima herders' community using standard criteria. Ankole cattle keepers with this ability could therefore accurately use these criteria in other herds that they may not have been exposed to earlier.

We assessed only the accuracy of assignments that the farmer felt confident to tell us. These represent only a subset of all first-, second- and third-degree relative dyads present in the herds. While a breeding improvement programme will rely on the knowledge of a high number of pedigree relationships and associated relevant phenotypic data, our results clearly indicate that on the pedigree side, herder memory-based information might be accurate enough for the design of such a programme. In the context of the pastoral agro-ecosystem of Uganda where pedigree written records at the farm level or herd level may prove to be difficult to be implemented and maintained, memory recording of relationships is likely the best option today for the development and the implementation of breeding improvement programmes of Ankole cattle and their crossbreds. However, the approach may not be so reliable in the future. The production system is changing with increasing reliance on fencing and paddocking to restrict animal movements and animal owners spending much less time with their animals and also more use of labourers. The use of records and more strict identification methods might be inevitable in the future. However, in the meantime, as indigenous cattle keepers transform to record keeping, the reliance on memorized pedigrees cannot be underestimated, and efforts should focus on the use of this indigenous knowledge to conserve and enhance selective improvement in the uniquely important and gracefully long-horned Ankole cattle.

Acknowledgements

This work was funded by NORAD Institutional Support to the Ankole cow–Mubende goat project in the Department of Agricultural Production (formerly Department of Animal Science), Makerere University. Part of this work was undertaken under an International Atomic Energy Agency fellowship (UGA/03020 to D.R.K). Thanks also to Francis Mugisha and Deo Ndumu for participating in the screening and selection of the studied cattle populations and blood sample collection. The collaboration of the cattle herders is greatly acknowledged. We are grateful for the comments of the anonymous reviewers of our manuscript. International Livestock Research Institute (ILRI) research is funded by programme grants from the UK, Japan, the EU, Ireland and France and funding from donors to the Consultative Group on International Agricultural Research (CGIAR).

References

Barbará A., Palma-Silva C., Paggi G.M., Bered F., Fay M.F., Lexer C. (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Mol. Ecol.*, **17**, 3759–3767.

Baron E.E., Martinez M.L., Verneque R.S., Coutinho L.L. (2002) Parentage testing and effect of misidentification on the estimation of breeding value in Gir cattle. *Genet. Mol. Biol.*, **25**, 389–394.

Bharwada C., Mahajan V. (2006) ''Gujarat: Quiet Transfer of Commons''. *Econ. Polit. Wkly*, **41**, 313–315.

Blouin M.S. (2003) DNA-Based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends Ecol. Evol.*, **18**, 503–511.

Blouin M.S., Parsons M., Lacaille V., Lotz S. (1996) Use of microsatellite loci to classify individuals by relatedness. *Mol. Ecol.*, **5**, 393–401.

Dodds K.G., Tate M.L., Sise J.A. (2005) Genetic evaluation using parentage information from genetic markers. *J. Anim. Sci.*, **83**, 2271–2279.

Falconer D.S. (1989) Introduction to Quantitative Genetics, 3rd edn. Longmans Green/John Wiley & Sons, Harlow, Essex, UK/New York.

FAO (2002) Secondary Guidelines for development of national farm animal genetic resources management plans. Measurement of domestic animal diversity (MoDAD): Recommended microsatellite markers. FAO/UNEP/iDAD, Rome, Italy. pp. 24.

FAO (2003) Understanding the indigenous knowledge and information systems of pastoralists in Eritrea. Communication for Development Case Study **26**, 1–40.

FAO (2004) An investigation into the dynamics of local knowledge among Maasai communities, with a particular focus on preferences and criteria for livestock breeding and selection. LinKS-FAO, Biodiversity-Gender-Knowledge, Phase I Final Report. FAO, Rome, Italy. **pp**. 50.

Garbe J.R., Da Y. (2005) Pedigraph^{TM}. User manual Version 2.2. University of Minnesota, Minneapolis, MN, pp. 23.

Garcia D., Carleos C., Parra D., Cañón J. (2002) Sib-parentage testing using molecular markers when parents are unknown. *Anim. Genet.*, **33**, 364–371.

Geldermann H., Pieper U., Weber W.E. (1986) Effect of misidentification on the estimation of breeding value and heritability in cattle. *J. Anim. Sci.*, **63**, 1759–1768.

Gomez-Raya L., Priest K., Rauw W.M., Okomo-Adhiambo M., Thain D., Bruce B., Rink A., Torell R., Grellman L., Narayanan R., Beattie C.W. (2007) The value of DNA paternity identification in beef cattle: examples from Nevada's free-range ranches. *J. Anim. Sci.*, **86**, 17–24.

Goodnight K.F., Queller D.C. (1999) Computer software for performing likelihood tests of pedigree relationship using genetic markers. *Mol. Ecol.*, **8**, 1231–1234.

Heyen D.W., Beever J.E., Da Y., Evert R.E., Green C., Bates S.R.E., Ziegle J.S., Lewin H.A. (1997) Exclusion probabilities of 22 bovine microsatellite markers in fluorescent multiplexes for semi-automated parentage testing. *Anim. Genet.*, **28**, 21–27.

Infield M. (2002) The Culture of Conservation: Exclusive Landscapes, Beautiful Cows and Conflict Over Lake Mburo National Park, Uganda. PhD thesis. Sch. Dev. Stud., University of East Anglia, Norwich, UK.

Jones A.G., Arden W.R. (2003) Methods of parentage analysis in natural populations. *Mol. Ecol.*, **12**, 2511–2523.

Kalinowski S.T., Taper M.L., Marshall T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.*, **16**, 1099–1106.

Krakauer A.H. (2005) Kin selection and cooperative courtship in wild turkeys. *Nature*, **434**, 69–72.

Kugonza D.R. (2008) Genetic Diversity, Husbandry, Selection Criteria and Verification of Kinship Assignment of Ankole Cattle Populations in Uganda. PhD Thesis, Makerere University, Kampala, Uganda. pp. 222.

Kugonza D.R., Nabasirye M., Mpairwe D., Hanotte O., Okeyo A.M. (2011) Productivity and morphology of Ankole cattle in three livestock production systems in Uganda. *Anim. Genet. Res.*, **48**, 13–22.

Luis C., Cothran G.E., Oom M.M. (2002) Microsatellites in Portuguese autochthonous horse breeds: usefulness for parentage testing. *Genet. Mol. Biol.*, **25**, 131–134.

MacOpiyo L., Stuth J., Wu J. (2006) Use of GIS and Agent-Based Modeling to Simulate Pastoralist Mobility in the Rangelands of East Africa. Proceedings of the GSDI-9 Conference. Santiago (Chile), 6–10 November 2006. pp. 13.

Museveni Y.K. (1997) Sowing the Mustard Seed: The Struggle for Freedom and Democracy in Uganda. Macmillan, London.

Nakimbugwe H., Muchunguzi C. (2003) Bahima, Uganda. In: I. Köhler-Rollefson, J. Wanyama (eds), The Karen Commitment. Proceedings of a Conference on Indigenous Livestock Breeding Communities, German NGO Forum on Environment and Development. pp. 34–40.

Putnová L., Knoll A., Dvořák V., Dvořák J. (2003) A novel porcine microsatellite panel for the identification of individuals and parentage control in the Czech Republic. *Czech J. Anim. Sci.*, **48**, 307–314.

Ron M., Blanc Y., Band M., Ezra E., Weller J.I. (1996) Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. *J. Dairy Sci.*, **79**, 676–681.

Sherman G.B., Kachman S.D., Hungerford L.L., Rupp G.P., Fox C.P., Brown M.D., Feuz B.M., Holm T.R. (2004) Impact of candidate sire number and sire relatedness on DNA polymorphism based measures of exclusion probability and probability of unambiguous parentage. *Anim. Genet.*, **35**, 220–226.

Tozaki T., Kakoi H., Mashima S., Hirota K., Hasegawa T., Ishida N., Miura N., Choi-Miura N., Tomita M. (2001) Population study and paternity testing for thoroughbred horses by 15 microsatellite loci. *J. Vet. Med. Sci.*, **63**, 1191–1197.

Usha A.P., Simpson S.P., Williams J.L. (1995) Probability of random sire exclusion using microsatellite markers for parentage verification. *Anim. Genet.*, **26**, 155–161.

Van Eenennaam A.L., Weaber R.L., Drake D.J., Penedo M.C.T., Quaas R.L., Garrick D.J., Pollak E.J. (2007) DNA-based paternity analysis and genetic evaluation in a large, commercial cattle ranch setting. *J. Anim. Sci.*, **85**, 3159–3169.

Van Horn R.C., Altmann J., Alberts S.C. (2008) Can't get there from here: inferring kinship from pairwise genetic relatedness. *Anim. Behav.*, **75**, 1173–1180.

Waters-Bayer A., Bayer W., Critchley W. (2003) Anything new in the rangelands? Participatory approaches to innovations by pastoralists. In: Proceedings of the 7th International Rangelands Congress. Durban (South Africa), 26 July–1 August 2003.

Supporting Information

Additional Supporting information may be found in a online version of this article:

Figure S1 Reconstructed pedigrees for eight cattle populations.

Table S1 Pairwise geographical distance (km) and proportion of alleles shared between the eight Ankole cattle populations.

Table S2 Accuracy of parentage (%) and kinship assignment of Ankole cattle dyads by herdsmen.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any Supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding for the article.