

High variability of milk protein genes in *Bos indicus* cattle breeds of Cameroon and Nigeria and characterization of a new α_{s1} -casein promoter allele

Eveline M Ibeagha-Awemu, Eva-Maria Prinzenberg and Georg Erhardt*

Department of Animal Breeding and Genetics, Justus-Liebig University, Giessen, Ludwigstrasse 21b, 35390 Giessen, Germany

Received 17 October 2003 and accepted for publication 25 February 2004

The study provides the first comprehensive information on the variability of milk protein genes of *Bos indicus* and *Bos taurus* cattle breeds in Cameroon and Nigeria. The investigations indicate a high diversity of milk protein genes for the zebu populations. Of the investigated alleles, 21 out of 29 were observed. The method of single strand conformation polymorphism (SSCP) was a particularly useful technique because it allowed discrimination of alleles, including zebu-specific alleles at the *CSN2* (*I*) and *CSN3* (*A₁* and *H*) loci, not separated by protein electrophoretic techniques and also made possible the detection of a further *CSN1S1* 5' promoter allele (*CSN1S1Prom⁵*), which is also zebu-specific. Characterization of *CSN1S1Prom⁵* showed that it was the most variable of all described *CSN1S1* promoter alleles. A potential GATA consensus motif is created by mutations in *CSN1S1Prom⁵*. Intra-breed diversity measured as mean effective number of alleles was higher in the zebu populations than in the taurine breeds. Of the expected casein haplotypes, 96 out of 320 were present in the studied breeds. 2-C-A-A²-H (*CSN1S1Prom²*-*CSN1S1^C*-*CSN1S2^A*-*CSN2^{A2}*-*CSN3^H*) and 5-C-A-A²-H were zebu-specific while 1-B-A-A²-B was specific to the taurines. Overall distribution of alleles and haplotypes clearly separated the zebu populations from the taurine breeds. Zebu influence on the taurine breed Namchi was detected through the occurrence of zebu alleles and haplotypes. High variability of milk proteins also means availability of resources for breed development, phylogenetic studies, and conservation and management decisions.

Keywords: Variability, milk protein genes, haplotype, *Bos indicus*, *Bos taurus*.

Over the last five decades, improvement in genotyping techniques including chromatographic, electrophoretic and DNA-based methods have established the existence of the four caseins (α_{s1} -casein, *CSN1S1*; α_{s2} -casein, *CSN1S2*; β -casein, *CSN2* and κ -casein, *CSN3*) and two major whey proteins (α -lactalbumin, LAA and β -lactoglobulin, LGB) in different allelic forms, controlled by codominant autosomal genes (for reviews see Ng-Kwai-Hang & Grosclaude, 1992; Formaggioni et al. 1999). DNA-based methods for the discrimination of milk protein alleles are especially useful because they are not sex- or age-limited and also allow the discrimination of variants not detectable at the protein level (Prinzenberg et al. 1999).

In addition to the important function of milk as a provider of essential nutrients to man and animals, several studies have established varied levels of relationships

between milk protein variants and production/adaptability traits (Jairam & Nair, 1983; Hill et al. 1996; Ng-Kwai-Hang, 1997; Freyer et al. 1999), reproductive efficiency (Russo & Mariani, 1978), biological fitness of the newborn (Mercier et al. 1976), manufacturing properties of milk (Aleandri et al. 1990), as well as possible effects on human health (Elliott et al. 1999; McLachlan, 2001). Some of the associations have been confirmed by reports of significant QTL for several traits including birth weight, milk yield, milk percentage, fat percentage, milking speed and udder depth on chromosome 6 and other chromosomes (Kühn et al. 1999; Velmala et al. 1999; Boichard et al. 2003; Hiendleder et al. 2003). In addition, milk protein variants have been effectively used in the study of relationships between different gene pools, tracing of evolutionary history and in studying the unique characters of breeds necessary for breed conservation programmes (Baker & Manwell, 1980; Mahé et al. 1999; Pieragostini et al. 2000). Linkage relationship of the casein genes first

*For correspondence; e-mail: Georg.Erhardt@agr.uni-giessen.de

provided evidence of a segregating QTL for milk yield traits (Geldermann et al. 1985) and has also indicated zebu gene flow in Southern European breeds (Beja-Pereira et al. 2002; Jann et al. 2004). Polymorphism has also been observed at promoter regions of milk protein genes (Koczan et al. 1993; Schild & Geldermann, 1996; Prinzenberg et al. 2003) and there are also reports of alteration of putative transcription factor binding sites (Schild & Geldermann, 1996; Prinzenberg et al. 2003), which could have effects on gene expression (Rosen et al. 1999).

Data on the occurrence and frequencies of milk protein variants in the different cattle breeds are far from being complete, especially for African *Bos indicus* and *Bos taurus* breeds (Grosclaude et al. 1974; Mahé et al. 1999). Mahé et al. (1999) noted that this lack of data on milk protein allelic variations on African breeds is unfortunate, since the populations as a result of their admixed origin present a high level of genetic diversity. MacHugh et al. (1997) through microsatellite analysis, reported high diversity levels for African zebu populations. At the protein level, Ibeagha-Awemu et al. (2004) demonstrated further alleles at the albumin and transferrin blood protein loci in African zebu populations. Two new milk protein alleles of African *Bos indicus* and *Bos taurus* populations were detected by Mahé et al. (1999). This indicates that diversity levels at milk protein loci in African cattle populations are high and have not yet been fully exploited.

The present study therefore aimed to assess allelic variations at milk protein loci of the main cattle breeds of Cameroon and Nigeria, which might be necessary for breed improvement, relationship studies and conservation decisions and also serve as basis for other African breeds.

Materials and Methods

Samples

A total of 521 blood samples were obtained from five zebu (Red Bororo, $n=52$; White Fulani, $n=53$; Sokoto Gudali, $n=65$; Wadara, $n=36$; Adamawa Gudali, $n=11$) and two *Bos taurus* (N'Dama, $n=26$; Muturu, $n=20$) breeds in Nigeria and from four zebu (Red Bororo, $n=52$; White Fulani, $n=44$; Banyo Gudali, $n=77$; Ngaoundere Gudali, $n=55$) and one taurine (Namchi, $n=30$) breed in Cameroon. Animals sampled were unrelated and possessed typical breed characteristics. Details of the characteristics of the breeds are found in FDLPCS (1992), Felius (1995) and Mason (1996). DNA was isolated from white blood cells according to Montgomery & Sise (1990).

Genotyping of alleles

Alleles A and D of *CSN1S2* were assessed by amplification-created restriction site for *DdeI* enzyme as described by Prinzenberg (1998). Alleles A and B of *LAA* and *LGB* were investigated by PCR-RFLP according to, respectively, Mitra et al. (1998) and Medrano & Aguilar-Cordova (1990).

PCR-SSCP analysis was used to investigate variations at the *CSN1S1-5'* promoter region (*CSN1S1Prom*) (promoter alleles 1 to 4, Prinzenberg et al. 2003), *CSN1S1* (alleles B and C, Jann et al. 2002), *CSN2* (alleles A¹, A², A³, B, C and I, Barroso et al. 1999; Jann et al. 2001) and *CSN3* (alleles A, A₁, B, C, E, F, G, H and I, Prinzenberg et al. 1999) genes. Allele-specific PCR and PCR-RFLP were used to determine the presence of, respectively, *CSN1S1^F* (Prinzenberg et al. 1998) and *CSN1S1^A* (Prinzenberg & Erhardt, 1994).

Characterization of new *CSN1S1Prom* allele

A new allele of *CSN1S1Prom* identified by SSCP analysis was characterized by cloning and sequencing. One homozygous carrier of the new allele was PCR amplified, gel purified (E.Z.N.A.[®] Gel Extraction Kit, Peqlab Biotechnologie GmbH, Erlangen, Germany) and cloned into pCR[®] 2.1 TOPO plasmid using the TOPO-TA Cloning Kit (Invitrogen Corporation, Carlsbad, CA 92008, USA). Both directions of three clones and additional PCR products were cycle sequenced using dye terminator technique and reactions were run on an ABI PRISM[®] 377 DNA Sequencer (Applied Biosystems, PE Corporation, Foster City, CA 94404, USA) according to the manufacturer's instructions. Sequences were processed with Chromas Version 1.45 (<http://www.technelysium.com.au/chromas.html>) and compared with promoter sequences type 1 to 4 (GenBank Accession No. AF549499-502) with GeneDoc (Nicholas & Nicholas, 1997; <http://www.psc.edu/biomed/genedoc>). Possible transcription factor binding sites on the new sequence were identified with DNASIS[™] for Windows[®] (Hitachi Software Engineering America Ltd., San Bruno, CA 94066, USA) and TRANSFAC softwares (Wingender et al. 2000).

Statistical analysis

Allele frequencies at the analysed loci and test for Hardy-Weinberg equilibrium (HWE) were estimated using the GENEPOP program (Raymond & Rousset, 2001). Casein haplotype (*CSN1S1Prom-CSN1S1-CSN1S2-CSN2-CSN3*) frequencies were determined on the basis of genotype combinations using the EH software (Xie & Ott, 1993). The program uses a Maximum Likelihood algorithm to resolve the haplotype distribution including multiple heterozygous animals and provides data under the assumption of allelic association. Mean effective number of alleles, the reciprocal of heterozygosity (Hartl & Clark, 1989) was estimated per population using the POPGENE program (Yeh et al. 1999).

Results

Between the breeds, differences in occurrence and frequencies of the different alleles were observed (Table 1).

Table 1. Allele frequencies of milk proteins in *Bos indicus* and *Bos taurus* cattle breeds of Cameroon and Nigeria

Locus	Alleles	<i>Bos indicus</i>								<i>Bos taurus</i>			
		White Fulani (Nigeria)	White Fulani (Cameroon)	Red Bororo (Nigeria)	Red Bororo (Cameroon)	Sokoto Gudali	Banyo Gudali	Ngaoundere Gudali	Wadara	Adamawa Gudali	Namchi	Muturu	N'Dama
	<i>n</i>	53	44	52	52	65	77	55	36	11	30	20	26
CSN1S1 <i>Prom</i>	1	0.094	0.058	0.105	0.170	0.167	0.169	0.133	0.095	0.100	0.342	0.140	0.179
	2	0.395	0.558	0.527	0.390	0.491	0.537	0.547	0.527	0.650	0.658	0.860	0.679
	3	0.279	0.174	0.145	0.180	0.114	0.051	0.028	0.068	0.150	—	—	0.142
	4	0.058	0.082	0.039	0.060	0.044	0.088	0.094	0.054	—	—	—	—
	5	0.174	0.128	0.184	0.200	0.184	0.154	0.198	0.256	0.100	—	—	—
CSN1S1	A	—	—	—	—	—	—	—	—	—	—	—	—
	B	0.344	0.432	0.349	0.402	0.379	0.333	0.312	0.340	0.500	0.618	0.789	0.920
	C	0.656	0.568	0.651	0.598	0.621	0.667	0.688	0.660	0.500	0.382	0.211	0.080
	F	—	—	—	—	—	—	—	—	—	—	—	—
CSN1S2	A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.950	1.000	1.000
	D	—	—	—	—	—	—	—	—	—	0.050	—	—
CSN2	A ¹	0.059	0.113	0.104	0.174	0.145	0.176	0.064	0.203	0.500	0.383	0.579	0.680
	A ²	0.784	0.784	0.792	0.740	0.774	0.746	0.845	0.716	0.400	0.584	0.421	0.320
	B	0.079	0.080	0.052	0.019	0.040	0.070	0.082	0.054	0.050	—	—	—
	C	—	—	—	—	—	—	—	—	—	—	—	—
	I	0.078	0.023	0.052	0.067	0.041	0.008	0.009	0.027	0.050	0.033	—	—
CSN3	A	0.135	0.244	0.229	0.294	0.210	0.153	0.400	0.242	0.100	0.167	0.053	0.320
	A ₁	0.135	0.186	0.167	0.167	0.096	0.146	0.064	0.068	0.050	0.133	—	—
	B	0.192	0.233	0.188	0.206	0.210	0.306	0.118	0.149	0.550	0.433	0.947	0.680
	H	0.538	0.337	0.416	0.333	0.484	0.396	0.418	0.541	0.300	0.267	—	—
	C	—	—	—	—	—	—	—	—	—	—	—	—
	E	—	—	—	—	—	—	—	—	—	—	—	—
	F	—	—	—	—	—	—	—	—	—	—	—	—
	G	—	—	—	—	—	—	—	—	—	—	—	—
	I	—	—	—	—	—	—	—	—	—	—	—	—
LAA	A	0.235	0.091	0.202	0.147	0.210	0.153	0.082	0.270	0.250	0.117	—	—
	B	0.765	0.909	0.798	0.853	0.790	0.847	0.918	0.730	0.750	0.883	1.000	1.00
LGB	A	0.048	0.116	0.021	0.154	0.081	0.083	0.028	0.028	0.111	0.350	0.368	0.220
	B	0.952	0.884	0.979	0.846	0.919	0.917	0.972	0.972	0.889	0.650	0.632	0.780
¹ MNA		2.857	2.857	2.857	2.857	2.857	2.857	2.857	2.857	2.714	2.571	1.714	1.857
² MNE		1.906	1.938	1.901	2.134	1.910	1.895	1.705	1.820	1.839	1.879	1.464	1.365

¹ MNA=Mean absolute number of alleles

² MNE=Mean effective number of alleles

High variability of milk protein genes in zebu cattle

Table 2. Haplotype frequencies of casein genes (*CSN1S1Prom-CSN1S1-CSN1S2-CSN2-CSN3*) of *Bos indicus* and *Bos taurus* cattle breeds of Cameroon and Nigeria. Only frequencies of haplotypes ≥ 0.001 are shown. 38 haplotype combinations involving *CSN1S1Prom*⁴ and *CSN1S1Prom*⁵ were only present in the zebu populations and not included in table

*Haplotype	n	<i>Bos indicus</i>								<i>Bos taurus</i>				
		White Fulani (Nigeria)	White Fulani (Cameroon)	Red Bororo (Nigeria)	Red Bororo (Cameroon)	Sokoto Gudali	Banyo Gudali	Ngaoundere Gudali	Wadara	Adamawa Gudali	Namchi	Muturu	N'Dama	
1-B-A-A ¹ -A _I	—	—	—	0.014	—	—	—	—	—	—	—	—		
1-B-A-A ¹ -B	—	—	—	—	0.010	—	—	—	—	—	0.025	—	0.023	
1-B-A-A ² -A	—	0.023	—	0.001	—	0.008	—	—	—	—	0.050	—	0.097	
1-B-A-A ² -A _I	—	—	—	—	—	0.008	—	—	—	—	—	—	—	
1-B-A-A ² -B	—	—	—	—	—	—	—	—	—	—	0.025	0.289	0.020	
1-B-A-A ² -H	0.013	—	—	—	0.025	0.018	—	—	—	—	—	—	—	
1-B-D-A ² -A	—	—	—	—	—	—	—	—	—	—	0.050	—	—	
1-C-A-A ¹ -A	—	—	—	—	—	0.008	—	—	—	—	—	0.001	—	
1-C-A-A ¹ -A _I	—	—	—	—	—	—	—	—	0.020	—	—	—	—	
1-C-A-A ¹ -B	0.011	0.011	—	—	—	—	—	—	—	0.100	—	0.026	—	
1-C-A-A ¹ -B	—	—	—	—	—	—	—	—	—	0.001	—	—	—	
1-C-A-A ² -A	—	—	—	—	0.020	—	0.008	—	—	0.022	—	0.026	—	
1-C-A-A ² -A _I	0.024	0.023	—	—	0.067	0.020	0.008	—	—	—	—	—	—	
1-C-A-A ² -B	0.013	—	—	0.014	0.031	0.036	0.129	0.012	—	—	—	—	—	
1-C-A-A ² -H	0.015	—	—	0.068	—	0.067	0.019	0.130	—	—	0.001	—	—	
1-C-A-I-B	—	—	—	0.014	—	—	—	—	—	—	—	—	—	
2-B-A-A ¹ -A	—	—	—	0.056	0.050	0.008	—	0.001	—	—	—	—	0.065	
2-B-A-A ¹ -B	—	0.036	—	—	—	0.086	0.123	0.011	—	0.001	0.113	0.100	0.447	0.539
2-B-A-A ¹ -H	—	—	—	—	—	0.001	0.015	—	—	0.058	0.186	—	—	—
2-B-A-A ² -A	0.015	0.145	—	0.083	0.060	0.032	0.045	0.148	—	0.158	—	0.100	0.026	0.153
2-B-A-A ² -A _I	0.035	—	—	0.027	—	—	—	0.011	—	—	—	—	—	—
2-B-A-A ² -B	0.041	0.001	—	0.042	0.051	0.027	0.023	0.001	—	—	—	0.200	0.026	0.023
2-B-A-A ² -H	0.013	0.156	—	—	0.019	0.070	—	0.016	—	0.024	—	—	—	—
2-B-A-B-A	—	—	—	—	—	—	0.001	—	—	—	—	—	—	—
2-B-A-B-H	—	—	—	—	—	—	0.016	—	—	—	—	—	—	—
2-B-A-I-A	0.013	0.029	—	—	—	—	—	—	—	0.020	—	—	—	—
2-B-D-A ² -A	—	—	—	—	—	—	—	—	—	—	—	0.050	—	—
2-C-A-A ¹ -A	—	—	—	—	—	0.018	—	0.042	—	0.020	—	—	—	—
2-C-A-A ¹ -A _I	—	0.011	—	—	—	—	—	—	—	—	—	—	—	—
2-C-A-A ¹ -B	0.014	—	—	—	0.024	—	—	—	—	—	—	—	0.105	0.053
2-C-A-A ¹ -H	—	—	—	—	—	0.008	—	—	—	—	—	—	—	—
2-C-A-A ² -A	0.047	0.045	—	0.028	0.031	0.058	0.089	0.166	—	—	—	—	—	0.006
2-C-A-A ² -A _I	0.026	0.101	—	0.042	0.041	0.069	0.072	—	—	0.080	0.050	—	—	—
2-C-A-A ² -B	—	0.107	—	0.014	0.011	0.050	0.017	0.064	—	0.073	0.186	—	0.053	0.021
2-C-A-A ² -H	0.178	0.058	—	0.208	0.110	0.064	0.138	0.061	—	0.107	0.113	0.200	—	—
2-C-A-B-A	—	—	—	—	—	—	—	0.032	—	—	—	—	—	—
2-C-A-B-B	0.030	—	—	—	—	—	—	0.001	—	—	—	—	—	—

EM Ibeagha-Awemu and others

Table 2 continued on page 5

populations and 25% (24) in the taurine breeds. Each of 41.5% (40) of the observed haplotypes occurred only in one population and were considered private or breed-specific haplotypes. The haplotypes 2-C-A-A²-H (*CSN1S1Prom*²-*CSN1S1*^C-*CSN1S2*^A-*CSN2*^{A2}-*CSN3*^H) and 5-C-A-A²-H occurred at moderate to highest frequencies in the zebu populations, and were considered zebu-specific haplotypes. Nine other haplotypes (1-C-A-A²-A₁, 1-C-A-A²-B, 1-C-A-A²-H, 2-B-A-A²-H, 2-C-A-A²-A, 2-C-A-A²-A₁, 3-B-A-I-A, 3-C-A-A²-A₁ and 5-C-A-A²-B) were present at intermediate frequencies in most (5 to 8 populations) of the zebu breeds and totally absent or present at very low frequencies in the taurine breeds. These could also be specific to this breed group. All haplotypes containing the newly detected zebu allele *CSN1S1Prom*⁵ were absent, as expected, in the taurine breeds. The haplotype 2-B-A-A¹-B occurred at highest frequencies in the taurine breeds. 1-B-A-A²-B was present only in the taurine group and was considered a taurine haplotype. Three haplotypes, 2-B-A-A²-A, 2-B-A-A²-B, 2-C-A-A²-B, were common, occurring in 10 to 11 populations.

Five populations (White Fulani of Nigeria, Red Bororo of Nigeria, Sokoto Gudali and Banyo Gudali) deviated from HWE at the *CSN1S1Prom* locus while one population each deviated at the *CSN1S1* (Muturu, $P < 0.001$), *CSN2* (Wadara, $P < 0.05$) and *CSN3* (White Fulani of Nigeria, $P < 0.01$) loci.

Discussion

In addition to demonstrating 21 alleles at the six milk protein loci for all populations, further alleles-*CSN2*^I and *CSN3*^{A₁} and ^H, not detectable under protein electrophoretic conditions (Prinzenberg et al. 1999; Jann et al. 2001), were demonstrated and a new *CSN1S1* 5' promoter allele (*CSN1S1Prom*⁵) was detected. Three of these alleles, *CSN3*^{A₁} and ^H, and *CSN1S1Prom*⁵ are zebu-specific, indicating that protein electrophoretic procedures alone are not sufficient to detect milk protein variants in African cattle breeds and that more variation occurs at the DNA level. This is also demonstrated at the *CSN3* locus where Grosclaude et al. (1974) and Mahé et al. (1999) demonstrated *CSN3*^A as the major allele in zebu cattle. Our study, however, shows that *CSN3*^H is the predominant allele. *CSN3*^A is found at frequencies above 0.5 in most breeds (Grosclaude et al. 1974; Erhardt, 1993a; Del Lama & Zago, 1996; Lien et al. 1999; Mahé et al. 1999; Malik et al. 2000). This is true for some European taurine cattle but not for zebu breeds, owing to the failure of previous routine protein and DNA typing techniques to discriminate alleles *CSN3*^{A₁} and *CSN3*^H from *CSN3*^A. The high resolution power of SSCP clearly gave a lead in frequency values to the H allele followed by A and then A₁ for the zebu populations in our study. Frequencies of between 0.60 and 0.95 reported for *CSN3*^A in Madagasy zebu (Grosclaude et al. 1974), Gyr and Nelore (Del Lama &

Zago, 1996), Sudanese Fulani and Shuwa Arab (Mahé et al. 1999) and Indian Sahiwal (Malik et al. 2000) on the basis of protein electrophoretic or PCR-RFLP techniques are actually the combined frequencies for *CSN3* alleles A, A₁ and H. Recent studies using also PCR-SSCP differentiate the three alleles in Zebu Peul Soudanaise, Borgou (Ceriotti, 2002), Anatolian Black, Brahman, Casta Navarra, Nelore and Santa Gertrudis (Jann et al. 2004). The frequency value of *CSN3*^H (0.608) for Nelore (Jann et al. 2004) is higher than for all the zebu populations in this study. Brazilian Nelore originates from Indian Nelore, implying higher frequency values for Indian zebu breeds for this allele and a possible Asiatic origin for it. The *CSN3*^{A₁} and ^H alleles also occurred in the Namchi breed, thus suggesting zebu gene influence.

*CSN1S1Prom*² dominated in the taurine (0.658 to 0.860) more clearly than zebu populations (0.390 to 0.650) in this study, which therefore confirms its higher distribution in the German Holstein breed (0.738, Prinzenberg et al. 2003) and its status as the 'wild type' allele. Absence of *CSN1S1Prom*³ allele in two taurine populations and *CSN1S1Prom*⁴ in all three taurines is in contrast to their occurrence in the Holstein breed (Prinzenberg et al. 2003) and could provide a further explanation for the large genetic divergence between African taurines and European taurines (MacHugh et al. 1997). The African populations analysed were clearly separated at this locus with types 4 and 5 seen only in the zebu populations. Since type 4 is seen at similar frequencies in the Holstein breed (Prinzenberg et al. 2003), we suppose that *CSN1S1Prom*⁵, which occurred at reasonable frequencies (0.100 to 0.257) in the zebu populations analysed, is zebu-specific. Sequence analysis of this variant clearly differentiated it from types 1 to 4 and its highly variable nature may provide a further split between *Bos indicus* and *Bos taurus* breeds. Based on the sequence characteristics of the promoter alleles, it is likely that all other alleles evolved from type 2, and that while allele 4 further evolved from allele 3, alleles 1 and 5 had a common path from allele 2, but separated from each other at a certain point (parsimony tree not shown).

In comparing *CSN1S1Prom*¹ and *CSN1S1Prom*⁵, which share three mutation sites, potential YY1 and AP1 transcription factor binding sites lost in *CSN1S1Prom*¹ (Prinzenberg et al. 2003) were not affected in *CSN1S1Prom*⁵. Instead, mutations on *CSN1S1Prom*⁵ created a potential GATA transcription factor binding site. This is the first report of the presence of a potential GATA motif in the promoter region of cattle casein genes. The function and particular GATA binding-proteins that it may attract are not yet known.

Investigation of the effects of *CSN1S1Prom*⁵ on milk production traits may provide useful information explaining the low milk yields of zebu cattle. Studies on the effect of different *CSN1S1* promoter genotypes in Holstein cattle reveal significant effects on protein percentage and possible linked loci affecting milk yield (Prinzenberg et al. 2003).

Compared with European cattle, African taurines share common dominant alleles for some loci (Erhardt, 1993a; Lien et al. 1999) with great discrepancies between the two groups at the *CSN2* and *CSN3* loci. While *CSN2*^{A1} and *CSN3*^B were the dominant alleles in two taurine breeds in this study, and most other African taurines (Mahé et al. 1999), *CSN2*^{A2} and *CSN3*^A are the dominant alleles in about 60% of European taurines (Erhardt, 1993a; Lien et al. 1999; Beja-Pereira et al. 2002). High frequencies of *CSN2*^{A2} in the European populations in these reports could be the combined frequencies of *CSN2*^{A2} and *CSN2*^I. *CSN2*^I was first detected in four European cattle breeds by DNA-based techniques at frequencies from 0.03 to 0.14 (Jann et al. 2001) and recently observed in other African (Ceriotti et al. 2002) and European (Jann et al. 2004) populations at similar frequencies. The presence of *CSN2*^{A1}, recently reported to play a possible role in human diseases such as ischaemic heart disease and type 1 diabetes mellitus in some European countries, Canada, New Zealand and the United States (Elliott et al. 1999; McLachlan 2001), at frequencies above 10% in the majority of the zebu breeds in this study should be given serious consideration in improvement decisions on the breeds. Rare alleles *A* and *F* reported for the *CSN1S1* locus of some European cattle breeds (Bech & Kristiansen, 1990; Erhardt, 1993b) were not present in the investigated populations, even the African taurines. Even though *CSN1S1*^B dominated in the taurines in this study and occurs at frequencies above 0.5 in most European breeds, similar frequencies as observed for the studied zebu breeds have been reported for the Jersey (Larsen et al. 1974) and the Italian Calabrian (Bettini & Masina, 1972). Jersey and Podolic breeds also share several other alleles with zebu breeds (Porter, 1991; Pieragostini et al. 2000; Kaupe et al. 2004).

Observation of *CSN1S2*^D in the Namchi breed contradicts the lack of observation of polymorphism at this locus in a previous report on African cattle (Mahé et al. 1999). *CSN1S2*^D was until now only observed in some European cattle breeds (Grosclaude et al. 1978; Erhardt, 1993a) and therefore considered a variant of limited distribution. Its presence in the African Namchi therefore expands its borders and its importance in relationship studies between African and European gene pools.

Whey proteins allelic distributions in the studied populations agree with earlier observations on African cattle breeds (Blumberg & Tombs, 1958; Aschaffenburg, 1968; Grosclaude et al. 1974; Mahé et al. 1999) and Indian *Bos indicus* breeds (Aschaffenburg, 1968; Mitra et al. 1998). Insufficient data on *LAA*^A distribution in African zebu breeds led to the belief that *LAA*^A occurred at higher frequencies in Indian zebu (0.22 to 0.40) than in African zebu (0.03 to 0.15) (Aschaffenburg, 1968). Our results and other available data (Mitra et al. 1998; Mahé et al. 1999) indicate that *LAA*^A is similarly distributed in the two groups.

Resolution of alleles *CSN2*^I and *CSN3*^{A1} and *CSN3*^H and the new *CSN1S1* promoter allele (*CSN1S1Prom*⁵)

allowed the detection of more haplotype combinations (96) as opposed to the 20 detected by Mahé et al. (1999), 12 by Beja-Pereira et al. (2002) and 21 by Jann et al. (2004) and also allowed the postulation of 2-C-A-A²-H (*CSN1S1Prom*²-*CSN1S1*^C-*CSN1S2*^A-*CSN2*^{A2}-*CSN3*^H) and 5-C-A-A²-H as the predominant haplotypes in zebu populations, instead of C-A²-A (*CSN1S1*^C-*CSN2*^{A2}-*CSN3*^A). C-A²-A and B-C-A-A²-H (*CSN1S1Prom*^B-*CSN1S1*^C-*CSN1S2*^A-*CSN2*^{A2}-*CSN3*^H), C-C-A-A²-H, B-C-A-A²-A₁ and C-C-A-A²-A₁, which were contained within the zebu haplotypes in this study have also been found at moderate frequencies in the Mertolenga, Alentejana and Arouquesa breeds of Spain (Beja-Pereira et al. 2002) and some Eastern European cattle breeds (Jann et al. 2004), and may indicate zebu gene introgression. This is supported by observations at mtDNA (Cymbron et al. 1999). The occurrence of more casein haplotypes in the zebu than in taurine populations should be seen in connection with the higher level of genetic variation in these populations, as has been demonstrated by observations at blood (Ibeagha-Awemu et al. 2004) and microsatellite (MacHugh et al. 1997) loci, and implies an advantageous position for the zebu over the taurines in diversity and association studies.

Higher intra-breed diversity levels in the zebu than the taurine breeds were influenced by the occurrence of more alleles at the analysed loci in these populations. Higher diversity level in the Namchi breed than Muturu and N'Dama could be seen as an influence from the zebu populations. This is justified by reports of higher levels of zebu gene introgression in the Namchi than in N'Dama and Muturu with microsatellite (Hanotte et al. 2002) and blood protein (Ibeagha-Awemu et al. 2004) markers. A similar level of diversity of Adamawa Gudali breed with the other zebu populations (measured as MNE) indicates that its low sample size did not affect distribution of its alleles, as would have been expected. Identified HWE disturbances for some of the populations could be the effects of gene flow and/or genetic admixture.

The distribution of alleles, haplotypes and intra-breed diversities shows individuals under the forces of natural selection and present allele frequency gradients different from those due to the effects of selection by man (e.g., selection for different milk production traits; MacHugh et al. 1994; Lien et al. 1999; Jann et al. 2004). The populations therefore present resources that can be positively exploited for the development of specific products and in management and conservation decisions.

Financial support for EM Ibeagha-Awemu from the German Academic Exchange Service (DAAD) is gratefully acknowledged. The authors also thank Heike Wagner for skilful technical support.

References

- Aleandri R, Buttazzoni LG, Schneider JC, Caroli A & Davoli R 1990 The effects of milk protein polymorphisms on milk components and cheese producing ability. *Journal of Dairy Science* 73 241–255

- Aschaffenburg R** 1968 Reviews of the progress of dairy science. Section G. Genetics. Genetic variants of milk proteins: their breed distribution. *Journal of Dairy Research* **34** 447–460
- Baker CMA & Manwell C** 1980 Chemical classification of cattle. 1. Breed groups. *Animal Blood Groups and Biochemical Genetics* **11** 127–150
- Barroso A, Dunner S & Cañón J** 1999 A Multiplex PCR-SSCP test to genotype bovine β -casein alleles A¹, A², A³, B and C. *Animal Genetics* **30** 322–324
- Bech AM & Kristiansen KR** 1990 Milk protein polymorphism in Danish dairy cattle and the influence of genetic variants on milk yield. *Journal of Dairy Research* **57** 53–62
- Beja-Pereira A, Erhardt G, Matos C, Gama L & Ferrand N** 2002 Evidence for a geographical cline of casein haplotypes in Portuguese cattle breeds. *Animal Genetics* **33** 295–300
- Bettini TM & Masina P** 1972 [Milk protein polymorphism in dairy cows]. *Produzion Animali* **11** 107–126
- Blumberg BS & Tombs MP** 1958 Possible polymorphism of bovine alpha-lactalbumin. *Nature* **181** 683–684
- Boichard D, Grohs C, Bourgeois F, Cerqueira F, Faugeras R, Neau A, Rupp R, Amigues Y, Boşcher MY & Levéziel H** 2003 Detection of genes of economic traits in three French dairy cattle breeds. *Genetic Selection and Evolution* **35** 77–101
- Cerioti G, Tirella G, Chessa S, Caroli A & Crimella C** 2002 Milk protein variability in African Bos Genus breeds. In *Proceedings of the XXVIII International Conference on Animal Genetics*. Göttingen Germany, 11–15 August 2002, ISAG, pp. 88
- Cymbron T, Loftus RT, Malheiro MI & Bradley DG** 1999 Mitochondria sequence variation suggests an African influence in Portuguese cattle. *Proceedings Royal Society of London, B Biological Sciences* **266** 597–603
- Del Lama SN & Zago MA** 1996 Identification of the κ -casein and β -lactoglobulin genotypes in Brazilian *Bos indicus* and *Bubalus bubalis* populations. *Brazilian Journal of Genetics* **19** 73–77
- Elliott RB, Harris DP, Hill JP, Bibby NJ & Wasmuth HE** 1999 Type 1 (insulin-dependent) diabetes mellitus and cow milk: casein variant consumption. *Diabetologia* **42** 292–296
- Erhardt G** 1993a Allele frequencies of milk proteins in German cattle breeds and demonstration of α_{s2} -casein variant by isoelectric focusing. *Journal of Animal Breeding and Genetics* **36** 145–152
- Erhardt G** 1993b A new α_{s1} -casein allele in bovine milk and its occurrence in different breeds. *Animal Genetics* **24** 65–66
- FDLPCS** 1992 *Nigerian Livestock Resources. Volume II. National Synthesis*. Abuja, Nigeria: Federal Department of Livestock and Pest Control Services
- Felius M** 1995 *Cattle Breeds. An Encyclopedia*. Doetinchem, The Netherlands: Misset uitgeverij bv
- Formaggioni P, Sumner A, Malacarne M & Mariani P** 1999 Milk protein polymorphism: detection and diffusion of the genetic variants in *Bos genus*. *Annali della Facoltà di Medicina Veterinaria, Estratti, Università Di Parma, XIX* 127–165
- Freyer G, Liu Z, Erhardt G & Panicke L** 1999 Casein polymorphism and relation between milk production traits. *Journal of Animal Breeding and Genetics* **116** 89–97
- Geldermann H, Pieper U & Roth B** 1985 Effects of marked chromosome sections on milk performance in cattle. *Theoretical and Applied Genetics* **70** 138–146
- Grosclaude F, Mahé MF, Mercier JC & Ribadeau-Dumas B** 1974 [Comparison of genetic polymorphism of milk proteins of zebu and other cattle]. *Annales de Génétique et de Sélection Animale* **6** 305–309
- Grosclaude F, Joudrier P & Mahé MF** 1978 [Polymorphism of bovine casein α_{s2} : close linkage of the α_{s2} -Cn locus with the loci for α -Cn, β -Cn and κ -Cn; evidence for a deletion in the α_{s2} -Cn^D variant]. *Annales de Génétique et de Sélection Animale* **10** 313–327
- Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW & Rege JEO** 2002 African pastoralism: genetic imprints of origins and migrations. *Science* **296** 336–339
- Hartl DL & Clark AG** 1989 *Principles of Population Genetics*, 2nd Edn. Sunderland, MA, USA: Sinauer Associates
- Hiendleder S, Thomsen H, Leyhe-Horn B, Reinsch N, Looft C, Xu N, Medjugorac I, Russ I, Grupe S, Kühn C, Brockmann GA, Blümel J, Brenig B, Reinhardt F, Reents R, Averdunk G, Schwerin M, Förster M, Kalm E & Erhardt G** 2003 Mapping of QTL for body conformation and behaviour in cattle. *Journal of Heredity* **94** 496–506
- Hill JP, Boland MJ, Creamer LK, Anema SG, Otter DE, Peterson GR, Lowe R, Motion RL & Thresher WC** 1996 Effect of bovine β -lactoglobulin, milk composition, and dairy products. In *Macromolecular Interactions in Food Technology*, pp. 281–294 (Eds N Parris, A Kato, K Creamer & JR Pearce). ACS Symposium Series No. 650. Washington DC: American Chemical Society
- Ibeagha-Awemu EM, Jäger S & Erhardt G** 2004 Polymorphisms in blood proteins of *Bos indicus* and *Bos taurus* cattle breeds of Cameroon and Nigeria, and description of new albumin variants. *Biochemical Genetics* **42** 181–197
- Jairam BT & Nair PG** 1983 Genetic polymorphism of milk proteins and economic characters in dairy animals. *Indian Journal of Animal Science* **53** 1–8
- Jann O, Cerioti G, Caroli A & Erhardt G** 2001 A new variant in exon VII of bovine β -casein gene (CSN2) and its distribution among European cattle breeds. *Journal of Animal Breeding and Genetics* **119** 65–68
- Jann OC, Prinzenberg EM, Brandt H, Williams JL, Ajmone-Marsan P, Zaragoza P, Özbeyaz C & Erhardt G** 2002 Intragenic haplotypes at the bovine CSN1S1 locus. *Archiv für Tierzucht, Dummerstorf* **45** 11–19
- Jann OC, Ibeagha-Awemu EM, Özbeyaz C, Zaragoza P, Williams JL, Ajmone-Marsan P, Lenstra JA, Moazami-Goudarzi K & Erhardt G** 2004 Geographic distribution of haplotype diversity at the bovine casein locus. *Genetics Selection Evolution* **36** 243–257
- Kauep B, Winter A, Fries R & Erhardt G** 2004 DGAT1 polymorphism in *Bos indicus* and *Bos taurus* cattle breeds. *Journal of Dairy Research* **71** 182–187
- Koczan D, Hobom G & Seyfert HM** 1993 Characterization of the bovine α_{s1} -casein gene C-allele based on a MaeIII polymorphism. *Animal Genetics* **24** 74
- Kühn C, Freyer G, Weikard R, Goldammer T & Schwerin M** 1999 Detection of QTL for milk production traits in cattle by application of a specifically developed marker map of BTA6. *Animal Genetics* **30** 333–340
- Larsen B, Gruchy CL & Moustgaard J** 1974 Studies on blood groups and polymorphic protein systems in Jersey cattle on the Isle of Jersey. *Acta Agriculturae Scandinavica A* **24** 99–110
- Lien S, Kantanen J, Olsaker I, Holm L, Erythorsdottir E, Sandberg K, Dalsgard B & Adalsteinsson S** 1999 Comparison of milk protein allele frequencies in Nordic cattle breeds. *Animal Genetics* **30** 85–90
- MacHugh DE, Bradley DG, Sharp PM & Cunningham P** 1994 Microsatellite DNA variation within and among European cattle breeds. In *Proceedings of The Royal Society of London B* **256** 25–31
- MacHugh DE, Shriver MD, Loftus RT, Cunningham P & Bradley DG** 1997 Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics* **146** 1071–1086
- Mahé MF, Miranda G, Queval R, Bado A & Zafindrajaona F** 1999 Genetic polymorphism of milk proteins in African *Bos taurus* and *Bos indicus* populations. Characterization of variants α_{s1} -CN H and κ -CN J. *Genetics Selection Evolution* **34** 239–253
- Malik S, Kumar S & Rani R** 2000 κ -Casein and β -casein alleles in cross-bred and zebu cattle from India using polymerase chain reaction and sequence-specific oligonucleotide probes. *Journal of Dairy Research* **67** 295–300
- Mason IL** 1996 *A World Dictionary of Livestock Breeds, Types and Varieties*, 4th Edn. Wallingford, UK: CAB International
- McLachlan CN** 2001 Beta-casein A1, ischaemic heart disease mortality, and other illnesses. *Medical Hypotheses* **56** 262–272
- Medrano JF & Aguilar-Cordova E** 1990 Polymerase chain reaction amplification of bovine β -lactoglobulin genomic sequences and

- identification of genetic variants by RFLP analysis. *Animal Biotechnology* **1** 73–77
- Mercier JC, Addeo F & Pelissier JP** 1976 Primary structure of the casein macropeptide of caprine kappa casein. *Biochimie* **58** 1303–1310
- Mitra A, Sashikanth & Yadav BR** 1998 Alpha-lactalbumin polymorphism in three breeds of Indian Zebu cattle. *Journal of Animal Breeding and Genetics* **115** 403–405
- Montgomery GW & Sise JA** 1990 Extraction of DNA from sheep white blood cells. *New Zealand Journal of Agricultural Research* **33** 437–441
- Ng-Kwai-Hang KF** 1997 A review of the relationship between milk protein polymorphism and milk composition/milk production. In *Milk Protein Polymorphism*, pp. 22–37. Proceedings of the IDF Seminar held in Palmerston North, New Zealand, February 1997, IDF Brussels, Belgium
- Ng-Kwai-Hang KF & Grosclaude F** 1992 Genetic polymorphism of milk proteins. In *Advanced Dairy Chemistry, 1*, pp. 105–455 (Ed. PF Fox). London: Elsevier Science Publishers
- Nicholas KB & Nikolas HB** 1997 Genedoc: a tool for editing and annotating multiple sequence alignment. www.psc.edu/biomed/genedoc
- Pieragostini E, Scaloni A, Rullo R & Di Luccia A** 2000 Identical marker alleles in Podolic cattle (*Bos taurus*) and Indian zebu (*Bos indicus*). *Comparative Biochemistry and Physiology B* **127** 1–9
- Porter V** 1991 *Cattle – A Handbook to the Breeds of the World*. New York, USA: Facts on File, Inc
- Prinzenberg EM** 1998 Development of gene diagnostic tests for rare milk protein variants in cattle with special consideration of their occurrence in endangered breeds. Fachverlag Koehler, Germany, ISBN 3-922306-68-3
- Prinzenberg EM & Erhardt G** 1994 Characterization of bovine α_{S1} -casein A by PCR-RFLP. *Animal Genetics* **25** (Suppl. II) 30–31
- Prinzenberg EM, Anglade P, Ribadeau-Dumas B & Erhardt G** 1998 Biochemical characterization of bovine alpha S1-casein F and genotyping with sequence-specific primers. *Journal of Dairy Research* **65** 223–231
- Prinzenberg EM, Krause I & Erhardt G** 1999 SSCP analysis at the bovine CSN3 locus discriminates six alleles corresponding to known protein variants (A, B, C, E, F, G) and three new DNA polymorphisms (H, I, A₁). *Animal Biotechnology* **10** 49–62
- Prinzenberg EM, Weimann C, Brandt C, Bennewitz J, Kalm E, Schwerin M & Erhardt G** 2003 Polymorphism of the bovine CSN1S1 promoter: linkage mapping, intragenic haplotypes and effects on milk production traits. *Journal of Dairy Science* **86** 2696–2705
- Raymond M & Rousset F** 2001 Population genetics software and ecumenicism. *Journal of Heredity* **86** 248–249. <http://vbiomed.curtin.edu.au/genepop/>
- Rosen JM, Wyszomierski SL & Hadsell D** 1999 Regulation of milk protein gene expression. *Annual Review of Nutrition* **19** 407–436
- Russo V & Mariani P** 1978 [Polymorphism of milk protein in relation to genetic variations of interest to animal production] *Rivista Di Zootechnia E Veterinaria* **6** 289–304 and 365–379
- Schild TA & Geldermann H** 1996 Variants within the 5' flanking regions of bovine milk protein encoding genes. III genes encoding the Ca-sensitive caseins α_{S1} , α_{S2} and β . *Theoretical and Applied Genetics* **93** 887–893
- Velmala RJ, Vilkki HJ, Elo KT, De Koning DJ & Mäki-Tanila AV** 1999 A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle. *Animal Genetics* **30** 136–143
- Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Prüss M, Reuter I & Schacherer F** 2000 TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Research* **28** 316–319
- Xie X & Ott J** 1993 Testing linkage disequilibrium between a disease gene and marker loci. *American Journal of Human Genetics* **53** 1107
- Yeh FC, Yang R-C & Boyle T** 1999 POPGENE Version 1.31. Microsoft windows-based freeware for population genetics analysis. <ftp://ftp.microsoft.com/softlib/MSLFILES/HPGL.EXE>