ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi:10.1111/j.1365-2052.2011.02245.x

Genetic diversity and admixture of indigenous cattle from North Ethiopia: implications of historical introgressions in the gateway region to Africa

M. Zerabruk*,[1], M.-H. Li[†,‡,1], J. Kantanen[†], I. Olsaker[§], E. M. Ibeagha-Awemu[¶], G. Erhardt and O. Vangen***

*Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Aas, Norway. [†]Biotechnology and Food Research, MTT Agrifood Research Finland, FI-31600, Jokioinen, Finland. [‡]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beichen West Road No. 1-5, Beijing 100101, China. [§]Department of Basic Sciences and Aquatic Medicine, Norwegian School of Veterinary Science, P.O. Box 8146 Dep, 0033 Oslo, Norway. [¶]Dairy and Swine Research & Development Centre, Agriculture and Agri-Food Canada, 2000 College Street, Sherbrooke, QC J1M 0C8, Canada. **Department of Animal Breeding and Genetics, Justus Liebig University, 35390 Giessen, Germany

Summary

Microsatellite variation was surveyed to determine the genetic diversity, population structure and admixture of seven North Ethiopian cattle breeds by combining multiple microsatellite data sets of Indian and West African zebu, and European, African and Near-Eastern taurine in genetic analyses. Based on allelic distribution, we identified four diagnostic alleles (*HEL1*-123 bp, *CSSM66*-201 bp, *BM2113*-150 bp and *ILSTS6*-285 bp) specific to the Near-Eastern taurine. Results of genetic relationship and population structure analyses confirmed the previously established marked genetic distinction between taurine and zebu, and indicated further divergence among the bio-geographical groupings of breeds such as North Ethiopian, Indian and West African zebu, and African, European and Near-Eastern taurine. Using the diagnostic alleles for bio-geographical groupings and a Bayesian method for population structure inference, we estimated the genetic influences of major historical introgressions in North Ethiopian cattle. The breeds have been heavily (>90%) influenced by zebu, followed by African, European and the Near-Eastern taurine. Overall, North Ethiopian cattle show a high level of within-population genetic variation (e.g. observed heterozygosity = 0.659–0.687), which is in the upper range of that reported for domestic cattle and indicates their potential for future breeding applications, even in a global context. Rather low but significant population differentiation (F_{ST} = 1.1%, $P < 0.05$) was recorded as a result of multiple introgression events and strong genetic exchanges among the North Ethiopian breeds.

Keywords admixture, *Bos indicus*, *Bos taurus*, Ethiopia, genetic diversity, microsatellite.

Introduction

The Ethiopian highlands and adjacent areas in the Horn of Africa have the highest concentration of domesticated cattle (*Bos indicus* and *Bos taurus*) in the continent (Rege 1999). Due to its geographical proximity to the entry points of Arabian and Indian zebu, and of European and the Near-Eastern taurine (e.g. Loftus *et al.* 1994, 1999; MacHugh *et al.* 1997; Hanotte *et al.* 2002), Ethiopia is believed to have been a gateway for cattle immigrations into Africa. Modern Ethiopian breeds were also introgressed with West African taurine during the spread of pastoralism (e.g. Hanotte *et al.* 2002). For these reasons, as well as because of the diverse agro-ecology and farming systems in the country, Ethiopia is a centre of secondary diversification for livestock in the continent (e.g. IBCR, 2001; Li *et al.* 2007b). For example, Ethiopia is the origin of the Sanga cattle, one of the major bovine groups in Africa (Albero & Haile-Mariam 1982; Rege 1999). Historically, Sanga cattle resulted from

Address for correspondence

O. Vangen, Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Aas, Norway.
E-mail: odd.vangen@umb.no

[1]These authors contributed equally to this study and should be considered co-first authors.

Accepted for publication 11 April 2011

cross-breeding of Arabian zebus with local long-horned taurines (Rege 1999). A second introduction of Indian zebu led to the emergence of Zenga (Zebu × Sanga) breeds, and their various strains adapted to the diverse ecological environments of the East African highlands (e.g. Li *et al.* 2007b).

Indigenous breeds constitute the bulk of the cattle genetic resources in Ethiopia, and they belong to four of the 10 major groups in Africa: Sanga, Zenga, Small East African Zebu and Large East African Zebu (Rege & Tawah 1999). African cattle, and particularly the zebu breeds, are characterized by high allelic diversity, and high levels of heterozygosity and genetic diversity compared with some European and Indian breeds (e.g. MacHugh *et al.* 1997). Due to lack of resources and political instability during most of the last century in Ethiopia, there is very little existing genetic knowledge about the indigenous cattle resources, besides the molecular characterization based on paternal lineages (Li *et al.* 2007b). However, there is a growing interest in this subject among local and international researchers involved in the study of palaeontology, phylogeography and the genetic origin of African cattle breeds (e.g. Albero & Haile-Mariam 1982; Alemseged & Geraads 2000). A more objective and reliable assessment and characterization of indigenous Ethiopian breeds at the molecular level will shed light on the current status of within-population genetic diversity and the genetic differentiation among breeds. In addition, molecular characterization will contribute to the understanding of the history and origins of breeds and to the establishment of strategies for their sustainable utilization in the future. Moreover, although African cattle (*B. indicus* and *B. taurus*) pastoralism has been clarified at the continent level (Hanotte *et al.* 2002), our study will help to understand the origins and migrations of African cattle at a fine scale, close to the gateway of zebu and Near-Eastern and European taurine expansions in the continent.

The main aim of the present study was to clarify the genetic admixture and genetic relationships of Ethiopian cattle relative to Near-Eastern, European and African taurine, and Indian and West African zebu, by combining multiple microsatellite data sets. The study also attempted to characterize intra-breed genetic diversity, gene flow and the genetic structure and differentiation of the seven north Ethiopian breeds, to provide basic information for designing conservation and genetic improvement programmes. In this study, the results were compared with previous results based on paternal lineage (Y-chromosomal microsatellite) analysis and are discussed in the larger geographic context of African cattle (*B. indicus* and *B. taurus*) pastoralism.

Materials and methods

Cattle breeds

A total of 315 animals, representing seven breeds belonging to three major cattle groups in Ethiopia, were included in this study: Sanga [Afar (ETAF) and Raya (ETRA)], Zenga [Abergelle (ETAB), Arado (ETAR), Irob (ETIR) and Fogera (ETFO)] and Large East African Zebu [Begait (ETBA)]. Sample size for each breed is detailed in Table 1. The sampling locations are provided by Li *et al.* (2007b). Hair samples were plucked from the tails of unrelated animals (according to farmers' information, as pedigree records were unavailable), placed in paper envelopes and stored at room temperature until further processing. Ten to 15 hair roots per animal were used to extract DNA using the Qiagen DNeasy 96 Tissue kit (Qiagen) according to the manufacturer's protocol.

The seven North Ethiopian breeds were reported to have adaptive advantages in their respective production environment, such as resistance of the Abergelle to tick infestation, adaptation of the Afar and Begait to arid and semi-arid environments, suitability (draught power and temperament) of Raya bulls for ploughing, the ability of the Arado to cope with seasonal feed and water shortages, and the adaptation of the Irob to graze succulents on extremely mountainous terrain. Detailed information on the

Table 1 Breed group, sample size (n), mean observed (H_O) and expected (H_E) heterozygosity, mean number of observed alleles (A_O), sample size corrected mean allelic richness (A_R) per breed based on minimum sample size of 17 diploid individuals, number of private alleles (A_P), inbreeding coefficient (F_{IS}), significance for Hardy–Weinberg equilibrium (HWE) test and the number of pairwise loci deviating from linkage equilibrium (LE) for the seven north Ethiopian cattle breeds.

Cattle breeds	Breed group[1]	n	H_E	H_O	A_O	A_R	A_P	F_{IS}[2]	HWE[3]	LE
Afar	Sanga	42	0.720	0.680	7.1	6.250	1	0.050**	n.s.	12
Raya	Sanga	41	0.709	0.674	6.7	6.073	5	0.041*	n.s.	8
Arado	Zenga	49	0.718	0.686	7.5	6.254	8	0.045**	n.s.	10
Abergelle	Zenga	47	0.709	0.672	7.2	6.272	4	0.054**	**	7
Fogera	Zenga	42	0.721	0.673	6.6	5.974	3	0.052*	**	5
Irob	Zenga	47	0.713	0.687	7.1	6.040	4	0.036*	n.s.	5
Begait	Large E.A. zebu	47	0.670	0.659	6.3	5.677	1	0.025	n.s.	4
Overall.	–	315	0.708	0.676	9.5	6.230	26	0.054	n.s.	51

[1]E.A., East Africa.

[2]*Significant at $P < 0.05$ and **Significant at $P < 0.01$ after sequential Bonferroni corrections (Rice 1989).

[3]n.s., not significant.

characteristics of these breeds has been previously reported by Rege & Tawah (1999), Zerabruk & Vangen (2005) and Zerabruk *et al.* 2007.

Data on additional breeds (see Table S1) representing the Indian [Nellore (INNEL), $n = 27$] and African Zebu [Red Bororo (REBO, Nigeria), $n = 47$ and Sokoto Gudali (SOGU, Nigeria), $n = 62$], and European [Finnish Holstein-Friesian (FIHFR), $n = 43$; German Simmental (DESIM), $n = 50$], African [N'Dama (AFNDFN), $n = 44$; Muturu (MUTURU), $n = 18$], and the Near-Eastern taurine [Turkish Grey Steppe (TUTGY), $n = 42$, East Anatolian Red (TUEAR), $n = 43$ and South Anatolian Red (TUSAR), $n = 43$] were also included in admixture and comparative analyses.

Microsatellite markers and genotyping

Twenty microsatellites (*BM1818*, *BM1824*, *BM2113*, *CSSM66*, *ETH3*, *ETH10*, *ETH152*, *ETH225*, *HEL1*, *HEL5*, *HEL9*, *HEL13*, *ILSTS005*, *ILSTS006*, *INRA005*, *INRA023*, *INRA032*, *INRA035*, *INRA037* and *INRA063*) out of the 30 markers recommended by the International Society for Animal Genetics and the Food and Agriculture Organization of the United Nations (FAO) for cattle genetic diversity studies were used for genotyping. Primer sequences, detailed information of PCR protocols and references for the microsatellites can be found on the CadBase website (http://www.projects.roslin.ac.uk/cdiv/markers.html). The markers were amplified using one unlabelled primer combined with either a primer labelled with ^{32}P or a primer labelled with fluorescent dye. The allele sizes were determined manually on autoradiograms with a sequence ladder or scored according to the TAMRA 500 size standard on an ABI PRISM 377 sequencer (Applied Biosystems) respectively. Two Norwegian and one international reference animal were included in all gel-runs, allowing adjustment of allele sizes to those agreeing with international reference samples.

Microsatellite data for the Near-Eastern breeds (Turkish Grey Steppe, East Anatolian Red, South Anatolian Red) were from the EU ResGen Project, while data for other additional cattle breeds were previously reported by Loftus *et al.* (1999), Ibeagha-Awemu *et al.* (2004) and Li *et al.* (2005, 2007a). Genotypes for the additional breeds have 12–19 loci in common with the markers genotyped in North Ethiopian cattle and 11 loci are shared among all the breeds included in this study. All breed information is detailed in Table S1. At least one international reference sample genotyped in all the additional breeds was used to standardize the allele sizes.

Data analysis

Allele frequencies, number of observed alleles (A_O), observed (H_O) and expected (H_E) heterozygosity as a measure of population genetic diversity, tests for deviations from genotypic linkage equilibrium (LE) for each locus pair within population and for each pair of loci across populations, and tests for deviation from Hardy–Weinberg equilibrium (HWE) were analysed as implemented in GENEPOP version 3.4 (Raymond & Rousset 1995). Results of the tests were corrected for multiple comparisons by applying sequential Bonferroni corrections (Rice 1989). The presence of non-amplifying null alleles was tested using MICRO-CHECKER version 2.2.3 (Oosterhout *et al.* 2004). Allelic richness (A_R) and Weir & Cockerham (1984) fixation indexes (F_{IT}, F_{ST} and F_{IS}) were calculated using FSTAT version 2.9.3.2 (Goudet 2002). Markov chain analyses were applied to estimate significance (10 000 dememorization steps, 500 batches and 5000 iterations per batch).

Nei's D_A genetic distance (Nei *et al.* 1983), which is suitable for reconstruction of phylogeny irrespective of the mutation model and is efficient in obtaining the correct topology (e.g. Takezaki & Nei 1996), was computed with DISPAN (Ota 1993). The individual distances based on the proportion of shared alleles $D_{PS} = -ln(PS)$ (Bowcock *et al.* 1994) were calculated using the MICROSAT version 1.4 computer program (Minch *et al.* 1995). A D_A-based topological tree was constructed using the neighbour-joining (NJ) algorithm (Saitou & Nei 1987) as implemented in PHYLIP version 3.6 (Felsenstein 1993) with 1000 bootstrapping values. The 10 additional cattle breeds were also included in the genetic relationship analysis on the basis of 11 common loci. The pattern of population differentiation among North Ethiopian breeds was further described by a factorial correspondence analysis of the individual multilocus scores, computed using GENETIX version 4.05 (Belkhir *et al.* 2004). The factorial correspondence analysis deals with data collected on more than one variable and can condense the information from a large number of alleles and loci into fewer synthetic variables. In this approach, allele frequencies of the breeds at all loci were used as variables, and the population clusters were identified graphically. The first three major components were plotted on a three-dimensional scatter diagram for the seven breeds.

We first made an estimation of introgressions of the different groups of breeds, such as Indian zebu and European, Near-Eastern and African taurine, using the grouping-diagnostic/specific-allele method as described in MacHugh *et al.* (1997) and Loftus *et al.* (1999). In summary, the frequencies of such alleles, or groups of alleles at a particular locus (see MacHugh *et al.* 1997; Ibeagha-Awemu *et al.* 2004; Li *et al.* 2007a) were averaged to give an estimate of proportion of introgression from the cattle groupings for each of the North Ethiopian breeds.

A Bayesian clustering method was then employed to assess population structure using the program STRUCTURE version 2.2 (Pritchard *et al.* 2000). This method uses multilocus genotypes to infer the fraction of population/individual genetic ancestry that belongs to a cluster, for a given number of clusters (K). The program assumes HWE in

the ancestral populations and complete LE between markers used, and the criteria for grouping individuals are to minimize the HWE and the gametic phase disequilibrium between loci within groups. We performed 10 runs for each K value at 1–17 and ran the program assuming a model of admixture and correlated allele frequencies. We did not use any prior information about the population origin of the animals. A burn-in period of 200 000 generations and Markov chain Monte Carlo (MCMC) simulations of 500 000 iterations were used in all the above runs. The values of LnP(D) (the log probability of data) were estimated assigning priors from 2 to 17, and the optimal K was chosen based on the delta K (ΔK) value. This criterion was originally described in Evanno *et al.* (2005) and was shown to be effective in later studies (Medugorac *et al.* 2009; Li & Kantanen 2010). We then evaluated the average membership coefficients of predefined (sampled) cattle populations for the K inferred clusters. Genotypes of all the populations at the 11 common loci (see Table S1) were included in this analysis.

The likelihoods calculated by STRUCTURE are often interpreted as indicating admixture ratios, but can be confounded by the fixation of alleles from one of the parental populations (see Freeman *et al.* 2004). Therefore, a recently developed implementation of a MCMC method, likelihood-based estimation of admixture (LEA) (Chikhi *et al.* 2001), was employed in estimating the relative contributions of zebu and taurine in each of the North Ethiopian breeds. This method estimates admixture proportions at the population level, taking into account drift since the admixture event, variation due to sampling and uncertainty in the estimation of the ancestral allele frequencies (e.g. Chikhi *et al.* 2001; Li *et al.* 2005). It estimates admixture proportions P_1 and $P_2 = 1 - P_1$ from parental (source) populations 1 and 2, respectively. Each analysis was based on 100 000 MCMC steps, and the pdf (probability density function) values of parental population 1 in parallel with 95% probability intervals of its posterior distribution obtained using the LOCIFIT package for *R*. As there are multiple introgression events in North Ethiopian cattle, we pooled the taurine (Finnish Holstein-Friesian, German Simmental, N'Dama, Muturu, Turkish Grey Steppe, East Anatolian Red and South Anatolian Red) to represent parental population 1 and the zebu (Nellore, REBO and Sokoto Gudali) to represent parental population 2.

Results

Allelic distribution and locus variability

A total of 198 alleles were detected. The observed number of alleles at a single locus across populations ranged from 6 (*INRA063* and *INRA005*) to 16 (*INRA032*), and the overall mean number of alleles per locus was 9.9 (Table S2). The statistics on the basis of 11 common loci detected only 3 (2.26%, 3/133) breed-specific (private) alleles in North Ethiopian cattle. A total of 16 (12.03%, 16/133) alleles were shared between the taurine and at least one North Ethiopian breed, but not with the zebu. In contrast, 2 (1.50%, 2/133) alleles were present in the zebu and North Ethiopian cattle, but not in the taurine breeds (data not shown). A further examination of the previously identified diagnostic/specific alleles for bio-geographical groupings (MacHugh *et al.* 1997; Ibeagha-Awemu *et al.* 2004; Li *et al.* 2007a; see also Table S3) in the present sample suggests definite introgression of the three groups into North Ethiopian cattle, with a smaller proportion of European taurine alleles. In addition, four loci (*HEL1*, *CSSM66*, *BM2113* and *ILSTS6*) displayed alleles that were present at higher frequency in the Near-Eastern taurine, were absent in the European taurine and the Indian zebu and were either absent or at low frequencies or in the African groupings (Fig. S1). Following the definition of MacHugh *et al.* (1997), the alleles (*HEL1*-123 bp, *CSSM66*-201 bp, *BM2113*-150 bp and *ILSTS6*-285 bp) were thus considered diagnostic or specific for the Near-Eastern *B. taurus*. These Near-Eastern taurine diagnostic alleles are further confirmed by the allelic frequency comparison with that in more than 40 breeds from Europe and Asia (Li & Kantanen 2010). They were used to evaluate gene flow and genetic admixture in the North Ethiopian cattle as described below.

Within-population diversity and tests for HWE and LE

Genetic diversity estimates within populations across all 20 loci showed that the variability was comparable among them: 0.670–0.721 for H_E, 0.659–0.687 for H_O, 6.300–7.500 for A_O and 5.677–6.272 for A_R (Table 1). The Arado had the highest number of private alleles ($A_P = 8$), while the lowest number of private alleles was detected in the Afar and Begait ($A_P = 1$). The Begait showed the lowest degree of within-breed genetic variation (Table 1). In terms of basic statistics (e.g. A_O, H_E and H_O) for within-population genetic diversity, on average North Ethiopian cattle had values lower than those for the Near-Eastern taurine and West African zebu, but higher than those for Indian zebu and European and African taurines (Table S4).

Of the 140 locus/population combinations (20 loci × 7 North Ethiopian cattle breeds), 19 (13.6%, 19/140) showed significant ($P < 0.05$) deviation from HWE (data not shown). Abergelle and Fogera showed significant deviation from HWE after the Bonferroni corrections (Table 1). Nevertheless, no significant deviation from HWE was detected when the test was applied to all loci across populations. Significant (Fisher's method: approximate value of $\chi^2 > 24$, d.f. = 14, $P < 0.05$) deviations from genotypic LE within each population were found in 51 (3.8%) of 1330 [$pn(n-1)/2$, for n loci and p populations; $p = 7$, $n = 20$] comparisons and in 8 (4.2%) of 190 for the tests across populations (data not shown; both deviations <5%, which is

significant). Thus, the two loci (*HEL5* and *ILSTS006*) showing significant evidence of deviation from HWE were not excluded in the STRUCTURE analysis.

Genetic differentiation and genetic distances

Low but statistically significant ($P < 0.05$) global genetic differentiation among North Ethiopian breeds ($F_{ST} = 0.011 \pm 0.002$; 95% CI, 0.006–0.015) was recorded. A pairwise population differentiation test carried out after sequential Bonferroni corrections revealed that there was significant differentiation ($P < 0.05$) in six of the 21 pairwise multiple comparisons among North Ethiopian breeds (Table S5). On average, the Begait had the largest genetic distances and were most differentiated from the other breeds. However, the NJ tree constructed from the D_A distance matrix revealed a slight genetic division between the group comprising the Irob, Arado and Begiat and the remaining group (Fig. S2). Genetic differentiation among the breeds was further explored using three-dimensional factorial correspondence analysis (FCA) at the individual level, based on allele frequencies of 20 microsatellites (Fig. S3). Close genetic relationships among the populations are evident. Nevertheless, the first factor, which accounted for approximately 29% of the total variation, coincided with the levels of zebu introgression in the populations. It roughly separated the populations with relatively greater zebu influence, such as the Begait, Irob and Arado, from the remaining populations.

Pairwise F_{ST} estimates (above the diagonal) and the D_{PS} genetic distances (below the diagonal) for the 17 breeds were estimated based on 11 common microsatellites (Table S6). Typically, low F_{ST} and D_{PS} estimates were recorded among North Ethiopian breeds, followed by those between North Ethiopian and West African zebu breeds, while larger estimates were generated between the zebu (including North Ethiopian cattle, and West African and Indian zebu) and taurine (including African, European, Near-Eastern taurine) pairs. Compared with the Near-Eastern taurine, European and African taurine showed greater genetic differentiation than from the zebu.

Figure 1 is a topological tree for all 17 breeds based on Nei's D_A distances. It illustrates the large divergence between *B. taurus* and *B. indicus* cattle, which is consistent with earlier results from mtDNA (Loftus *et al.* 1994) and autosomal (MacHugh *et al.* 1997; Loftus *et al.* 1999) analysis of European, African and Indian cattle, as well as from Y-chromosomal microsatellite characterization of European and African cattle (Li *et al.* 2007b). As reported by Loftus *et al.* (1999), breeds from the Near-East (TUEAR, TUSAR and TUTGY) were positioned proximal to the *B. taurus* radiation, with high bootstrap values (Fig. 1). The tree also exhibits a compact cluster of European taurine (FIHFR and DESIM), African taurine (AFNDFN and MUTURU) and West African zebu breeds (SOGU and REBO), respectively.

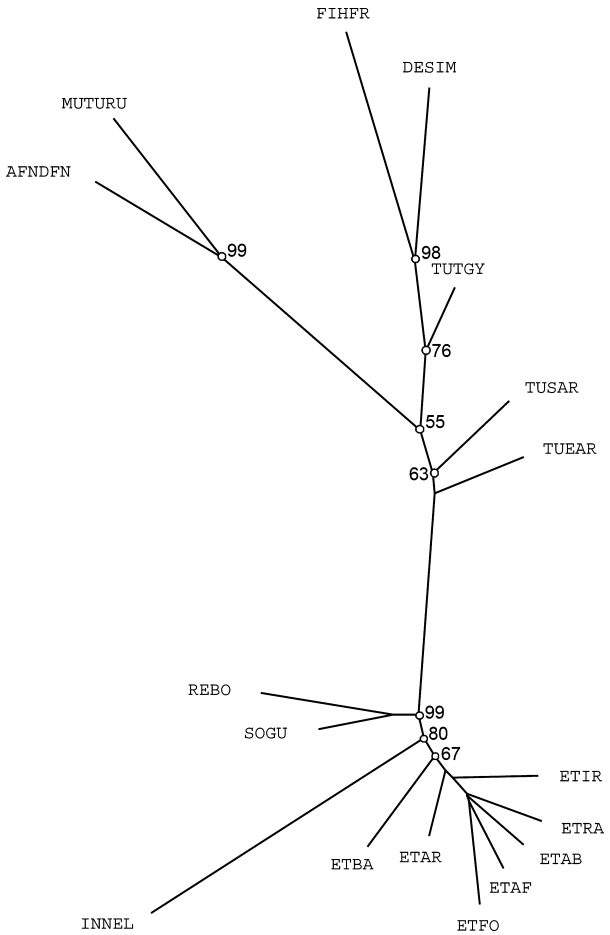

Figure 1 Genetic relationships among all the 17 cattle breeds constructed using NJ (Saitou & Nei 1987) with 1000 bootstraps based on 11 common microsatellites.

However, there is no compact clustering of the North Ethiopian (ETBA, ETAR, ETFO, ETAF, ETAB, ETRA and ETIR) or Near-Eastern breeds. The seven North Ethiopian cattle in general have shorter branches with lower bootstrap values, possibly owing to the multiple introgression events and strong gene flow among them.

Genetic admixture

It was possible to estimate the proportions of admixture in North Ethiopian cattle from different bio-geographical groupings by assigning diagnostic alleles for Indian zebu, and African, European and the Near-Eastern taurine ancestry, identified previously and in this study (Table S3). The Indian zebu influence is considerable (55.16–63.78%), and in terms of magnitude is followed by African (9.20–14.02%) and European taurine (1.51–4.44%), while the least introgression was from the Near-Eastern taurine (0.37–1.88%). The Begait, Arado and Irob breeds received most Indian zebu genes, but relatively few from Near-Eastern taurine introgression.

Table 2 Admixture in North Ethiopian cattle breeds estimated by different methods, i.e. STRUCTURE program (Pritchard *et al.* 2000) at $K = 2$ and 6, and the LEA program (Chikhi *et al.* 2001), and the distribution of private alleles.

	Two main genetic groups				Multiple genetic groups									
	STRUCTURE		LEA		Private alleles				STRUCTURE					
Breed	*Bos taurus*	*Bos indicus*	*Bos taurus*	*Bos indicus*	Indian zebu	AF taurine	EU taurine	NE taurine	Indian zebu	WF zebu	AF taurine	EU taurine	NE taurine	NEt cattle
ETAB	0.095	0.905	0.109	0.991	0.588	0.140	0.021	0.019	0.077	0.288	0.046	0.038	0.028	0.523
ETAF	0.085	0.915	0.094	0.906	0.552	0.137	0.045	0.006	0.045	0.161	0.042	0.027	0.024	0.701
ETAR	0.051	0.949	0.087	0.913	0.615	0.133	0.019	0.009	0.070	0.364	0.032	0.023	0.014	0.497
ETBA	0.086	0.914	0.072	0.928	0.638	0.092	0.035	0.004	0.174	0.305	0.060	0.030	0.024	0.408
ETFO	0.103	0.897	0.112	0.888	0.588	0.137	0.015	0.013	0.111	0.307	0.056	0.043	0.024	0.458
ETIR	0.041	0.959	0.091	0.909	0.611	0.094	0.029	0.010	0.068	0.286	0.017	0.016	0.014	0.600
ETRA	0.102	0.898	0.125	0.875	0.586	0.138	0.022	0.005	0.088	0.241	0.048	0.038	0.031	0.554

LEA, likelihood-based estimation of admixture; ETAB, Abergelle; ETAF, Afar; ETAR, Arado; ETBA, Begait; ETFO, Fogera; ETIR, Irob; ETRA, Raya; AF, African; EU, European; NE, Near-Eastern; WF, West African; NEt, North Ethiopian.

Using the LEA program, the analysis produced estimates of admixture proportions, suggesting that the contemporary gene pool of North Ethiopian cattle is heavily influenced by zebu. The genetic contribution from the taurine populations (P_1) was estimated to be low, ranging from 0.72% (95% CI: 0.12–18.4%) in the Begait to 12.34 (95% CI: 1.4–27.3%) in the Raya (MCMC method, Table 2). The remainder (P_1 = 87.66–99.28%) was contributed by the zebu.

The model-based clustering analysis provided in the STRUCTURE program complemented the two aforementioned methods by allowing the genetic contribution from the two known major source populations to be estimated. Figure 2 presents the most frequent pattern of the results of admixture analysis. The mean likelihood function values for $L(K)$ were an increasing function of K for all the examined values of 1–17 (10 runs for each). However, the distribution of ΔK had two peaks, one at $K = 2$ and the other at $K = 6$, suggesting an uppermost level of hierarchy at these two K values (Fig. S4). Results of the clustering analysis strongly support previous findings and suggest that the contemporary North Ethiopian cattle are highly introgressed by the zebu (Fig. 2 and Table 2). The 10 runs at $K = 2$ all indicated a uniform division into two main genetic clusters, *B. taurus* and *B. indicus*. A high level of zebu genetic introgression was indicated in the seven breeds, although there were slight differences among the breeds. The contribution from the *B. taurus* lineage was very low and ranged from 1% in the Irob to 5% in the Abergelle (Table 2). A further analysis at $K \geq 3$ revealed the presence of a population structure within the two main clusters. At $K = 3$, most of the runs (7/10) resulted in differentiation of the African taurine (N'Dama and Muturu) and the other taurine breeds. At $K = 4$, the most frequent pattern (7/10) cluster placed the Indian and West African zebu in separate clusters, but they, in turn, split into separate clusters at $K = 5$. The subdivision at $K = 6$ identified the clusters corresponding to the six bio-geographical groupings of populations: North Ethiopian, Indian and West African zebu, African, European and the Near-Eastern taurine (Fig. 2). Table 2 summarizes the mean proportions of shared ancestry among the six

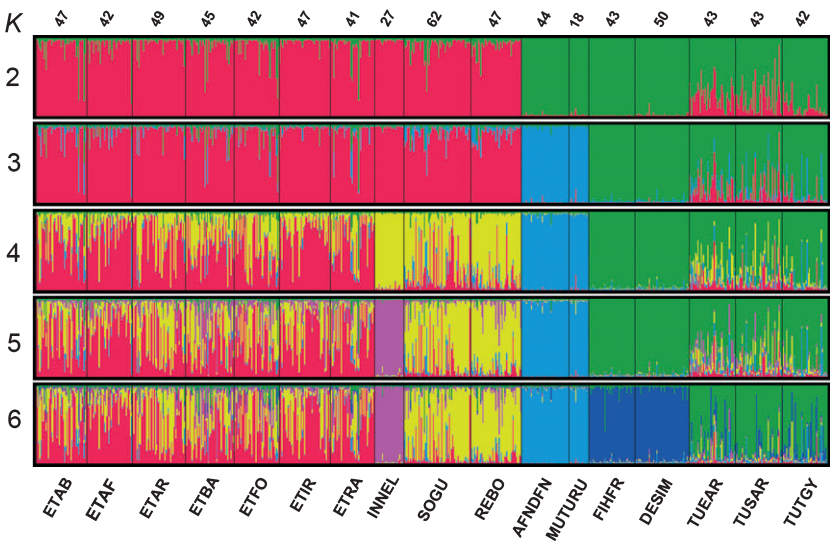

Figure 2 Population structure of 17 cattle populations using the model-based STRUCTURE program (Pritchard *et al.* 2000). Each animal is represented by a single vertical line divided into K colours, where K is the number of clusters assumed and the coloured segment shows the individual's estimated proportion of membership (averaged across 10 runs at each K = 2–6) in that cluster. Black lines separate the populations labelled above the figure. The labels above the figure indicate the number of animals analysed in each breed.

genetically differentiated bio-geographical clusters across 10 runs at $K = 6$. We found the highest proportion of West African zebu ancestry (0.161–0.364) in North Ethiopian cattle, followed by Indian zebu (0.045–0.174). All seven North Ethiopian breeds are characterized by low taurine ancestry, with consistently more African taurine (0.017–0.060), less European taurine (0.016–0.043) and least Near-Eastern taurine (0.014–0.031) ancestry.

Discussion

Genetic admixture

Our microsatellite allelic distribution indicates multiple introgressions, particularly highlighting the previously non-quantified influence of Arabian zebu and Near-Eastern taurine. Our results support the use of diagnostic alleles for Indian zebu, and African and European taurine that have been identified previously (MacHugh *et al.* 1997; Loftus *et al.* 1999; Magee *et al.* 2002; Ibeagha-Awemu *et al.* 2004). However, the Near-Eastern taurine diagnostic alleles (*HEL1*-123 bp, *CSSM66*-201 bp, *BM2113*-150 bp and *ILSTS6*-285 bp) in this study are reported for the first time. These alleles will be useful in future investigations on the immigration and admixture of cattle from the Near-Eastern centre of domestication, especially for clarifying the migration routes from the Near-East to central and north-eastern Asia, where cattle genetic resources have been less well characterized (but see Li *et al.* 2007a).

The model-based Bayesian method (Pritchard *et al.* 2000) differentiated the two Bovidae subspecies, *B. taurus* and *B. indicus*, and the six bio-geographical groups of cattle breeds, Indian, West African and East African zebu, African, European and Near-Eastern taurine, at $K = 2$ and $K = 6$, respectively (Fig. 2). The *B. indicus* admixture level of the breeds inferred in this study, using the STRUCTURE and LEA methods, is in a similar range as that earlier reported for East African and Madagascan breeds (>90%, Freeman *et al.* 2006), but is a higher level than that calculated previously for West and Central African breeds (59.8–83.2%, MacHugh *et al.* 1997; 58; 1–74.0%, Ibeagha-Awemu *et al.* 2004). This fits in well with earlier results of a gradient of zebu influence that peaked in East Africa (Hanotte *et al.* 2002). However, we noted differences between the estimates for total zebu ancestry generated by the STRUCTURE and LEA methods, and the proportions of Indian zebu by the private-allele method (Table 2) in North Ethiopian cattle. These differences may be due to two components of zebu influence in North Ethiopian cattle, one from Indian zebu and the other probably from Arabian zebu, thus reflecting the two models of zebu introgression suggested by archaeological and genetic evidence (Clutton-Brock 1989; Hanotte *et al.* 2002).

The breeds showed high levels of zebu introgression, with a cline of *B. indicus* influence from the north-western part of Ethiopia, where the Begait is dominant, to the north-eastern Afar region of Ethiopia, where the Afar is dominant. This conforms with the ongoing process of replacement of the Sanga by Zebu and the formation of intermediate breeds of Zenga, for example Arado, as a result of the last wave of Zebu introduction to Ethiopia through North Africa (Albero & Haile-Mariam 1982). Our previous study on the Y-chromosome haplotype diversity of these breeds also revealed the predominance of indicine alleles and a single taurine haplotype in these breeds (Li *et al.* 2007b). We also noted the lack of alleles shared solely between indicine and North Ethiopian cattle ($n = 2$) vs. those shared by taurine and North Ethiopian cattle ($n = 16$). However, one might expect more alleles shared between indicine and North Ethiopian cattle, as the North Ethiopian cattle are mainly indicine. This discrepancy is due to the strong Indian zebu introgression into the Near-Eastern taurine breeds (Loftus *et al.* 1999), which results in fewer alleles shared only between indicine and North Ethiopian cattle.

We also detected three sources of *B. taurus* influence on North Ethiopian cattle: West African taurine ranked the highest, European cattle were intermediate and the Near-Eastern cattle ranked lowest. The highest portion of the genetic source from West African taurine lends support to an indigenous origin for the earliest African domestic cattle (see Hanotte *et al.* 2002). The detection of non-African *B. taurus* genetic influence from Europe and the Near East is probably due to a more recent colonial legacy, suggesting that it must have been secondary to the origin of African cattle pastoralism (Hanotte *et al.* 2002). However, we noted a difference in the estimates of different introgressions in North Ethiopian cattle using the various methods, especially the considerably higher level of Indian zebu ancestry estimated using the private-allele method than was indicated by the STRUCTURE method. One caveat is that we definitely cannot establish all the diagnostic alleles for the bio-geographical groupings and, therefore, the diagnostic-allele method cannot provide an exact value for admixture proportions, instead providing approximations. Another caveat stems from the STRUCTURE program, the results of which are sensitive to the number of populations sampled, the number of individuals typed in each population and the number of loci scored (e.g. Pritchard *et al.* 2000; Evanno *et al.* 2005; Li and Kantanen 2010). Therefore, we ascribe the difference in the estimates of Indian zebu ancestry mostly to the small number of Nellore samples that represent the Indian zebu. The discrepancy between the two methods can also arise from the stochasticity of markers included in the analysis, but we assume this effect to have been minimal because this set of markers has proved useful and efficient in cattle genetic diversity studies across populations with global geographic origins (e.g. Li and Kantanen 2010). However, the estimation using the diagnostic-allele method seems more reasonable in terms of the results of Y-chromosomal microsatellite analysis (e.g. European Cattle Genetic Diversity Consortium 2006; Li *et al.* 2007b) and knowledge of cattle pastoralism in Africa (e.g. Hanotte *et al.*

2002). Inclusion of additional population samples, especially the samples of Indian zebu genotyped using additional microsatellites, may enable better estimates of membership and introgressions to be made, especially from the STRUCTURE program. Our data also suggest a relatively higher proportion of zebu ancestry in the Near-Eastern taurine than in African and European taurine (see $K = 2$ in Fig. 2). This is consistent with the conclusion regarding zebu introgression into the Near-Eastern region from the East (Loftus *et al.* 1999).

Within-population genetic diversity and genetic differentiation

There was a high level of genetic variation, in terms of basic population diversity indices, among the seven North Ethiopian breeds. The results are within the upper range of those reported for African zebu (MacHugh *et al.* 1997; Rege *et al.* 2001; Ibeagha-Awemu *et al.* 2004; Freeman *et al.* 2006) and are only exceeded by results for taurines from the Near East and for West African zebu. Increased genetic diversity in African cattle is believed to be a result of high levels of introgression of *B. indicus* and *B. taurus* (MacHugh *et al.* 1997; Freeman *et al.* 2004). The overall high level of genetic diversity within North Ethiopian cattle (Table S5) could also be because the supposed breeds might actually be cross-bred animals of no particular refinement (Fig. S3) relative to traditional breeding practices for type. Thus, a much finer resolution analysis based on dense markers is needed to accurately identify those animals that still contain portions of the genome unique to native breeds.

The level of genetic differentiation ($F_{ST} = 1.1\%$) observed among North Ethiopian cattle is comparable to previous results of Y-chromosomal lineage analysis among the same set of populations ($F_{ST} = 2.1\%$, Li *et al.* 2007b), but is much lower than the range reported for other breeds based on autosomal microsatellite analysis, such as the 3.5% for Belgian breeds (Mommens *et al.* 1999), 6% for West African cattle (Ibeagha-Awemu & Erhardt 2005), 10.7% for North European breeds (Kantanen *et al.* 2000) and 8.5% for North Eurasian breeds (Li *et al.* 2007a). In addition, some Asian cattle breeds maintained using similar management practices also showed low F_{ST} values ($F_{ST} = 0.76\%$ based on mtDNA sequence analysis; Berthouly *et al.* 2010) and a similar range of high levels of observed heterozygosity (e.g. Berthouly *et al.* 2010) to that observed in this study. The rather low genetic differentiation reported here could be due to the high level of genetic exchange or strong male-mediated gene flow (Li *et al.* 2007b) among the breeds, resulting from geographical proximity, uncontrolled breeding and recent divergence.

Genetic relationships and conservation implications

Our analyses that included all the breeds confirmed the genetic relationships among breeds form different bio-geographical groupings established previously (e.g. MacHugh *et al.* 1997; Loftus *et al.* 1999; Ibeagha-Awemu *et al.* 2004; Ibeagha-Awemu & Erhardt 2005), that is, the large genetic divergence between cattle of *B. indicus* and *B. taurus* origin and the breeds from the Near East proximal to the *B. taurus* populations. Among the zebu populations, the North Ethiopian cattle showed closer genetic relationships to the Indian zebu (i.e. Nellore) than to the West African zebu, which have geographical origins similar to the North Ethiopian cattle. This finding is in line with African cattle pastoralism, for which substantial Indian zebu immigration was identified in East Africa, and the genetic influence of *B. taurus* was found to be most important in West Africa (Hanotte *et al.* 2002).

A recent on-farm characterization of North Ethiopian cattle indicated differences in some phenotypic characteristics and adaptive advantages of some of the breeds with respect to their production environments. An increasing number of farmers preferred intermediate (cross-bred) to indigenous animals (Zerabruk & Vangen 2005). As phenotypes including tick resistance and cactus eating are important for the cattle breeds in the region, it is necessary to conserve the germplasm until a fine analysis of the cause of phenotypic diversity can be identified. From a genetic perspective, North Ethiopian cattle seem to represent a unique gene pool. Their generally hybridized status and allele combinations may well represent crucial contributions to future sustainable animal production systems. The high levels of genetic diversity in North Ethiopian cattle located at the gateway to Africa serve as a basis for future exploitation of cattle genetic resources.

In conclusion, we have made for the first time, an estimation of the genetic influence of all possible bio-geographical groupings of cattle in North Ethiopia, which represents the gateway to Africa. The cattle breeds were characterized as having a high level of zebu genetic background that have resulted from major immigrations from the Arabian region and a secondary influx from India, with less influence from West African and European taurines, and the least contribution coming from the Near-Eastern taurines. Our results show a high level of within-breed genetic diversity and a low level of population differentiation among the breeds, probably due to multiple cross-breeding events in East Africa. The unique adaptive traits of these breeds make them a conservation priority.

Acknowledgements

The present work was funded by the Mekelle University/NORAD cooperation project. We thank Ole A. Guttersrud (Norwegian School of Veterinary Science) for his assistance in the laboratory work, Dr J.A. Lenstra for the verification of comparable allele sizes on the reference breeds and our data and for his help in obtaining the data for Turkish Grey Steppe, East Anatolian Red and South Anatolian Red

populations, and Dr Miika Tapio (MTT Agrifood Research Finland, Jokioinen, Finland) for his useful discussions during the initial data analysis stage. We are grateful to Dr Dan Bradley for sharing part of the data with us.

References

Albero M. & Haile-Mariam S. (1982) The indigenous cattle of Ethiopia: part I. *World Animal Review* **41,** 2–10.

Alemseged Z. & Geraads D. (2000) A new Middle Pleistocene Fauna from the Busidima-Telalak region of the Afar, Ethiopia. *Comptes Rendus de l'*Academie des Sciences, Paris: Sciences de la Terre et des planètes, Earth and Planetary Sciences **331,** 549–56.

Belkhir K., Borsa P., Chikhi L., Raufaste N. & Bonhomme F. (2004) GENETIX 4.05, Logiciel Sous Windows TM pour la Génétique des Populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.

Berthouly C., Maillard J.C., Pham Doam L. *et al.* (2010) Revealing fine scale subpopulation structure in the Vietnamese H'mong cattle breed for conservation purposes. *BMC Genetics* **11,** 45.

Bonin A., Nicole F., Pompanon F., Miaud C. & Taberlet P. (2007) Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology* **21,** 697–708.

Bowcock A.M., Ruiz-Linares A., Tomfohrde J., Minch E., Kidd J.R. & Cavalli-Sforza L.L. (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368,** 455–7.

Chikhi L., Bruford M.W. & Beaumont M.A. (2001) Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics* **158,** 1347–62.

Clutton-Brock J. (1989) Cattle in ancient North Africa. In: *The Walking Larder: Patterns of Domestication, Pastoralism, and Predation* (Ed. by J. Clutton-Brock), pp. 200–14. Unwin Hyman, London.

European Cattle Genetic Diversity Consortium (2006) Marker-assisted conservation of European cattle breeds: an evaluation. *Animal Genetics* **37,** 475–81.

Evanno G., Regnaut S. & Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14,** 2611–20.

Felsenstein J. (1993) PHYLIP – Phylogeny Interference Package. Department of Genetics, Washington University, Seattle, WA, USA.

Freeman A.R., Meghen C.M., MacHugh D.E., Loftus R.T., Achukwi M.D., Bado A., Sauveroche B. & Bradley D.G. (2004) Admixture and diversity in West African cattle populations. *Molecular Ecology* **13,** 3477–87.

Freeman A.R., Bradley D.G., Nagda S., Gibson J.P. & Hanotte O. (2006) Combination of multiple microsatellite data sets to investigate genetic diversity and admixture of domestic cattle. *Animal Genetics* **37,** 1–9.

Goudet J. (2002) *FSTAT: Computer package for PCs*. Institute of Ecology, UNIL, Lausanne, Switzerland.

Hanotte O., Bradley D.G., Ochieng J.W., Emmeline Y.V., Hill W. & Rege J.E.O. (2002) African pastoralism: genetic imprints of origins and migrations. *Science* **296,** 336–9.

Hedrick P.W. (1999) Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53,** 313–8.

IBCR (2001) *Institute of Biodiversity Conservation and Research*. IBCR, Addis Ababa, Ethiopia. Available at: http://www.telecom.net.et/~ibcr/.

Ibeagha-Awemu E.M. & Erhardt G. (2005) Genetic structure and differentiation of 12 African *Bos indicus* and *Bos taurus* cattle breeds, inferred from protein and microsatellite polymorphisms. *Journal of Animal Breeding and Genetics* **122,** 2–20.

Ibeagha-Awemu E.M., Jann O.C., Weimann C. & Erhardt G. (2004) Genetic diversity, introgression and relationships among West/Central African cattle breeds. *Genetics, Selection, Evolution* **36,** 673–90.

Kantanen J., Olsaker I., Holm L.-E., Lien S., Vilkki J., Brusgaard K., Eythorsdottir E., Danell B. & Adalsteinsson S. (2000) Genetic diversity and population structure of 20 North European cattle breeds. *The Journal of Heredity* **91,** 446–57.

Li M.H. & Kantanen J. (2010) Genetic structure of Eurasian cattle (*Bos taurus*) based on microsatellites: clarification for their breed classification. *Animal Genetics* **41,** 150–8.

Li M.H., Nogovitsina E., Ivanova Z., Erhardt G., Vilkki J., Popov R., Ammosov I., Kiselyova T. & Kantanen J. (2005) Genetic contribution of indigenous Yakutian Cattle to two hybrid populations, revealed by microsatellite variation. *Asian-Australasian Journal of Animal Sciences* **18,** 613–9.

Li M.H., Tapio I., Vilkki J. *et al.* (2007a) The genetic structure of cattle populations (*Bos taurus*) in northern Eurasia and the neighbouring Near Eastern regions: implications for breeding strategies and conservation. *Molecular Ecology* **16,** 3839–53.

Li M.H., Zerabruk M., Vangen O., Olsaker I. & Kantanen J. (2007b) Reduced genetic structure of north Ethiopia cattle revealed by Y-chromosome analysis. *Heredity* **98,** 214–21.

Loftus R.T., MacHugh D.E., Bradley D.G., Sharp P.M. & Cunningham P. (1994) Evidence for two independent domestications of cattle. *Proceeding of the National Academy of Sciences of the United States of America* **91,** 2757–61.

Loftus R.T., Ertugrul O., Harba A.H., El-Barody M.A., MacHugh D.E., Park S.D. & Bradley D.G. (1999) A microsatellite survey of cattle from a centre of origin: the near East. *Molecular Ecology* **8,** 2015–22.

MacHugh D.E., Shriver M.D., Loftus R.T., Cunningham P. & Bradley D.G. (1997) Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurine* and *Bos indicus*). *Genetics* **146,** 1071–86.

Magee D.A., Meghen C., Harrison S., Troy C.S., Cymbron T., Gaillard C., Morrow A., Maillard J.C. & Bradley D.G. (2002) A partial African ancestry for the Creole cattle populations of the Caribbean. *The Journal of Heredity* **93,** 429–32.

Medugorac I., Medugorac A., Russ I., Veit-Kensch C.E., Taberlet P., Luntz B., Mix H.M. & Förster M. (2009) Genetic diversity of European cattle breeds highlights the conservation value of traditional unselected breeds with high effective population size. *Molecular Ecology* **18,** 3394–410.

Minch E., Ruiz-Linares A., Goldstein D.B., Feldman M.W. & Cavalli-Sforza L.L. (1995) Microsat (Version 1.4): a computer program for calculating various statistics on microsatellite allele data. Available at: http://hpgl.stanford.edu/projects/microsat/. Last accessed on 6 October 2011.

Moazami-Goudarzi K. & Laloe D. (2002) Is multivariate consensus representation of genetic relationships among populations always meaningful? *Genetics* **162,** 473–84.

Mommens G., Peelman L.J., Zeveren A.V., Ieteren G.D. & Wissocq N. (1999) Microsatellite variation between an African and five European taurine breeds results in a geographical phylogenetic tree with a bison outgroup. *Journal of Animal Breeding and Genetics* **116,** 325–30.

Nei M., Tajima F. & Tateno Y. (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19,** 153–70.

Oosterhout C.V., Hutchinson W.F., Wills D.P.M. & Shipley P. (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology* **4,** 535–8.

Ota T. (1993) *DISPAN: Genetic Distance and Phylogenetic Analysis*. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA, USA.

Pritchard J.K., Stephens M. & Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155,** 945–59.

Raymond M. & Rousset F. (1995) GENEPOP: a population genetics software for exact tests and ecumenicism. *The Journal of Heredity* **86,** 248–9.

Rege J.E.O. (1999) The state of African cattle genetic resources I. Classification framework and identification of threatened and extinct breeds. *Animal Genetic Resources Information Bulletin* **25,** 1–25.

Rege J.E.O. & Tawah C.L. (1999) The state of African cattle genetic resources II. Geographical distribution, characteristics and uses of present-day breeds and strains. *Animal Genetic Resources Information Bulletin* **26,** 1–25.

Rege J.E.O., Kahi A.K., Okomo-Adhiambo M., Mwacharo J. & Hanotte O. (2001) Zebu cattle of Kenya: uses, performance, farmer preferences, measures of genetic diversity and options for improved use. *Animal Genetic Resources Research* **1,** 103, ILRI (International Livestock Research Institute), Nairobi, Kenya.

Rice W.R. (1989) Analyzing tables of statistical tests. *Evolution* **43,** 223–5.

Saitou N. & Nei M. (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4,** 406–25.

Takezaki N. & Nei M. (1996) Genetic distances and reconstruction of phylogenetic tree from microsatellites DNA. *Genetics* **144,** 389–99.

Weir B.S. & Cockerham C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38,** 1358–70.

Wright S. (1978) *Evolution and the Genetics of Populations – Variability Within and Among Natural Populations. Vol. 4*. University of Chicago Press, Chicago, IL, USA.

Zerabruk M. & Vangen O. (2005) The Abergelle and Irob cattle breeds of north Ethiopia: description and on-farm characterization. *Animal Genetic Resources Information Bulletin* **36,** 7–20.

Zerabruk M., Bennewitz J., Kantanen J., Olsaker I. & Vangen O. (2007) Analysis of genetic diversity and conservation priorities for six North Ethiopian cattle breeds. *Journal of Animal Breeding and Genetics* **124,** 236–41.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 Allele frequency distribution for four loci in six cattle groupings.

Figure S2 Phylogenetic relationship between seven North Ethiopian cattle breeds constructed using NJ.

Figure S3 Factorial correspondence analysis based on allele frequencies of 20 microsatellites in seven North Ethiopian cattle breeds at the individual level.

Figure S4 Description of mean *Ln*P(D) and Δ*K* values computed by the software STRUCTURE.

Table S1 Summary information for the additional cattle breeds included in the analysis and the number of shared microsatellites genotyped in them, referenced to the 20 markers genotyped for the seven north Ethiopian cattle.

Table S2 Number of alleles, Nei's unbiased gene diversity, *F*-statistics (Weir & Cockerham 1984) estimates, frequency of null alleles, and the significance of deviation from Hardy–Weinberg expectation (HWE) at each locus across seven North Ethiopia cattle breeds.

Table S3 Frequencies of group (Indian zebu, Near-Eastern taurine, European taurine and African taurine) diagnostic alleles in the seven North Ethiopian cattle breeds.

Table S4 Summary of sample origin, breed group, sample size (n) and within-population genetic diversity, such as number of private alleles (A_P), mean number of observed alleles (A_O), expected (H_E) and observed (H_O) heterozygosity, fixation index (F_{IS}) in all the 17 cattle breeds based on the 11 common microsatellites.

Table S5 Pairwise estimates of F_{ST} (above diagonal) and D_{PS} (below diagonal) between seven North Ethiopian cattle breeds on the basis of 20 microsatellites.

Table S6 F_{ST} (above the diagonal) and D_{PS} (below the diagonal) estimates among the 17 cattle breeds on the basis of 11 common microsatellites.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.