ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi:10.1111/j.1365-2052.2009.01976.x

Origins and genetic diversity of New World Creole cattle: inferences from mitochondrial and Y chromosome polymorphisms

C. Ginja*, M. C. T. Penedo†, L. Melucci‡, J. Quiroz§, O. R. Martínez López¶, M. A. Revidatti, A. Martínez-Martínez††, J. V. Delgado‡‡ and L. T. Gama§§**

*Veterinary Genetics Laboratory, University of California, One Shields Avenue, Davis, CA 95616, USA; Instituto Superior de Agronomia, Lisboa, Portugal; Departamento de Genética e Melhoramento Animal, Instituto Nacional dos Recursos Biológicos, Fonte Boa, 2005-048 Vale de Santarém, Portugal. †Veterinary Genetics Laboratory, University of California, One Shields Avenue, Davis, CA 95616, USA. ‡Facultad Ciencias Agrarias, Universidad Nacional de Mar del Plata, Balcarce, Argentina; Estación Experimental Agropecuaria Balcarce, Instituto Nacional de Tecnología Agropecuaria, Balcarce, Argentina. §Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, México. ¶Departamento de Producción Animal, Facultad de Ciencias Agrarias, Universidad Nacional de Asunción, San Lorenzo, Paraguay; PDIZ, Universidade Federal Rural de Pernambuco, Pernambuco, Brazil. **Universidad Nacional del Nordeste, Corrientes, Argentina. ††Departamento de Genética, Universidad de Córdoba, Campus Rabanales C-5 14071, Córdoba, Spain. ‡‡Departamento de Genética, Universidad de Córdoba, Campus Rabanales C-5 14071, Córdoba, Spain. §§Departamento de Genética e Melhoramento Animal, Instituto Nacional dos Recursos Biológicos, Fonte Boa, 2005-048 Vale de Santarém, Portugal

Summary

The ancestry of New World cattle was investigated through the analysis of mitochondrial and Y chromosome variation in Creoles from Argentina, Brazil, Mexico, Paraguay and the United States of America. Breeds that influenced the Creoles, such as Iberian native, British and Zebu, were also studied. Creoles showed high mtDNA diversity ($H = 0.984 \pm 0.003$) with a total of 78 haplotypes, and the European T3 matriline was the most common (72.1%). The African T1a haplogroup was detected (14.6%), as well as the ancestral African-derived AA matriline (11.9%), which was absent in the Iberian breeds. Genetic proximity among Creoles, Iberian and Atlantic Islands breeds was inferred through their sharing of mtDNA haplotypes. Y-haplotype diversity in Creoles was high ($H = 0.779 \pm 0.019$), with several Y1, Y2 and Y3 haplotypes represented. Iberian patrilines in Creoles were more difficult to infer and were reflected by the presence of H3Y1 and H6Y2. Y-haplotypes confirmed crossbreeding with British cattle, mainly of Hereford with Pampa Chaqueño and Texas Longhorn. Male-mediated *Bos indicus* introgression into Creoles was found in all populations, except Argentino1 (herd book registered) and Pampa Chaqueño. The detection of the distinct H22Y3 patriline with the *INRA189-90* allele in Caracú suggests introduction of bulls directly from West Africa. Further studies of Spanish and African breeds are necessary to elucidate the origins of Creole cattle, and determine the exact source of their African lineages.

Keywords Creole cattle, genetic diversity, mitochondrial DNA, native breeds, Y chromosome.

Introduction

Cattle were introduced in the Americas after the first trip of Columbus in 1492. The Spanish introduced animals primarily through the Caribbean Islands, whereas the Portuguese route occurred for the most part through Brazil (Rouse 1977). The Atlantic Islands such as Canary and Cape Verde were important intermediate ports between the Iberian Peninsula and the Western Hemisphere and were used as livestock depositories by the Spanish and the Portuguese respectively (Rouse 1977; Rodero *et al.* 1992). Information regarding the appearance and types of cattle brought to the Americas is scarce. Even though cattle brought from southern Iberia are thought to be the main source, some authors refer to animals being transported from other regions, such as the north of Portugal and

Address for correspondence

C. Ginja, Veterinary Genetics Laboratory, University of California, One Shields Avenue, Davis, CA 95616, USA.
E-mail: cjginja@ucdavis.edu

Accepted for publication 14 August 2009

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*

Galicia (Primo 1992; Rodero *et al.* 1992). Cattle spread throughout the American continent, became well adapted to a wide range of environmental conditions, and formed the so called Creoles (Rouse 1977). The introduction of Northern European breeds such as Angus, Hereford and Shorthorn, as well as zebu cattle (*Bos indicus*) during the 19th century, threatened the Creole cattle which remained in isolated regions (Rouse 1977; Giovambattista *et al.* 2001). Iberian cattle share a common Middle Eastern ancestry with other European breeds and it is believed that the flow of livestock to the Peninsula occurred through the mainland and the Mediterranean region (Cortes *et al.* 2008). Another source of cattle are African animals thought to have arrived in Iberia during the Bronze Age (Anderung *et al.* 2005; Beja-Pereira *et al.* 2006), but also introduced later during the Moorish occupation from the 8th to the 13th century (Cymbron *et al.* 1999). The contribution of aurochs to the ancestry of Iberian and Mediterranean cattle breeds is debated (Beja-Pereira *et al.* 2006; Achilli *et al.* 2008).

The analysis of mitochondrial DNA (mtDNA) sequence variation has been widely used to study the origins of domestic cattle and breed relationships (Bruford *et al.* 2003). Ten major mtDNA haplogroups were identified in cattle, corresponding to the ancient *Bos primigenius* matriline (P), the indicine I1 and I2, and the taurine T, T1, T2, T3, T4, T5 and Q (Loftus *et al.* 1994; Troy *et al.* 2001; Mannen *et al.* 2004; Lai *et al.* 2006; Achilli *et al.* 2008). These haplogroups are geographically structured with predominance of T1 and T1a in Africa (Troy *et al.* 2001; Achilli *et al.* 2008), T2 in Western Asia (Lai *et al.* 2006), T3 in Europe (Troy *et al.* 2001) and T4 in East Asia (Mannen *et al.* 2004). T5 was found in a few animals from Italy and Iraq (Achilli *et al.* 2008). Haplogroup Q was detected in native Italian breeds and appears to represent aurochs mtDNA (Achilli *et al.* 2008). A new African-derived lineage (AA) was found in Creole cattle (Miretti *et al.* 2002).

One limitation of mtDNA analysis is that it retrieves an incomplete history, as it reflects exclusively the genetic relationship through matrilines. Y chromosome markers complement this information, and are useful to study the origins and evolution of domestic animals (Gotherstrom *et al.* 2005; Meadows *et al.* 2006; Anderung *et al.* 2007; Bollongino *et al.* 2008). In cattle, Y chromosome microsatellites (STRs) were used to distinguish between *B. taurus* and *B. indicus* ancestry (Bradley *et al.* 1994). Gotherstrom *et al.* (2005) used SNPs to define three cattle Y-haplogroups (Y1 and Y2 in *B. taurus* and Y3 in *B. indicus*) and to investigate aurochs contributions to the genetic composition of modern breeds. Recently, Ginja *et al.* (2009) used a combination of SNPs and STRs specific to the non-recombining region of the Y chromosome to describe 13 haplotypes in Portuguese cattle.

The first study of mtDNA lineages in Creole cattle was reported by Miretti *et al.* (2002) in one Argentinean and four Brazilian breeds. Other mtDNA studies corroborated the European ancestry of Creoles, but also detected the influence of African cattle (Magee *et al.* 2002; Carvajal-Carmona *et al.* 2003; Mirol *et al.* 2003; Miretti *et al.* 2004). However, none of these studies included North American Creoles. Information on Creole patrilines is also very scarce and, to our knowledge, limited to one report in which the STR *INRA124* was used to detect male-mediated *B. indicus* introgression (Giovambattista *et al.* 2000). We analysed the genetic diversity of Creole cattle using mtDNA and Y chromosome markers. The origins of North and South American Creoles and the contribution of breeds from the Iberian Peninsula, the Atlantic Islands, and Continental Europe to their genetic makeup were investigated. Male-mediated *B. indicus* introgression was also assessed based on Y-haplotypes.

Materials and methods

Sample collection and DNA extraction

Samples were collected from 413 animals (106 females, 307 males) of 13 native Portuguese cattle breeds: Alentejana (8, 31), Arouquesa (8, 31), Barrosã (8, 33), Brava de Lide (8, 26), Cachena (8, 25), Garvonesa (10, 6), Marinhoa (8, 17), Maronesa (8, 23), Mertolenga (8, 17), Minhota (8, 28), Mirandesa (8, 23), Preta (8, 29) and Ramo Grande (8, 18). Samples from Spain included three native breeds: Canaria (8, 14), Mostrenca (8, 21) and Palmera (8, 25). For the related Creoles of the Americas, we collected a total of 352 (97, 158) samples. From Argentina, we sampled Argentino1 (8, 18), which included registered animals from Central Argentina, and Argentino2 (6, 6) which were free ranging animals from the north-eastern part of the country. From Brazil, we sampled the Caracú breed (48 males). The Mexican Creoles were sampled in Baja California (20 females), Chiapas (8, 12), Chihuahua (15, 4), Nayarit (16 females) and Puebla (16 females). Additionally, we sampled one Creole population in Paraguay, Pampa Chaqueño (8, 25) and the North American Texas Longhorn (45 males). European (35, 195) and Zebu (19, 28) breeds which may have influenced Creoles were also analysed: Angus (6, 41), British White (21 males), Charolais (8, 13), Friesian (8, 27), Hereford (5, 45), Jersey (20 males), Limousin (8, 17), Shorthorn (11 males), Brahman (13, 7) and Gyr (6, 21). Samples from British cattle were obtained in the USA, except for 12 Angus and 11 Hereford which were from Argentina. The zebu samples were collected in Mexico, except for five Brahman bulls from which semen samples were obtained from a commercial source (Bovine Elite, Llc.). Charolais, Friesian and Limousin were sampled in Portugal. For most breeds, care was taken to select individuals unrelated to the first and often to the second degree based on herd book registration records. For

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Argentino2, Caracú, Mexican Creoles and Pampa Chaqueño pedigree information was not available; samples were collected from various herds at different locations to minimize relatedness. DNA was extracted from whole blood, semen or hair roots using the Gentra kit (PUREGENE®, Gentra Systems, Inc.).

PCR amplification and mtDNA sequencing

A fragment of approximately 1.1 kb was amplified with forward primer CTGCAGTCTCACCATCAACC (L15737; Loftus *et al.* 1994) and reverse primer CCTAGAGGGC-ATTCTCACTGGG (H516; Miretti *et al.* 2002). Platinum® High fidelity *Taq* DNA polymerase and SuperMix (Invitrogen) were used in 25 μl volume PCRs following manufacturer's recommendations. PCR cycling consisted of initial denaturation at 94 °C for 2 min and then 40 cycles at 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 90 s with a final extension at 72 °C for 10 min. Direct cycle sequencing was performed with the BigDye Terminator v3.1 kit (Applied Biosystems), with the forward primer and an internal forward primer CTTAATTACCATGCCGCGTG (MacHugh *et al.* 1999). Products were separated on ABI 3730 DNA Analyzer instruments (Applied Biosystems) and sequences analysed with SEQMAN™ II v6.1 (DNASTAR Inc.).

Statistical analysis of mtDNA data

The mtDNA sequences were aligned with the ClustalW multiple alignment algorithm of BIOEDIT v7.0.1 (Hall 1999). Haplotype identification was performed with GENALEX v6.0 (Peakall & Smouse 2006). ARLEQUIN v2.0 (Schneider *et al.* 2000) was used to estimate haplotype diversity ($H \pm$ SD), nucleotide diversity ($\pi \pm$ SD) for each population (Nei 1987) and pairwise F_{ST} values (5% significance level obtained with 10000 permutations). The mean number of pairwise nucleotide differences (MNPD) (Tajima 1983) within each population was estimated assuming heterogeneity of substitution rates per site (Tamura & Nei 1993). A hierarchical likelihood ratio test was used with MODELTEST v3.7 (Posada & Crandall 1998; available at http://darwin.uvigo.es/) to determine the best-fitting evolutionary model for the data and to estimate gamma distribution parameters (Posada & Crandall 2001). Analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) was performed with ARLEQUIN to test the effect of geographical breed grouping (Iberian Peninsula, Atlantic Islands, Creoles, Zebu, British Islands, and Continental Europe) on total genetic variance. NETWORK v4.5.1.0 (Fluxus Technology Ltd, 2004–2009) was used to construct Median-Joining (MJ) networks (Bandelt *et al.* 1999) between haplotypes or haplogroups with nucleotide (nt) substitutions weighted 10 for transitions and 30 for transversions. First a reduced-median (RM) network was constructed using a binary dataset, and the outputs were then used to obtain the MJ networks.

Analysis of Y chromosome SNPs and STRs

Y chromosome haplotype analysis was performed with 13 Y-specific markers as described by Ginja *et al.* (2009). Haplotype frequency and diversity ($H \pm$ SD), pairwise F_{ST} values (10000 permutations for 5% significance level) and AMOVA were obtained with ARLEQUIN v2.0 (Schneider *et al.* 2000). MJ networks of haplotypes were constructed as described in Ginja *et al.* (2009).

Results

mtDNA haplotype diversity

Sequences of the complete *D-loop* region (919 bp) were aligned with the *B. taurus* (Eucons) reference sequence (Anderson *et al.* 1982). A total of 311 haplotypes (accession numbers FJ815445-FJ816013) were identified in 569 animals and 35 cattle populations. Among 142 total variable sites detected, 86 were phylogenetically informative, 37 were singletons and 19 were indels (Table S1). Nucleotide frequencies for this region were: A = 0.3287; C = 0.2430; G = 0.1473; and T = 0.2809. The transition/transversion rate of 21.1:1 determined by MODELTEST confirmed a strong bias towards transitions. Haplotype frequencies across breeds are shown in Table S2. Hap209 was shared by several breeds of the six geographical groups and was identical to the *Eucons*. In the Creole group, Hap209 was the most common and was found in five populations. Creoles shared the highest number of haplotypes (9) with the Iberian Peninsula group, followed by the British Islands (4), the Atlantic Islands (2) and Zebu (2). Three haplotypes (Hap054, Hap243 and Hap274) were shared exclusively among Iberian, Atlantic Islands and Creole cattle. Only one haplotype was shared between British (Angus), Zebu (Brahman) and the Creoles (Baja California and Nayarit). Hap039 was found in two individuals from the Atlantic Islands (Ramo Grande and Palmera) and in one from Mexico (Nayarit), but also in two Angus from Argentina. Hap211 was found in one Garvonesa, two Creoles from Mexico (Baja California and Chihuahua) and in one Angus.

Estimates of mtDNA genetic diversity within cattle populations and in each of the six geographical groups are shown in Table 1. A total of 78 haplotypes were detected among the Creoles, of which 48 were unique of this group. The overall diversity estimates for Creoles were $H = 0.984 \pm 0.003$, $\pi = 0.006 \pm 0.003$ and MNPD = 5.0 ± 2.5. Among all breeds analysed, Palmera had only six haplotypes and the lowest diversity for all estimates ($H = 0.857 \pm 0.056$, $\pi = 0.002 \pm 0.001$, MNPD = 2.0 ± 1.2). Cachena, Minhota, British White and Shorthorn had the highest diversity, with all animals in each breed showing different haplotypes. Haplotype differences, estimated as MNPD and π, tended to be uniform across breeds, with the

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Table 1 Estimates of mtDNA genetic diversity within 35 cattle populations and six geographical groups.

Breeds	*N*	*H*	SD	π	SD	MNPD	SD	TNH	EH	F_{ST} (%)	SD
Arouquesa	16	0.992	0.025	0.004	0.002	3.8	2.0	15	11	3.7	0.032
Barrosã	16	0.983	0.028	0.004	0.002	3.6	1.9	14	6	5.7	0.052
Cachena	16	1.000	0.022	0.005	0.003	4.9	2.5	16	6	2.3	0.055
Marinhoa	16	0.967	0.036	0.004	0.003	4.0	2.1	13	6	6.1	0.043
Maronesa	16	0.958	0.036	0.002	0.001	2.1	1.2	12	8	7.2	0.055
Minhota	15	1.000	0.024	0.006	0.003	5.6	2.8	15	9	2.0	0.050
Mirandesa	16	0.992	0.025	0.003	0.002	3.2	1.7	15	9	3.3	0.052
Alentejana	16	0.975	0.029	0.008	0.004	7.4	3.6	13	10	5.9	0.069
Brava de Lide	16	0.983	0.028	0.004	0.002	3.7	2.0	14	9	3.2	0.056
Garvonesa	16	0.958	0.036	0.004	0.002	3.3	1.8	12	5	2.7	0.041
Mertolenga	16	0.992	0.025	0.004	0.003	4.0	2.1	15	9	3.3	0.055
Preta	16	0.967	0.036	0.004	0.002	3.6	1.9	13	7	7.6	0.051
Mostrenca	16	0.967	0.031	0.003	0.002	3.1	1.7	12	10	4.9	0.051
Iberian Peninsula	*207*	*0.993*	*0.002*	*0.005*	*0.003*	*4.1*	*2.1*	*140*	*117*	*4.4*	*1.881*
Ramo Grande	16	0.958	0.036	0.004	0.002	3.3	1.8	12	4	3.6	0.049
Canaria	15	0.933	0.054	0.004	0.002	3.5	1.9	11	5	3.9	0.046
Palmera	14	0.857	0.056	0.002	0.001	2.0	1.2	6	3	8.7	0.066
Atlantic Islands	*45*	*0.946*	*0.017*	*0.003*	*0.002*	*3.1*	*1.6*	*23*	*12*	*5.4*	*2.867*
Argentino1&2	23	0.988	0.016	0.004	0.002	3.8	2.0	20	9	4.6	0.055
Baja California	20	0.895	0.043	0.007	0.004	6.0	3.0	10	6	7.7	0.026
Caracú	10	0.867	0.107	0.005	0.003	5.1	2.7	7	4	19.0	0.073
Chiapas	15	0.962	0.040	0.004	0.002	3.8	2.0	12	4	3.5	0.045
Chihuahua	19	0.936	0.047	0.005	0.003	4.4	2.3	14	7	3.9	0.036
Poblano	16	0.983	0.028	0.004	0.002	3.6	1.9	14	4	3.9	0.054
Nayarit	16	0.950	0.036	0.004	0.002	3.7	2.0	11	2	3.3	0.047
Pampa Chaqueño	16	0.917	0.049	0.007	0.004	6.8	3.4	10	4	21.0	0.060
Texas Longhorn	16	0.942	0.041	0.005	0.003	4.4	2.3	11	3	7.4	0.048
Creoles	*151*	*0.984*	*0.003*	*0.006*	*0.003*	*5.0*	*2.5*	*78*	*48*	*8.3*	*6.857*
Brahman	20	0.979	0.021	0.005	0.003	4.6	2.4	16	8	6.3	0.055
Gyr	9	0.972	0.064	0.003	0.002	3.1	1.8	8	5	11.3	0.055
Zebu	*29*	*0.988*	*0.012*	*0.005*	*0.003*	*4.4*	*2.2*	*24*	*13*	*8.8*	*3.537*
Angus	27	0.989	0.013	0.004	0.002	3.6	1.9	23	14	3.1	0.058
British White	10	1.000	0.045	0.003	0.002	2.9	1.7	10	6	6.5	0.052
Hereford	25	0.993	0.013	0.005	0.003	4.3	2.2	23	14	4.0	0.039
Jersey	18	0.974	0.029	0.003	0.002	2.5	1.4	15	10	6.0	0.054
Shorthorn	9	1.000	0.052	0.004	0.002	3.7	2.1	9	7	3.4	0.038
British Islands	*89*	*0.992*	*0.004*	*0.004*	*0.002*	*3.6*	*1.8*	*73*	*51*	*4.6*	*1.548*
Charolais	16	0.975	0.029	0.004	0.002	3.8	2.0	13	9	3.2	0.061
Friesian	16	0.983	0.028	0.005	0.003	4.2	2.2	14	10	3.4	0.068
Limousin	16	0.942	0.041	0.005	0.003	4.4	2.3	16	10	4.6	0.054
Continental Europe	*48*	*0.991*	*0.007*	*0.004*	*0.003*	*4.1*	*2.1*	*41*	*30*	*3.7*	*0.767*

N, sample size; *H*, haplotype diversity; π, population nucleotide diversity; MNDP, mean number of pairwise nucleotide differences (including indels) between haptotypes; TNH, total number of haplotypes; EH, exclusive haplotypes; SD, standard deviation.

exception of the higher values found in Alentejana, the Creoles Baja California and Pampa Chaqueño.

mtDNA differentiation and AMOVA

Pairwise F_{ST} values estimated between 35 cattle populations are shown in Table S3 and average pairwise F_{ST} values per breed and group are summarized in Table 1. Among all breeds, Pampa Chaqueño, Caracú and Gyr had the highest average F_{ST} (21.0 ± 0.060, 19.0 ± 0.073 and 11.3 ± 0.055 respectively). Among Iberian breeds, Preta and Maronesa had the highest average F_{ST} (7.6 ± 0.051 and 7.2 ± 0.055 respectively) whereas Minhota had the lowest (2.0 ± 0.050). Within the Atlantic Islands, Palmera had the highest average F_{ST} (8.7 ± 0.066). Within the Creoles, Nayarit had the lowest average F_{ST} (3.3 ± 0.047). For the British Islands and Continental Europe breed groups, Angus and Charolais had the lowest average F_{ST} (3.1 ± 0.058 and 3.2 ± 0.061 respectively), whereas British White had the highest (6.5 ± 0.052). When pairwise F_{ST} was calculated among breed groups, Creoles and Zebu were significantly ($P < 0.05$) differentiated from all

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

others (results not shown). The AMOVA analysis showed significant differences among breeds within groups ($P < 0.0001$) that accounted for 5.1% of the total genetic variation (Table S4). Geographical clustering of breeds had a significant effect ($P < 0.05$) but explained only 1.2% of the total variance, such that 93.7% of the variation corresponded to differences among individuals. When AMOVA was performed for major haplogroups (European T3, Iberian Q, African T1/T1a, and African-derived Creole AA) found in each region, results were highly significant ($P < 0.0001$), with differences among haplogroups accounting for most (52%) of the genetic variation (Table S4). A much lower effect of individual differences was observed (46%), whereas differences among regions within haplogroups explained only 1.5% of the total variation.

mtDNA phylogenetics

The best-fitting model of sequence evolution determined with MODELTEST was the *HKY + I + G* of Hasegawa *et al.* (1985). The shape parameter of the gamma distribution (α) was 0.7201 and the proportion of invariable sites (I) was 0.7693. The RM networks (Fig. 1) show the relationships between mtDNA haplotypes in five geographical regions: Americas (Creoles), Atlantic Islands, Iberian Peninsula, Continental Europe and the British Islands. The zebu breeds (Brahman and Gyr) were included within the Creoles. Reference sequences (Table S5) representing known mtDNA haplogroups were included in each network. In every geographical region, most haplotypes clustered within two distinct haplogroups: the African (T1a) and the European (T3). For the Creoles, a total of 22 animals had T1a African type (found in every breed) and 109 individuals belonged to the European T3 haplogroup. We also found one Chihuahua of T1 type; four Caracús, seven Pampa Chaqueños, and one Chihuahua of AA type; and one Chihuahua of T2 type. We identified new haplotypes related to the AA lineage in five animals from Baja California (Hap089) and one Chihuahua (Hap293) which clustered together in the Creoles RM network (Fig. 1a). A more detailed analysis of all African-derived haplotypes (Fig. S1) suggests that Hap089 is more closely related to the AA group and that Hap293 more likely belongs to the T1a haplogroup, even though it has the transition at nt 16222. Interestingly, all *B. indicus* individuals showed *B. taurus* matrilines, most of which belonged to the T3 group, except for 6 Brahman that showed African T1a types. In the Iberian Peninsula, 175 samples were of T3 type, and the T1a haplogroup was detected in 26 animals across different breeds, except for Marinhoa, Maronesa and Preta. Three Marinhoa and one Mertolenga had the T2 type, and one Alentejana the Q type. One Arouquesa showed a Q haplotype (Hap086); however, it clustered within T3. Coding-region sites are needed to infer Q reliably (Achilli *et al.* 2008), and even though this Arouquesa had the transversion typical of the Q haplogroup (nt 15953), it was similar to the reference T3. In the Atlantic Islands, Ramo Grande and Canaria showed African influence with six animals of T1a type and one Canaria of T1 type, whereas all Palmeras had T3 type. Among the Continental European and British breeds, most animals had T3 type, except two Charolais with T1a (Hap003) and one with T2 type (Hap147), two Angus from Argentina which were T1a (Hap075 and Hap079), and three Herefords from Argentina which were AA (one Hap070 and two Hap268). African-derived matrilines have not been reported in these breeds (Troy *et al.* 2001) and thus their presence in these seven animals most probably results from crossbreeding.

Y chromosome diversity

Among the 24 Y-haplotypes found in 748 bulls, seven were of the Y1 and 11 were of the Y2 haplogroups characteristic of *B. taurus* (Table 2). Additional variation was found within the *B. indicus* Y3 haplogroup, with six haplotypes identified. One SNP (*UTY* intron 19 AY936543*:* g.423C>A) and the indel (*ZFY* intron 10 AF241271: g.697_8indelGT) were sufficient to discriminate Y1 and Y2 haplogroups. The remaining SNPs (*DDX3Y* intron 1 AY928816: g.425C>T*; DDX3Y* intron 7 AY928819: g.123C>T; *ZFY* intron 9 AY928828: g.120C>T; and *ZFY* intron 10 AF241271: g.655C>T) and three STRs *(DDX3Y_1, BM861 and INRA124*) distinguished between *B. taurus* and *B. indicus* lineages. Three STRs (*INRA189, UMN0103* and *UMN0307*) were variable and defined new haplotypes in both taurine and zebu cattle. One STR (*UMN0504*) was monomorphic across breeds. The most frequent haplotypes were H6Y2 (0.227) and H11Y2 (0.148) followed by H3Y1 (0.131) and H14Y1 (0.128). Within the Y3 haplogroup, the most frequent haplotype was H18Y3 (0.053), which was fixed in Gyr. Overall, Y-haplotype diversity was generally low (mean $H = 0.240 \pm 0.295$), with several breeds showing a fixed haplotype (Table 3). We identified 12 private haplotypes with frequencies ranging between 0.04 in Brava de Lide (H5Y1) and 1.00 in Caracú (H22Y3). Furthermore, six haplotypes were found exclusively in the Iberian Peninsula, five in the Creoles, two in the British breeds, and one each in the Atlantic Islands and Continental European breed groups. The highest diversity was found among Creoles ($H = 0.779 \pm 0.019$), Iberian breeds ($H = 0.712 \pm 0.016$), and also within the Atlantic Islands ($H = 0.677 \pm 0.037$). While few breeds were sampled within these groups, the British Islands, Continental Europe and Zebu had lower diversity ($H = 0.636 \pm 0.025$, $H = 0.543 \pm 0.026$ and $H = 0.389 \pm 0.084$ respectively), which reflects a low number of effective males within these more intensively selected breeds. Among Iberian breeds of the Y1 haplogroup, Mostrenca was fixed for H3Y1 and Mertolenga had H2Y1 in high frequency. Within the Canary Islands, Canaria showed both Y1 and Y2 types, whereas Palmera was almost fixed for H23Y1 which was exclusive of

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

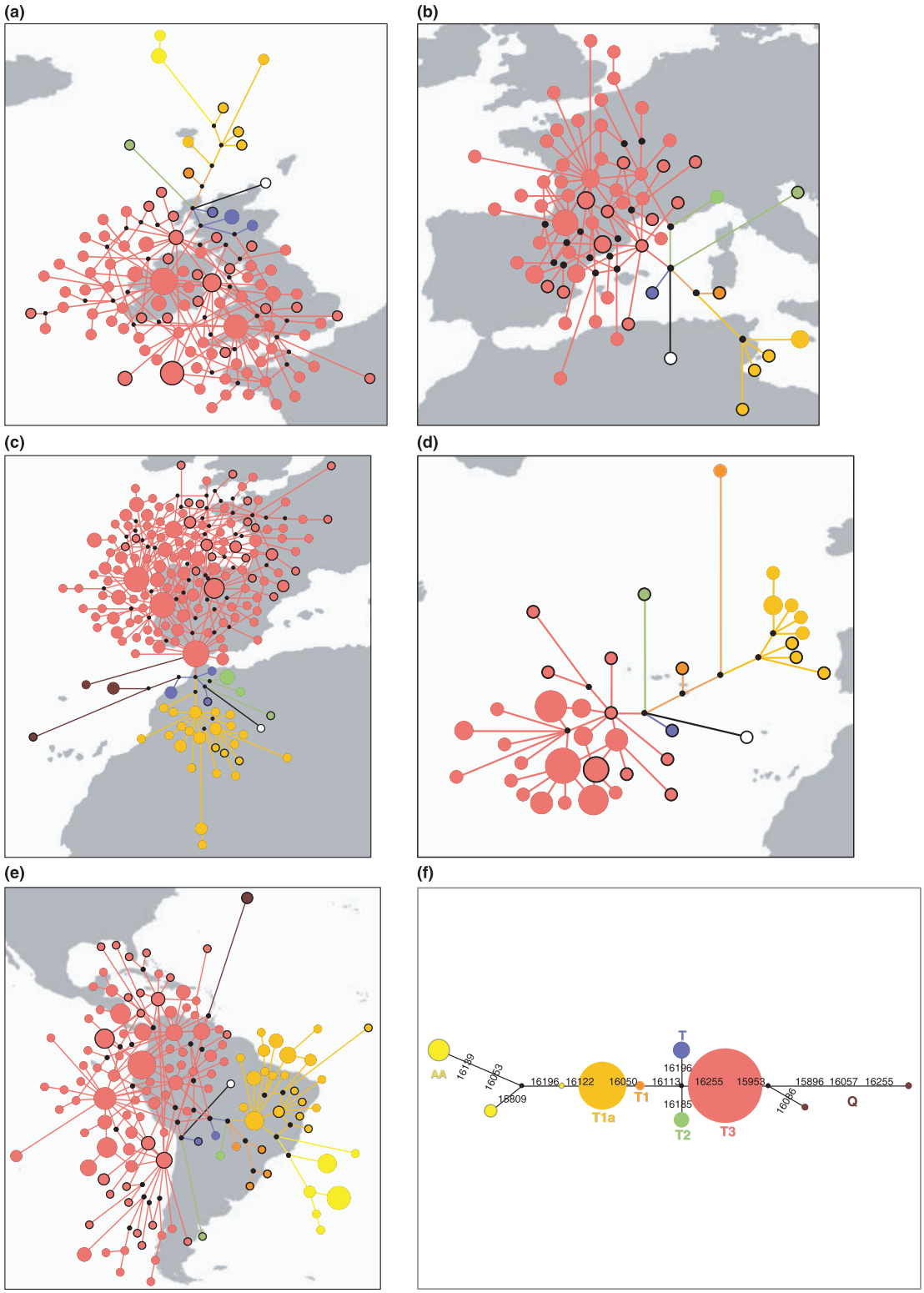

Figure 1 Reduced-median networks of mtDNA haplotypes observed in cattle from: (a) British Islands, (b) Continental Europe, (c) Iberian Peninsula, (d) Atlantic Islands, (e) Americas (Creoles and Zebu), and (f) network of haplogroups across regions. The networks were reduced at the following positions: 16049 and 16057 (a), 16057, 16131, 16133, and 16248 (b), 15947, 16057, 16122 and 351 (c), 16057, 16131, 16133 and 16302 (d) and 16057 (e and f). Haplogroups are colour coded as follows: T, purple; T1, orange; T1a, light orange; AA, yellow; T2, green; T3, red; Q, brown. Circle sizes are proportional to sequence frequency, reference sequences are denoted by bold circles and theoretical median vectors are represented by black dots.

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Table 2 Y chromosome haplotypes defined by five SNPs, one indel and seven STRs and their overall frequency (*N*).

		SNPs					*Indel*	STRs						
Haplotype	N	*DDX3_Y1*	*DDX3_Y7*	*UTY_19*[2]	*ZFY_9*	*ZFY_10*	*ZFY_10*	*DDX3_Y1*	*BM861*	*INRA124*	*INRA189*	*UMN0103*	*UMN0307*	*UMN0504*
H1Y2	0.005	C	C	A	C	C	GT	249	158	132	86	132	149	144
H2Y1	0.011	C	C	C	C	C	–	249	158	132	98	114	151	144
H3Y1	0.131	C	C	C	C	C	–	249	158	132	98	124	151	144
H4Y1	0.067	C	C	C	C	C	–	249	158	132	98	124	155	144
H5Y1	0.001	C	C	C	C	C	–	249	158	132	98	132	149	144
H6Y2	0.227	C	C	A	C	C	GT	249	158	132	102	132	149	144
H7Y2	0.008	C	C	A	C	C	GT	249	158	132	104	124	149	144
H8Y2	0.029	C	C	A	C	C	GT	249	158	132	106	132	149	144
H9Y3	0.024	T	T	A	T	T	GT	245	156	130	88	116/122[1]	149	144
H10Y2	0.001	C	C	A	C	C	GT	249	158	132	104	132	135	144
H11Y2	0.148	C	C	A	C	C	GT	249	158	132	104	132	149	144
H12Y2	0.024	C	C	A	C	C	GT	249	158	132	104	132	151	144
H13Y2	0.001	C	C	A	C	C	GT	249	158	132	102	138	149	144
H14Y1	0.128	C	C	C	C	C	–	249	158	132	100	124	155	144
H15Y1	0.001	C	C	C	C	C	–	249	158	132	100	124	151	144
H16Y2	0.004	C	C	A	C	C	GT	249	158	132	104	136	149	144
H17Y2	0.027	C	C	A	C	C	GT	249	158	132	104	128	149	144
H18Y3	0.053	T	T	A	T	T	GT	245	156	130	88	116/124[1]	149	144
H19Y3	0.001	T	T	A	T	T	GT	245	156	130	88	118/122[1]	149	144
H20Y3	0.001	T	T	A	T	T	GT	245	156	130	88	118/124[1]	149	144
H21Y3	0.001	T	T	A	T	T	GT	245	156	130	88	122/130[1]	149	144
H22Y3	0.064	T	T	A	T	T	GT	245	156	130	90	114/124[1]	151	144
H23Y1	0.033	C	C	C	C	C	–	249	158	132	102	124	151	144
H24Y2	0.005	C	C	A	C	C	GT	249	158	132	102	130	149	144

[1]Two fragments amplified in *Bos indicus*. [2]Results from the pyrosequencing assay (G/T) were converted to match Gotherstrom *et al.* 2005.

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Table 3 Y chromosome haplotype frequency and diversity ($H \pm SD$) within breeds.

	Haplotype																										
Breed	H2 Y1	H3 Y1	H4 Y1	H5 Y1	H14 Y1	H15 Y1	H23 Y1	H1 Y2	H6 Y2	H7 Y2	H8 Y2	H10 Y2	H11 Y2	H12 Y2	H13 Y2	H16 Y2	H17 Y2	H24 Y2	H9 Y3	H18 Y3	H19 Y3	H20 Y3	H21 Y3	H22 Y3	N	*H*	SD
Arouquesa									0.77	0.19			0.03												31	0.374	0.092
Barrosã									0.12				0.36	0.52											33	0.606	0.049
Cachena									0.04				0.92	0.04											25	0.157	0.096
Marinhoa									1.00																17	0.000	0.000
Maronesa													1.00												23	0.000	0.000
Minhota									1.00																28	0.000	0.000
Mirandesa									1.00																23	0.000	0.000
Alentejana													1.00												31	0.000	0.000
Brava de Lide		0.04		0.04				0.12	0.77				0.04												26	0.406	0.116
Garvonesa									1.00																6	0.000	0.000
Mertolenga	0.41								0.12			0.06	0.41												17	0.684	0.070
Preta	0.03								0.03		0.76		0.17												29	0.406	0.101
Mostrenca		1.00																							21	0.000	0.000
Iberian Peninsula																									*310*	*0.712*	*0.016*
Ramo Grande		0.28	0.72																						18	0.425	0.099
Canaria			0.50				0.07						0.43												14	0.604	0.076
Palmera							0.96		0.04																25	0.080	0.072
Atlantic Islands																									*57*	*0.677*	*0.037*
Argentino1		0.06	0.06		0.06				0.61									0.22							18	0.601	0.113
Argentino2																			0.83	0.17					6	0.333	0.215
Caracú																								1.00	48	0.000	0.000
Chiapas		0.08	0.17						0.17								0.08		0.33	0.08			0.08		12	0.879	0.075
Chihuahua									0.25											0.25	0.25	0.25			4	1.000	0.177
Pampa Chaqueño					1.00																				25	0.000	0.000
Texas Longhorn		0.02			0.56				0.02										0.04	0.36					45	0.575	0.049
Creoles																									*158*	*0.779*	*0.019*
Brahman																			1.00						7	0.000	0.000
Gyr																				1.00					21	0.000	0.000
Zebu																									*28*	*0.389*	*0.084*
Angus		1.00																							41	0.000	0.000
British White		0.76				0.05							0.05			0.14									21	0.414	0.124
Hereford					1.00																				45	0.000	0.000
Jersey													0.05				0.95								20	0.100	0.088
Shorthorn		1.00																							11	0.000	0.000
British Islands																									*138*	*0.636*	*0.025*
Charolais								0.08	0.92																13	0.154	0.126
Friesian			1.00																						27	0.000	0.000
Limousin									0.94						0.06										17	0.118	0.101
Continental Europe																									*57*	*0.543*	*0.026*

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

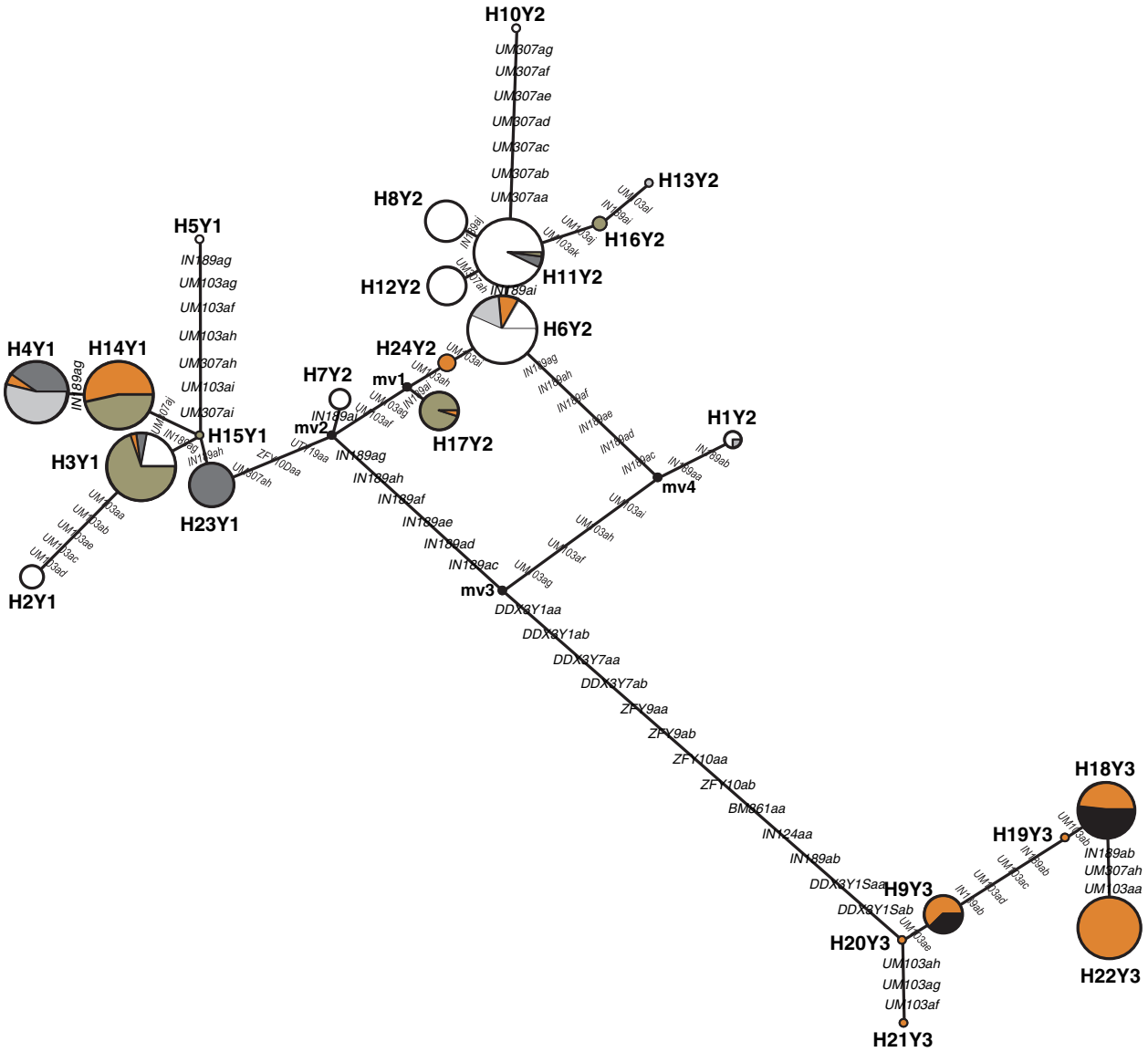

Figure 2 Median-joining network constructed with Y-haplotypes from 748 bulls representing 33 cattle populations. Circle sizes are proportional to the haplotype frequency. Breed groups were defined based on geographical origins and are represented according to the following: Iberian Peninsula (white), Atlantic Islands (dark grey), Creoles (orange), Zebu (black), British Islands (green) and Continental Europe (light grey). Theoretical median vectors are represented by black dots.

this breed group. Creole cattle had a heterogeneous composition with Y1, Y2 and Y3 haplogroups. Introgression from *B. indicus* was detected among all Creoles except Argentino1 and Pampa Chaqueño. British cattle were mainly of the Y1 haplogroup, with H3Y1 haplotype fixed in Angus and Shorthorns and H14Y1 fixed in Hereford. Jersey was an exception, with H17Y2 in high frequency. The Continental European breeds belonged to the Y2 haplogroup, except for Friesian, which was fixed for H4Y1.

Y chromosome differentiation and AMOVA

Pairwise F_{ST} values for Y-haplotypes (Table S6) ranged from nearly zero to one, and were mostly significant ($P < 0.05$). Because of their sharing of H3Y1, Mostrenca, Angus and Shorthorn were not significantly differentiated ($P > 0.05$). We found non-significant ($P > 0.05$) pairwise F_{ST} values among Argentino1 and the Portuguese breeds Brava de Lide ($F_{ST} = 0.050$) and Garvonesa ($F_{ST} = 0.103$). Based on average F_{ST} values, Caracú ($F_{ST} = 0.905 \pm 0.180$) was the most differentiated breed, whereas Chiapas was the least ($F_{ST} = 0.444 \pm 0.212$). Mexican Creoles from Chihuahua and Chiapas were not significantly differentiated ($F_{ST} = 0.010$, $P > 0.050$). Pampa Chaqueño was not significantly ($P > 0.05$) differentiated from Hereford. Mean pairwise F_{ST} among Creoles was 0.626 ± 0.183, similar to that found for the Iberian Peninsula (0.647 ± 0.113), Atlantic Islands (0.671 ± 0.144) and Continental European

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Table 4 Frequency (%) of mtDNA and Y chromosome haplogroups, and of African Y-alleles per breed and geographical group.

	mtDNA							Y chromosome				
Breeds	T1	T1a	AA	T2	T3	Q	*N*	Y1	Y2	Y3	African alleles	*N*
Arouquesa		12.5			81.2	6.3	16		100		22.6	31
Barrosã		6.3			93.8		16		100		87.9	33
Cachena		25.0			75.0		16		100		96.0	25
Marinhoa				18.8	81.3		16		100			17
Maronesa					100		16		100		100	23
Minhota		26.7			73.3		15		100			28
Mirandesa		6.3			93.8		16		100			23
Alentejana		31.3			62.5	6.3	16		100		100	31
Brava de Lide		12.5			87.5		16	7.7	92.3		3.8	26
Garvonesa		18.8			81.3		16		100			6
Mertolenga		6.3		6.3	87.5		16	41.2	58.8		47.1	17
Preta					100		16	3.4	96.6		17.2	29
Mostrenca		18.8			81.3		16	100				21
Iberian Peninsula		*12.6*		*1.9*	*84.5*	*1.0*	*207*	*10.0*	*90.0*		*41.3*	*310*
Ramo Grande		18.8			81.3		16	100				18
Canaria	6.7	20.0			73.3		15	57.1	42.9		42.9	14
Palmera					100		14	96.0	4.0			25
Atlantic Islands	*2.2*	*13.3*			*84.4*		*45*	*87.7*	*12.3*		*10.5*	*57*
Argentino1&2		4.3			95.7		23	12.5	62.5	25.0		24
Baja California		5.0	25.0		70.0		20					
Caracú		10.0	40.0		50.0		10			100	100	48
Chiapas		20.0			80.0		15	25.0	25.0	50.0	8.3	12
Chihuahua	5.3	5.3	10.5	5.3	73.6		19		25.0	75.0		4
Poblano		25.0			75.0		16					
Nayarit		25.0			75.0		16					
Pampa Chaqueño		6.3	43.7		50.0		16	100				25
Texas Longhorn		37.5			62.5		16	57.8	2.2	40.0		45
Creoles	*0.7*	*14.6*	*11.9*	*0.7*	*72.1*		*151*	*36.1*	*12.7*	*51.3*	*31.0*	*158*
Brahman		30.0			70.0		20			100		7
Gyr					100		9			100		21
Zebu		*20.7*			*79.3*		*29*			*100*		*28*
Angus		7.4			92.6		27	100				41
British White					100		10	81.0	19.0		19.0	21
Hereford			12.0		88.0		25	100				45
Jersey					100		18		100		100	20
Shorthorns					100		9	100				11
British Islands		*2.3*	*3.4*		*94.3*		*89*	*82.6*	*17.4*		*17.4*	*138*
Charolais		12.5		6.3	81.3		16		100			13
Friesian					100		16	100				27
Limousin					100		16		100			17
Continental Europe		*4.2*		*2.1*	*93.8*		*48*	*47.4*	*52.6*			*57*
Overall	*0.3*	*11.3*	*3.7*	*1.1*	*83.3*	*0.3*	*569*	*37.3*	*48.1*	*14.6*	*27.7*	*748*

(0.688 ± 0.145) breed groups. British breeds were slightly more differentiated (average F_{ST} = 0.752 ± 0.081) and the Zebu group the most differentiated (average F_{ST} = 0.830 ± 0.046).

We detected significant ($P < 0.0001$) differences among breeds based on the AMOVA analysis of the Y-haplotype data, with approximately 62% of the total genetic variation found between breeds within groups and 22% corresponding to individual differences (Table S4). Breed grouping based on geographical origin had a significant effect ($P < 0.001$) and accounted for approximately 16% of the total genetic variation.

Y chromosome phylogenetics

Relationships among Y-haplotypes and their breed distributions are depicted in the MJ network of Fig. 2. Haplotypes H2Y1, H3Y1, H4Y1, H5Y1, H14Y1, H15Y1 and H23Y1 were separated from H6Y2, H8Y2, H10Y2, H11Y2, H12Y2, H13Y2, H16Y2, H17Y2 and H24Y2 by two

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

median vectors forming two clusters corresponding to Y1 and Y2 haplogroups respectively. The Iberian H7Y2 from Arouquesa had an intermediate position and was separated from the Y1 cluster by only one median vector. H1Y2 clustered within the Y2 group and was separated from Y1 haplotypes by three median vectors. The *B. indicus* Y3 haplogroup was clearly separated from the *B. taurus* Y1 and Y2 clusters. Creole cattle shared H3Y1, H4Y1and H14Y1, H6Y2 and H17Y2 with several *B. taurus* breeds. With the exception of Argentino1 and Pampa Chaqueño, Creoles also clustered within the Y3 haplogroup together with Brahman and Gyr. The newly detected H22Y3 of Caracú had a more distant position in the network and may represent a more ancestral haplotype.

A summary of mtDNA and Y-haplogroups is shown in Table 4. African influence (mtDNA T1 and T1a; African Y-alleles) was detected in most Iberian and Atlantic Islands breeds. In Creoles, African matrilines (T1, T1a or AA) were detected in all populations, whereas African patrilines were only found in Caracú and Chiapas. In Continental European and British breeds, African matrilines were detected in Charolais (T1a), Angus and Hereford from Argentina (T1a and AA respectively), whereas African patrilines were found only in British White and Jersey.

Discussion

Creole cattle showed high genetic diversity of mtDNA and Y chromosome haplotypes, several of which were exclusive of this breed group. The Y-diversity found in the Iberian Peninsula and Creoles was higher compared with the more intensively selected European breeds, which were fixed or nearly so for a Y-haplotype. Our results did not reflect a founder effect from the dispersion of a few Iberian animals introduced in the Americas, but rather the heterogeneous genetic makeup of Creoles. The previous studies have also found high levels of diversity within Creole cattle (Giovambattista *et al.* 2001; Carvajal-Carmona *et al.* 2003). Creoles represent important reservoirs of cattle genetic diversity and should be the subject of conservation measures.

Signatures of Iberian ancestry in Creole cattle were inferred from mtDNA and Y- haplotypes. Among the matrilines found in Creoles, the common European T3 haplogroup was prevalent (≈72%) and most of the haplotypes shared exclusively with the Iberian Peninsula and/or the Atlantic Islands were of this group. The haplotypes shared with animals from the Atlantic Islands are consistent with the use of these Islands as intermediate ports between the Iberian Peninsula and the Americas (Rouse 1977). Iberian cattle had a high incidence of African T1a matrilines, which were also found in all Creole populations with an overall frequency of ≈ 15%. The more divergent African-derived AA lineage was found in Baja California, Caracú, Chihuahua and Pampa Chaqueño, and accounted for ≈12% of the Creole haplogroups. Although we did not detect the AA lineage in the Iberian samples that we analysed, this matriline has been found in the Spanish Retinta (Miretti *et al.* 2004) and in one de Lidia individual (Cortes *et al.* 2008). Moreover, three Alentejana animals (one Hap083 and two Hap301) which clustered with the Creole AA haplogroup in the network of all African-derived matrilines could represent old matrilines present in Iberia during the colonization period. Iberian origin of Creole patrilines was more difficult to infer but could be reflected by the presence of H3Y1 and H6Y2. H3Y1 was found in Mostrenca (fixed), Brava de Lide and Ramo Grande, but this haplotype was also fixed in Angus and Shorthorn, and was in high frequency in British White. Thus it is not possible to determine if the presence of H3Y1 in Creoles reflects Iberian ancestry, or more recent crossbreeding with British bulls, which is known to have occurred in the late 19th century (Rouse 1977). Interestingly, the finding that Mostrenca belongs to the Y1 haplogroup, which based on the analysis of ancient DNA was frequent in aurochsen (Gotherstrom *et al.* 2005), appears to indicate the contribution of wild bulls to this breed. Mostrenca is a semi-feral breed from the Doñana region and is kept isolated from other cattle populations (Rodero *et al.* 1992; Martinez *et al.* 2005). H6Y2 is the prevalent patriline in the Iberian Peninsula and is also found in Creoles from Argentina, Mexico and USA. The analysis of Y-haplotypes in other Spanish breeds thought to have been brought to the Americas, such as Retinta and Rubia Galega, would help clarify the Iberian origin of Creole patrilines.

Whereas mtDNA lineages reflected the genetic proximity among Creoles, Iberian and Atlantic Islands breeds, male-mediated *B. indicus* introgression was evident in most Creole populations except for Argentino1 and Pampa Chaqueño, and accounted for ≈51% of Y-haplogroups. No *B. indicus* matrilines were found in the Creoles, which suggested that the first cattle introduced in the Americas were strictly *B. taurus* and that crossbreeding with zebu bulls occurred later (Meirelles *et al.* 1999; Miretti *et al.* 2004). Interestingly, the Brahman and Gyr animals that we analysed also had taurine mtDNA sequences. These results are in agreement with other studies that have reported phenotypically zebuine animals with taurine mtDNA (Miretti *et al.* 2002; Mirol *et al.* 2003). Troy *et al.* (2001) attributed the absence of zebuine mtDNA in African humped cattle to ancestral crossbreeding with *B. taurus*.

The *B. indicus* patriline H22Y3 detected exclusively in Caracú, and fixed in this breed, supports the hypothesis of introduction of bulls directly from Africa during colonial times. H22Y3 carries the *INRA189-90* allele, which was also detected in the West-African taurine N'Dama (Edwards *et al.* 2000). Hanotte *et al.* (2000) found zebu patrilines in taurine African cattle, including the N'Dama from Guinea-Bissau, which was a Portuguese colony. Cattle were introduced in Brazil by Portuguese settlers (Primo 1992), and Portugal had colonies in West Africa, and the islands of Cape Verde and S. Tomé e Principe, since the 15th century

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

until 1974. It is possible that cattle from Africa may have been introduced in Brazil through slave trade routes (Rouse 1977). To determine the exact origin of the Caracú patrilines, it will be necessary to analyse additional samples from the African continent, as well as other zebu breeds from Brazil. The contribution of African bulls can also be inferred through the detection of the *INRA189-104* allele as discussed by Ginja *et al.* (2009), and this allele was found in Y-haplotypes of Canaria (H11Y2), Chiapas (H17Y2), British White (H11Y2 & H16Y2) and Jersey (H17Y2) breeds. Magee *et al.* (2002) used autosomal STRs to infer West African ancestry in Guadeloupe Creoles of the Caribbean. Therefore, the introduction of African cattle directly into the Americas appears to have occurred at least in Brazil and the Caribbean. The African influence detected in British White and Jersey breeds could be related to their putative Roman origin (Rouse 1970). The common H11Y2 haplotype was detected in animals from both breeds, as well as a patriline specific of each breed (H16Y2 and H17Y02 respectively). These haplotypes have the *INRA189-104* allele and differ by allele variation at locus *UMN0103*. The analysis of autosomal STRs has also shown a close genetic relationship among Jersey and African N'Dama cattle from Guinea (Cymbron *et al.* 2005). Even though our results suggest African influence in these two breeds, a more complete analysis of European cattle is needed to investigate the origin and diversity of patrilines.

We report a wider dispersal of the ancestral African-derived AA matriline, which is found in Brazil, Paraguay and as far north as in Baja California. The detection of this lineage in Caracú (40%), as well as in three other Creole cattle from Brazil (Miretti *et al.* 2002), and its absence in Portugal lend support to the hypothesis of direct introduction of African cattle, although AA haplotypes have yet to be found in Africa (Liron *et al.* 2006). Miretti *et al.* (2002, 2004) attributed the presence of the AA haplogroup in Caribbean Creoles to their Spanish ancestry but did not exclude the possibility that it existed in Portuguese cattle brought to Brazil during colonial times and disappeared since then. Sequencing additional mitochondrial regions could help to better characterize the AA haplogroup. Although variation across the control region has provided valuable information regarding the genetic origins of cattle, complete mtDNA sequences might be needed to determine more accurately the phylogeny of maternal haplogroups (Achilli *et al.* 2008; Hiendleder *et al.* 2008).

Our results showed that Creoles retain genetic signatures of their Iberian ancestry, but also confirmed more recent introgression of zebu and British breeds. The high incidence of African-derived lineages in Creoles is related to their Iberian origin, but direct contribution of cattle from West Africa cannot be excluded. Y-haplotypes were valuable to better understand the origins of Creoles and to complement mtDNA information, particularly regarding the contributions of African cattle. Nonetheless, more comprehensive studies of Spanish and African breeds with mtDNA and Y chromosome markers would be helpful to further clarify their contributions to the Creole cattle of the Americas.

Acknowledgements

We gratefully acknowledge the collaboration of the Portuguese breed associations and Red Iberoamericana sobre la conservación de la biodiversidad de animales domésticos locales para el desarollo rural sostenible (Red CYTED XII-H) for the sampling of animals. The establishment of the DNA bank was supported by Direcção Geral de Veterinária and Sociedade Portuguesa de Recursos Genéticos Animais. C. Ginja was supported by a grant from the Fundação para a Ciência e a Tecnologia (Ref. SFRH/BD/13502/2003). This work was partially funded by the Veterinary Genetics Laboratory, University of California, Davis. We thank Semex Alliance (Guelph, Canada) for providing six semen samples used in this study.

References

Achilli A., Olivieri A., Pellecchia M. *et al.* (2008) Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Current Biology* **18**, 157–8.

Anderson S., de Bruijn M.H., Coulson A.R., Eperon I.C., Sanger F. & Young I.G. (1982) Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. *Journal of Molecular Biology* **156**, 683–717.

Anderung C., Bouwman A., Persson P., Carretero J.M., Ortega A.I., Elburg R., Smith C., Arsuaga J.L., Ellegren H. & Gotherstrom A. (2005) Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8431–5.

Anderung C., Hellborg L., Seddon J., Hanotte O. & Gotherstrom A. (2007) Investigation of X- and Y-specific single nucleotide polymorphisms in taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle. *Animal Genetics* **38**, 595–600.

Bandelt H.J., Forster P. & Rohl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37–48.

Beja-Pereira A., Caramelli D., Lalueza-Fox C. *et al.* (2006) The origin of European cattle: evidence from modern and ancient DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8113–8.

Bollongino R., Elsner J., Vigne J.D. & Burger J. (2008) Y-SNPs do not indicate hybridisation between European aurochs and domestic cattle. *PLoS ONE* **3**, e3418, doi: 10.1371/journal.pone.0003418.

Bradley D.G., MacHugh D.E., Loftus R.T., Sow R.S., Hoste C.H. & Cunningham E.P. (1994) Zebu-taurine variation in Y chromosomal DNA: a sensitive assay for genetic introgression in west African trypanotolerant cattle populations. *Animal Genetics* **25**, 7–12.

Bruford M.W., Bradley D.G. & Luikart G. (2003) DNA markers reveal the complexity of livestock domestication. *Nature Reviews Genetics* **4**, 900–10.

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Carvajal-Carmona L.G., Bermudez N., Olivera-Angel M., Estrada L., Ossa J., Bedoya G. & Ruiz-Linares A. (2003) Abundant mtDNA diversity and ancestral admixture in Colombian criollo cattle (*Bos taurus*). *Genetics* **165,** 1457–63.

Cortes O., Tupac-Yupanqui I., Dunner S., Garcia-Atance M.A., Garcia D., Fernandez J. & Canon J. (2008) Ancestral matrilineages and mitochondrial DNA diversity of the Lidia cattle breed. *Animal Genetics* **39,** 649–54.

Cymbron T., Loftus R.T., Malheiro M.I. & Bradley D.G. (1999) Mitochondrial sequence variation suggests an African influence in Portuguese cattle. *Proceedings of the Royal Society of London, Series B, Biological sciences* **266,** 597–603.

Cymbron T., Freeman A.R., Isabel Malheiro M., Vigne J.D. & Bradley D.G. (2005) Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations. *Proceedings of the Royal Society of London, Series B, Biological sciences* **272,** 1837–43.

Edwards C.J., Gaillard C., Bradley D.G. & MacHugh D.E. (2000) Y-specific microsatellite polymorphisms in a range of bovid species. *Animal Genetics* **31,** 127–30.

Excoffier L., Smouse P.E. & Quattro J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131,** 479–91.

Ginja C., Gama L.T. & Penedo M.C.T. (2009) Y chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *Journal of Heredity* **100,** 148–57.

Giovambattista G., Ripoli M.V., De Luca J.C., Mirol P.M., Liron J.P. & Dulout F.N. (2000) Male-mediated introgression of *Bos indicus* genes into Argentine and Bolivian Creole cattle breeds. *Animal Genetics* **31,** 302–5.

Giovambattista G., Ripoli M.V., Peral-Garcia P. & Bouzat J.L. (2001) Indigenous domestic breeds as reservoirs of genetic diversity: the Argentinean Creole cattle. *Animal Genetics* **32,** 240–7.

Gotherstrom A., Anderung C., Hellborg L., Elburg R., Smith C., Bradley D.G. & Ellegren H. (2005) Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society of London, Series B, Biological sciences* **272,** 2345–50.

Hall T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41,** 95–8.

Hanotte O., Tawah C.L., Bradley D.G., Okomo M., Verjee Y., Ochieng J. & Rege J.E. (2000) Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-saharan African cattle breeds. *Molecular Ecology* **9,** 387–96.

Hasegawa M., Kishino H. & Yano T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22,** 160–74.

Hiendleder S., Lewalski H. & Janke A. (2008) Complete mitochondrial genomes of *Bos taurus* and *Bos indicus* provide new insights into intra-species variation, taxonomy and domestication. *Cytogenetic and Genome Research* **120,** 150–6.

Lai S.J., Liu Y.P., Liu Y.X., Li X.W. & Yao Y.G. (2006) Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Molecular Phylogenetics and Evolution* **38,** 146–54.

Liron J.P., Bravi C.M., Mirol P.M., Peral-Garcia P. & Giovambattista G. (2006) African matrilineages in American Creole cattle: evidence of two independent continental sources. *Animal Genetics* **37,** 379–82.

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

MacHugh D.E., Troy C.S., McCormick F., Olsaker I., Eythorsdottir E. & Bradley D.G. (1999) Early medieval cattle remains from a Scandinavian settlement in Dublin: genetic analysis and comparison with extant breeds. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **354,** 99–108.

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. *Journal of Heredity* **93,** 429–32.

Mannen H., Kohno M., Nagata Y. *et al.* (2004) Independent mitochondrial origin and historical genetic differentiation in North Eastern Asian cattle. *Molecular Phylogenetics and Evolution* **32,** 539–44.

Martinez A.M., Calderon J., Camacho E., Rico C., Vega-Pla J.L. & Delgado J.V. (2005) Genetic characterization of the Mostrenca cattle with microsatellites. *Archivos Zootecnia* **54,** 357–61.

Meadows J.R., Hanotte O., Drogemuller C., Calvo J., Godfrey R., Coltman D., Maddox J.F., Marzanov N., Kantanen J. & Kijas J.W. (2006) Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* **37,** 444–53.

Meirelles F.V., Rosa A.J.M., Lobo R.B., Garcia J.M., Smith L.C. & Duarte F.A.M. (1999) Is the American Zebu really *Bos indicus*? *Genetics and Molecular Biology* **22,** 543–6.

Miretti M.M., Pereira H.A., Jr, Poli M.A., Contel E.P. & Ferro J.A. (2002) African-derived mitochondria in South American native cattle breeds (*Bos taurus*): evidence of a new taurine mitochondrial lineage. *Journal of Heredity* **93,** 323–30.

Miretti M.M., Dunner S., Naves M., Contel E.P. & Ferro J.A. (2004) Predominant African-derived mtDNA in Caribbean and Brazilian Creole cattle is also found in Spanish cattle (*Bos taurus*). *Journal of Heredity* **95,** 450–3.

Mirol P.M., Giovambattista G., Liron J.P. & Dulout F.N. (2003) African and European mitochondrial haplotypes in South American Creole cattle. *Heredity* **91,** 248–54.

Nei M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, USA.

Peakall R. & Smouse P.E. (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6,** 288–95.

Posada D. & Crandall K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14,** 817–8.

Posada D. & Crandall K.A. (2001) Selecting the best-fit model of nucleotide substitution. *Systematic Biology* **50,** 580–601.

Primo A.T. (1992) El ganado bovino Iberico en las Americas: 500 anos despues. *Archivos Zootecnia* **41,** 421–32.

Rodero A., Delgado J.V. & Rodero E. (1992) Primitive Andalusian livestock and their implications in the discovery of America. *Archivos Zootecnia* **41,** 383–400.

Rouse J.E. (1970) *World Cattle*. University of Oklahoma Press, USA.

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x

Rouse J.E. (1977) *The Criollo: Spanish Cattle in the Americas*. NORMAN, University of Oklahoma Press, USA.

Schneider S., Roessli D. & Excoffier L. (2000) *Arlequin ver. 2.000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Tajima F. (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105,** 437–60.

Tamura K. & Nei M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10,** 512–26.

Troy C.S., MacHugh D.E., Bailey J.F., Magee D.A., Loftus R.T., Cunningham P., Chamberlain A.T., Sykes B.C. & Bradley D.G. (2001) Genetic evidence for Near-Eastern origins of European cattle. *Nature* **410,** 1088–91.

Supporting information

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

Figure S1 Reduced-median network of African-derived mtDNA haplotypes observed in cattle across the six regions studied. Haplogroups are colour coded as follows: T1, orange; T1a, light orange; AA, yellow. The mutation sites that differentiate these haplogroups are shown in red. Circle sizes are proportional to haplotype frequency, reference sequences are denoted by bold circles and theoretical median vectors are represented by black dots.

Table S1 Sequence variation in the control region of 569 cattle samples aligned with the *Bos taurus* reference sequence (Accession number V00654). A total of 311 haplotypes were identified (only variable nt shown). Dashes (–) indicate gaps and dots (.) denote identities with the reference sequence. Haplotype frequency (N) and the respective haplogroup are shown in the last two columns.

Table S2 Frequency of the 311 mtDNA haplotypes in each population of the six geographic groups.

Table S3 Pairwise F_{ST} values between populations for mtDNA sequences. Significant values ($P < 0.05$) are shown in bold.

Table S4 AMOVA results for mtDNA and Y-haplotypes.

Table S5 Reference mtDNA sequences used in the phylogenetic analysis.

Table S6 Pairwise F_{ST} values between populations for Y-haplotypes. Non-significant values ($P > 0.05$) are shown in bold.

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.

© 2009 The Authors, Journal compilation © 2009 Stichting International Foundation for Animal Genetics, *Animal Genetics*
doi:10.1111/j.1365-2052.2009.01976.x