www.nature.com/hdy

ORIGINAL ARTICLE

Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*)

J Kantanen[1], CJ Edwards[2], DG Bradley[2], H Viinalass[3], S Thessler[4], Z Ivanova[5], T Kiselyova[6], M Ćinkulov[7], R Popov[8], S Stojanović[9], I Ammosov[10] and J Vilkki[1]

[1]Biotechnology and Food Research, MTT Agrifood Research Finland, Jokioinen, Finland; [2]Smurfit Institute of Genetics, Trinity College, Dublin, Ireland; [3]Institute of Veterinary Medicine and Animal Sciences, Estonian University of Life Sciences, Tartu, Estonia; [4]Information Management Unit, MTT Agrifood Research Finland, Helsinki, Finland; [5]Yakut State Agricultural Academy, Yakutsk, Sakha, Russia; [6]Department of Genetics and Biotechnology, All-Russian Research Institute for Farm Animal Genetics and Breeding, St Petersburgh-Pushkin, Russia; [7]Animal Science Department, University of Novi Sad, Novi Sad, Serbia; [8]Yakutian Research Institute of Agriculture, Yakutsk, Sakha, Russia; [9]Division of Agro-Resources, Ministry of Agriculture and Water Management, Beograd, Serbia and [10]Batagay-Alyta, Sakha, Russia

Maternally inherited mitochondrial DNA (mtDNA) has been used extensively to determine origin and diversity of taurine cattle (*Bos taurus*) but global surveys of paternally inherited Y-chromosome diversity are lacking. Here, we provide mtDNA information on previously uncharacterised Eurasian breeds and present the most comprehensive Y-chromosomal microsatellite data on domestic cattle to date. The mitochondrial haplogroup T3 was the most frequent, whereas T4 was detected only in the Yakutian cattle from Siberia. The mtDNA data indicates that the Ukrainian and Central Asian regions are zones where hybrids between taurine and zebu (*B. indicus*) cattle have existed. This zebu influence appears to have subsequently spread into southern and southeastern European breeds. The most common Y-chromosomal microsatellite haplotype, termed here as H11, showed an elevated frequency in the Eurasian sample set compared with that detected in Near Eastern and Anatolian breeds. The taurine Y-chromosomal microsatellite haplotypes were found to be structured in a network according to the Y-haplogroups Y1 and Y2. These data do not support the recent hypothesis on the origin of Y1 from the local European hybridization of cattle with male aurochsen. Compared with mtDNA, the intensive culling of breeding males and male-mediated crossbreeding of locally raised native breeds has accelerated loss of Y-chromosomal variation in domestic cattle, and affected the contribution of genetic drift to diversity. In conclusion, to maintain diversity, breeds showing rare Y-haplotypes should be prioritised in the conservation of cattle genetic resources.
Heredity (2009) **103**, 404–415; doi:10.1038/hdy.2009.68; published online 15 July 2009

Introduction

Analyses of mitochondrial DNA (mtDNA) D-loop sequences and Y-chromosome-specific polymorphisms (such as single nucleotide polymorphisms, insertion–deletion mutations and microsatellites) have indicated that humpless taurine (*Bos taurus*) and humped zebu cattle (*B. indicus*) have clearly distinguishable mtDNA and Y-chromosomal haplotypic profiles (Loftus *et al.*, 1994; Bradley *et al.*, 1996; Hanotte *et al.*, 1997; Mannen *et al.*, 2004; Götherström *et al.*, 2005; Li *et al.*, 2007b). This observation points towards two independent domestication events from genetically differentiated aurochs (*B. primigenius*) populations for the two basic taxa of domestic cattle. The modern European and northern Asian domestic cattle are of humpless taurine type and descend from the aurochs populations domesticated 10 000 years ago in the Near Eastern region (Troy *et al.*,

2001; Edwards *et al.*, 2007a). However, in some areas of the Eurasian continent, phenotypically humpless cattle are known to have been influenced by historical admixture from zebu cattle. One of such cattle breeds is the Mongolian cattle, which belong to the so-called Turano-Mongolian type of taurine cattle (Mannen *et al.*, 2004; Lai *et al.*, 2006). Turano-Mongolian cattle are typically found in Eastern Asia and can be distinguished from the European taurine cattle by their characteristic cranial formation and the shape of the horns (Felius, 1995).

MtDNA D-loop sequences of the existing taurine cattle breeds are typically converged into five mtDNA haplogroups: T, T1, T2, T3 and T4 (Troy *et al.*, 2001; Mannen *et al.*, 2004 but see also; Achilli *et al.*, 2008). In the Eurasian continent, the groups T, T1, T2 and T3 have been found in modern Near Eastern cattle breeds, whereas T, T2 and T3 appear in European breeds (Troy *et al.*, 2001). Compared with the modern Near Eastern taurine cattle, European cattle show a higher frequency of T3 and much lower frequencies of T and T2 (Troy *et al.*, 2001). In East Asia, T and T1 have not been detected in the Turano-Mongolian cattle, whereas T4 has so far only been found in Asia (Mannen *et al.*, 2004; Lai *et al.*, 2006). In their recent whole mitochondrial genomic study,

Achilli *et al*. (2008) suggested that T4-carrying mt genomes are derived from T3-carrying genomes, which were captured during domestication in the Near East.

Frequency differences in alleles at Y-chromosomal microsatellites and in Y-chromosomal microsatellite haplotypes have been found in cattle (Edwards *et al*., 2000; Li *et al*., 2007b), suggesting their applicability for genetic diversity studies of domestic cattle. In addition to the more recent evolutionary events, Y-chromosomal microsatellites may have a potential to identify deeper lineages of Y-chromosomes (Bosch *et al*., 1999). In taurine cattle, to date two Y-chromosomal haplogroups Y1 and Y2 have been defined by sequence polymorphisms. The haplogroup Y1 dominates in modern northern and western European breeds, but has not yet been detected in Near Eastern samples (Götherström *et al*., 2005). This observation led Götherström *et al*. to conclude that the Y1 haplogroup was introduced to northern and western European taurine cattle by prehistoric backcrossing with local aurochs bulls, rather than descending from the domesticated Near Eastern aurochs population.

Our current knowledge of mtDNA variation of taurine cattle in Eurasia is based on studies of European and Near Eastern breeds (Loftus *et al*., 1994; Bradley *et al*., 1996; Troy *et al*., 2001). In recent years, a few East Asian breeds have also been examined (Mannen *et al*., 2004; Lai *et al*., 2006). Global surveys on Y-chromosomal polymorphisms in domestic cattle have not been published. In addition, no previous study has covered the vast area of the Russian Federation, and the distribution of taurine mtDNA haplogroups in the Baltic Sea region is also poorly known. The South of the Russian Federation borders archaeological domestication sites from where animals spread across the Russian territory (Epstein and Mason, 1984; Tapio *et al*., 2006b). The Eurasian continent is inhabited by a large number of unique cattle breeds, like Busha cattle in the Balkan region, Eastern Finncattle in Northern Europe and Yakutian cattle in the Sahka Republic of the Russian Federation. On the basis of autosomal microsatellite data, these breeds have shown genetic divergence from modern commercial West European breeds, such as Black-and-White Holstein-Friesian and Red-and-White Ayrshire cattle (Li *et al*., 2007a). Yakutian cattle belong to the Turano-Mongolian breeds and have adapted to extreme cold climatic conditions (Li *et al*., 2007a; Granberg *et al*., 2009).

Here, we examine the taurine cattle breeds of eastern, northern and southeastern Europe and Central and East Asia for mtDNA D-loop sequence variation. In addition to the mtDNA data, we present a first global survey on Y-chromosomal microsatellite haplotypes in taurine cattle using a large data set of cattle breeds originating from, or having paternal roots, in Europe, the Near East, Central Asia and Eastern Siberia.

Materials and methods

Animal samples
The present mtDNA D-loop sequence and Y-chromosomal marker analyses were based on a representative data set of different indigenous and international taurine cattle breeds originating from eastern, northern and southeastern Europe, Central Asia and Eastern Siberia (Supplementary Information, Supplementary Table S1).

A total of 268 male and female individuals from 34 cattle breeds were examined for mtDNA D-loop sequence polymorphisms: 235 animals from Europe and 33 from Central and East Asia. A total of 405 bulls, representing 54 cattle breeds, were assessed for five Y-chromosomal markers: 343 bulls from Europe, 27 from Central and East Asia, and 35 from the Near East and Anatolia. The breeds examined were from Byelorussia, Denmark, Estonia, Faroe Islands, Finland, Germany, Iraq, Iceland, Kazakhstan, Latvia, Lithuania, Norway, Poland, the Russian Federation, Serbia, Sweden, Syria, Turkey, Ukraine and Uzbekistan (Supplementary Table S1).

For the analysis of geographical patterns of mtDNA diversity, published taurine mtDNA D-loop sequences of breeds originating from northern Europe (Icelandic cattle, Norwegian Red, Telemark and Western Fjord cattle, $n = 37$), UK (Jersey, $n = 7$) and the Near East and Anatolia (Anatolian Black, East Anatolian Red, South Anatolian Red, Turkish Grey, Damascus, Iraqi and Kurdi breeds, $n = 35$) were included (Troy *et al*., 2001). With the addition of this GenBank data, it was possible to achieve a more global mtDNA dataset, which corresponded geographically with the sampling regions of the Y-chromosome data.

Laboratory analysis
The DNA samples of Byelorussian, Ukrainian, Russian, Central Asian and Serbian breeds were extracted from either 6–10 ml blood samples or semen samples by standard protocols, involving either organic extraction or salt precipitation (Miller *et al*., 1988; Zadworny and Kuhnlein, 1990). For the other breeds, DNA was available from previous cattle genetic diversity studies (Bradley *et al*., 1996; Edwards *et al*., 2000; Tapio *et al*., 2006a).

The mtDNA D-loop region was amplified using the primers published in Cymbron *et al*. (1999). A 25-µl PCR reaction mixture included 20 pmol of each primer, 200 µM of each dNTP, DynaZyme buffer (Finnzymes, Helsinki, Finland), 100 ng of DNA template and 1 U DynaZyme II DNA polymerase (Finnzymes). A thermo cycling protocol of 5 min at 94 °C, followed by 30 cycles of 1 min at 94 °C, 1 min at 55 °C, 2 min at 72 °C and a final extension step (72 °C for 5 min) was applied. Amplified DNA products of 375 bp were purified with ExoSAP-IT method (Amersham Biosciences, UK). Standard double-stranded sequencing was performed with DYEnamic ET Terminator Kit (Amersham Biosciences) using 10 µl of purified PCR-product and the PCR-primers on a Mega-BACE 500 DNA Sequencer (Amersham Biosciences). An additional PCR amplification and double-stranded sequencing analysis was performed to read a longer sequence from the five individuals, which showed nucleotide substitutions T>C and G>A at the mtDNA positions 16042 and 16093, respectively, with the forward primer 5′GCC CAT ACA CAG ACC ACA GA3′ and reverse primer 5′CAA GCA TCC CCC AAA ATA AA3′, to confirm that their mtDNA haplotypes could be assigned to haplogroup T4. The sequence analysis was performed using the SEQUENCHER v.4.6 software (Gene Codes Co, Ann Arbor, MI, USA).

The primer sequences for the Y-chromosomal microsatellites *INRA124*, *INRA189*, *BM861* and *BYM-1* were reported in Edwards *et al*. (2000) and Ward *et al*. (2001). The *DYZ-1* marker (Perret *et al*., 1990; Bradley *et al*., 1994)

was amplified using forward primer 5′CCT GGC GAC TGT GCA ATA TT3′ and reverse primer 5′CAC ACA CAC AAC CGG TTT CT3′. *INRA124*, *INRA189* and *BM861* were typed as described in Edwards *et al.* (2000). The markers *BYM-1* and *DYZ-1* were PCR-amplified with annealing temperatures of 66 and 60 °C, respectively, and analysed on an ALF-express Automated Sequencer (Pharmacia, Sweden). The consistency of the size of different alleles was assured by using an internal size standard (Sizer, Pharmacia, Sweden) and samples for which allele sizes had been previously determined by Edwards *et al.* (2000). No amplification product was obtained when using DNA from females (Edwards *et al.*, 2000).

Statistical analyses

Multiple mtDNA D-loop sequence alignments were performed using the Clustal X package (Thompson *et al.*, 1997). The size of the aligned D-loop fragment was 255 bp between the nucleotide positions 16021 and 16275 in relation to the entire taurine mtDNA sequence (GenBank accession number V00654) for the sequences PCR-amplified and sequenced with the primer pairs described by Cymbron *et al.* (1999), and 353 bp between the nucleotide positions 15987 and 16339 for the Yakutian cattle sequences examined with the additional primer pairs. Haplotype diversity (H), nucleotide diversity (π) and the mean number of differences between haplotypes were calculated with DnaSP v.3 (Rozas and Rozas, 1999, available at http://www.ub.es/dnasp/). The molecular relationships between the mtDNA haplotypes were studied by reduced median (RM) network analysis (NETWORK 4.1.1.1, available at http://www.fluxus-technology.com/; Bandelt *et al.*, 1995). The same program was used for the analysis of mtDNA sequence mismatch distributions as substitutional differences between pairs of haplotypes. The neighbor-joining (NJ) method (Saitou and Nei, 1987) was applied to infer the phylogenetic relationships among zebu and the T, T2 and T4 haplogroups, and the five most frequent T3 mtDNA haplotypes, from a pairwise matrix of the distances based on Kimura (1980) two-parameter model. The degree of support for phylogenetic tree branches was assessed by 1000 bootstrap replications. The phylogenetic analysis was performed with TREECON v1.3 program (Van De Peer and De Wachter, 1994).

Allele frequencies for Y-chromosomal loci were determined by standard gene counting. We compared interpopulation F_{ST} estimates of the Y-chromosomal microsatellites corrected for the four-times smaller effective population size (Pérez-Lezaun *et al.*, 1997), with F_{ST} estimates of 20 autosomal microsatellites. In the F_{ST} calculations computed with FSTAT1.2 (Goudet, 1995), the data set comprised the 40 taurine breeds analysed here for the Y-chromosomal markers, plus previously published autosomal microsatellite data from Li *et al.* (2007a) and Tapio *et al.* (2006a). After the marker-based analyses, the allelic combinations of Y-chromosomal markers were recorded as haplotypes, and the number of haplotypes and haplotypic diversity (Nei, 1987) were calculated. Mismatch distribution (allelic differences between pairs of haplotypes) and a median-joining (MJ) network were constructed using the program NETWORK 4.1.1.1. An input file for the MJ algorithm was obtained by the

RM algorithm of the program (Bandelt *et al.*, 1995, 1999). By this approach, phylogenetically unrealistic cycles within the network were reduced as suggested by Forster *et al.* (2000). A weighting scheme was used to compensate for genetic variance differences at the five microsatellites so that loci with the highest variances were given the lowest weights. The following weights were assigned: *INRA189* = 7, *DYZ-1* and *BYM1* = 8, *INRA124* and *BM861* = 10. Otherwise, program default settings were applied in the network construction.

The analysis of geographic patterns in mtDNA and Y-chromosome diversity was based on *PhiPT* distances between the haplotypes computed with GENALEX v.6 program (Peakall and Smouse, 2006). Any haplotype sequence or allelic combination comparison with the same nucleotide at the same nucleotide position, or allelic state at the same locus yields a value of 0, while any comparison of different nucleotides or allelic states yields a value of 1 (Huff *et al.*, 1993). When calculated across multiple positions for a given pair of samples, the *PhiPT* distance is equivalent to the tally of differences between two genetic profiles (Peakall and Smouse, 2006). Statistical significance of the pairwise distances was tested by random permutations, with the number of permutations set to 9999. A 38- and 46-dimensional data of population *PhiPT* genetic distances for mtDNA and Y-chromosomal data, respectively, was summarised into three dimensions and visualized effectively by ordination. First, principal co-ordinate (PCO) analyses of population *PhiPT* matrices for both the mtDNA and Y-chromosome data sets were performed with the GENALEX v.6. Next, the geographic patterns of inter-breed diversity were visualised by interpolating and mapping the scores of the first three coordinates of the PCO analysis. With the interpolation, we wanted to visualize the geographic patterns, not to estimate axes scores between the sampling points. Interpolation was conducted using minimum curvature method, which fits a two-dimensional spline function on the given sampling points and yields the smoothest possible surface for irregularly located data points (Briggs, 1974; Smith and Wessel, 1990). Minimum curvature interpolation was run in TNT Mips v 6.9, which first applied two-dimensional cubic spline function to the sampling point values and then smoothed the surface between sampling points. This initial phase was followed by iteration that smoothed the surface. The pixel size of the produced raster layer was defined to 20×20 km and large water areas were masked out. The parameterization detail is given in Supplementary Information.

To visualize the results of the geographic analysis, all the three grid layers (PCO axes) were mapped as a single red-green-blue (RGB) image, where scores from axis 1 were shown in red, scores from axis 2 in green and scores from axis 3 in blue (for interpretation, see the RGB image in Supplementary Information, Supplementary Figure S1). The RGB image summarizes a spatial variation of genetic distances in a single map showing the areas that have short genetic distances in similar colours. The map may be interpreted using the RGB colour cube (Supplementary Information, Supplementary Figures S1 and S2), in which three axes of the colour cube (red, green and blue) correspond to the three ordination axes, and the intensity of the colours increase as axes scores increase.

In the geographic analysis of mtDNA diversity, the taurine mtDNA data sequenced in this study was aligned with taurine GenBank sequence data. The calculation of genetic distances was based on an mtDNA fragment of 240 bp of size spanning from mtDNA nucleotide 16023–16262. The GenBank mtDNA data of Near Eastern and Anatolian breeds was pooled and the breeds were considered as one 'Near Eastern' population. Similarly, the taurine data from the two Central Asian breeds (Ala-Tau and Bushuev) were combined. Because of their remote location, the Yakutian cattle from East Siberia were excluded from the analysis, and so calculations were based on 38 populations. For the identical Y-chromosomal analyses, similar pooling and exclusion of breeds were conducted and the calculations were based on 46 populations.

In the spatial analysis, different latitude and longitude coordinates were used in mtDNA and Y-chromosomal analysis for Norwegian Red, Finnish Ayrshire and Finnish Holstein-Friesian breeds. These commercial breeds have their paternal origins in Scottish Ayrshire and Friesian (western and northwestern Europe, respectively) cattle populations, whereas their maternal ancestries are assumed to be mainly of Norwegian and Finnish indigenous cattle origin. Similarly, the paternal origin of the Lithuanian and Polish Holstein-Friesian bulls is from northwestern Europe, and that of Estonian Red, Danish Red in Latvia and Lithuanian Red from Denmark. The Mantel Test implemented in GENALEX was computed to examine the relationship between the geographic and genetic distances.

Results

MtDNA variation
A total of 60 mtDNA haplotypes from 268 sequences of 255 bp D-loop sequence fragment were defined by 49 transitions, two transversions and two 1 bp insertion–deletion mutations, demonstrating a typical strong transitional bias (Figure 1). We assigned the data to 58 taurine and two zebu haplotypes by inspecting the sequence correspondence with the taurine and zebu reference sequences (accession numbers V00654 and L27733, respectively). The nucleotide sequence of the haplotype H3 is identical with the taurine reference sequence, and the sequence of the haplotype H2 with the indicus reference sequence. The average number of nucleotide differences (gaps excluded) between the taurine and zebu haplotypes was 24.03.

Zebu mtDNA haplotypes were found in two Ukrainian Whitehead, two Ala-Tau and two Bushuev animals. The zebu mtDNA is assigned to two haplogroups I1 and I2 (Lai et al., 2006; Achilli et al., 2008). In our data set, the H1 haplotype was assigned to I2 and the H2 to I1. The taurine mtDNA haplotypes were grouped into four different haplogroups T, T2, T3 and T4 on the basis of the nucleotide differences at defining positions (Troy et al., 2001; Mannen et al., 2004). A total of 49 taurine haplotypes were assigned to the T3, three to the T (a defining nucleotide position 16255), four to the T2 (defining positions 16057 [G>C transversion], 16185 and 16255) and two to the T4 (defining positions 16042, 16093 and 16302) (Figure 1). The transitional mutation G>A on the nucleotide position 16302 was inspected in

five Yakutian cattle animals with additional sequencing as described in Materials and methods. T3 haplotypes were found across all study regions, whereas T haplotypes were only found in Serbia (at frequency 5/19), in the European part of the Russian Federation (2/58) and in Sweden (1/53). T2 haplotypes were found in Ukraine (3/18), Finland (2/47), Estonia (1/10) and in the Asian part of the Russian Federation (in the Sakha Republic 1/24), and T4 haplotypes only in Sakha (5/24) (Figure 1).

One zebu and 45 taurine haplotypes were breed specific, of which 36 taurine haplotypes were detected only once (62% of the taurine haplotypes). The zebu and taurine private haplotypes were distributed across 23 breeds.

The haplotypic diversity for the entire data set was 0.746 ± 0.029 and the respective nucleotide diversity value was 0.99%. Within each breed with at least five individuals analysed, between 2 (Danish Red) to 8 (Suksun) different haplotypes were found (on average 3.6 haplotypes within one breed). The intrabreed haplotypic diversity varied from 0 (Ukrainian Grey) to 0.956 (Suksun), being on average 0.603 within a breed. More detailed information on within-breed diversity estimates, including mean number of differences between haplotypes and nucleotide diversity, is given in Supplementary Table S1.

Y-chromosomal variation
At the Y-chromosomal markers, we detected two alleles at INRA124, three at BM861, four each at DYZ1 and BYM-1 and eight at INRA189 (Supplementary Table S2). The low level of diversity at the markers INRA124 and BM861 may be due to either a low mutation rate or some form of selection affecting the allelic diversity of the markers. The INRA189-allele sizes showed departure from the ladder-like distribution found at the other tested Y-chromosomal markers: the detected lengths of INRA189 alleles were in base pairs 82, 88, 90, 98, 100, 102, 104 and 106. The allele INRA189[82 bp] was found in six individuals from the two Norwegian indigenous breeds (the Telemark and Doela cattle), the INRA189[88 bp] allele, which is an indicus-specific allele (Edwards et al., 2000), was found in nine bulls from Near Eastern and Central Asian breeds, and the INRA189[90 bp] allele in seven individuals of the Ukrainian Grey and Iraqi breeds. The typing of the taurine bulls displaying an exceptional short allele length at INRA189 was repeated three times.

The taurine data gave locus-wise F_{ST} estimates ranging from 0.733 (INRA189) to 0.848 (DYZ1) [INRA124 and BM861 were not considered in the calculation], with an average of 0.787, which reduced to 0.477 when a correction for the smaller effective size of the Y-chromosomes (Pérez-Lezaun et al., 1997) was applied. In previous studies for autosomal markers (Tapio et al., 2006a; Li et al., 2007a), the mean F_{ST} for autosomal data was 0.0948.

Y-chromosomes were classified to be taurine ($n = 395$) or zebu ($n = 10$) based on the detection of zebu- and taurine-specific alleles INRA124[130 bp] and INRA124[132 bp], respectively (Table 1; Edwards et al., 2000; Hanotte et al., 2000) and the zebu-specific alleles BM861[156 bp] and INRA189[88 bp] (Edwards et al., 2000). A total of 26 haplotypes were observed among 405 bulls tested (Table 1). The data included five zebu and 21 taurine

```
                                   1111111111111111111111111111111111111111111111
                                   6666666666666666666666666666666666666666666666
                                   0000000000000000011111111111111111111112222222222
                                   2444556778888899900011111222333333344489001224455 6
                                   2029782461234523428903679126013578913756009127805 0
                                                                                    i
                                                              Frequency    Breed(s)
            H3        GATCGCATATGACTAGTGTTCTTGTGTTTTTTTTCTATGGG-ACCCCATC   132          in 27 breeds, all study area
            H1(zebu)  A..TAT....A..C....A.C.CCACAC.C...CC.C-C.A.AG.TTT...   2            Ala-Tau
            H2(zebu)  A..TAT.C..A.T....A.C.CCACAC.C...CC..-C.A.AG..TT...   4            Ukr Whiteheaded, Bushuev
            H4  (T)   ...........................................-......C.   5            Podolian
            H5  (T)   .........................................A.-......C.   2            Yaroslavl
            H6  (T)   .................................C...-....T.C.         1            Swedish Holstein-Fr
            H7  (T2)  ....C....................................A..-.....C.   2            Ukrain Whiteheaded
            H8  (T2)  ....C..........................C......A..-.....C.     3            Finn Ayrsh, West Finnc, Est Red
            H9  (T2)  ....C...........T.........C......A..-.....C.          1            Yakut
            H10 (T2)  ....C....................................A..-.T.T..C.  1            Ukrain Whiteheaded
            H11 (T4)  ..C...........A..................-.......             4            Yakut
            H12 (T4)  .GC...........A..................-.......             1            Yakut
            H13       .................................A-.......            1            Ala-Tau
            H14       A...........G....................-.......            1            Finn Ayrsh
            H15       ......G..........C..C............-.......            1            Faroe Islands
            H16       .........................C......-...T....            5            Faroe Islands, Swed Red-White
            H17       ...AT...........................-.......            2            Finn Holst-Fr
            H18       ...A............................-.....G..            3            East Finn, North Finn
            H19       ...............C..C..............-.......            1            Yakut
            H20       ..C.............C...............-.......            1            Yakut
            H21       ........................C.......-.......            3            Yakut, North Finn
            H22       ...T............................-.......            7            West Finn, Kholmogory, Suksun
            H23       .............A..................-.......            1            West Finn
            H24       ...A............................-.......            4            NorthFinn,SwedRedPoll,Jutland,Bushuev
            H25       ..C.............................-.......            5            NorthFinn, SwedHolsFr, Jutland,ByeloR
            H26       ..C..T........................-.T......               1            Ukr Whiteheaded
            H27       .............C......C...........-.T......            1            Kholmogory
            H28       ..................T.............-.......            2            Ukr Whiteheaded
            H29       ......C.........................-...G..               1            Kholmogory
            H30       .........G......................-.......            5            Kholmogory, Pechora, Suksun
            H31       ..................C.............-.......            1            Swed Mountain
            H32       ....T...........................-.......            1            Swed Mountain
            H33       ................T...............-.......            27           in 10 breeds from Swed,Den,Rus,Estonia
            H34       ...............................-.T......            1            Fjal
            H35       ......C........T.....G....-.......                    1            Fjal
            H36       ...........T....C...............-.......            1            Fjal
            H37       .......................C.....-...G..                  1            Fjal
            H38       ......C.........................-.......            2            Ringamala, ByeloR
            H39       ......C.........T.........G...A-......               4            Vane cattle
            H40       ...............C.......C.........-.......            1            Vane cattle
            H41       ...T...................C.........-.......            1            Pechora
            H42       ...A.........A..................-.......            1            Istoben
            H43       ......C........T................-.......            2            Istoben, Yurino
            H44       ........................T.......-.......            2            Istoben, Bushuev
            H45       ..........G.............C.......-.......            1            Istoben
            H46       ..........C.............C.......-.......            1            Istoben
            H47       ......C.G.......................-.......            1            Suksun
            H48       ...............................-.......T             1            Suksun
            H49       ......C.....C..........C........-.......            2            Suksun
            H50       ...............................-...T....            1            Suksun
            H51       ...........................-..T.....                  1            Jutland
            H52       ......C........T...............-.......T             1            Jutland
            H53       ........G...............T.....-.......              1            Red Gorbatov
            H54       ...A........................A-.......               1            Byelorussian Red
            H55       .............C..................-.......            1            Byelorussian Red
            H56       ...A...................C........-.......            3            Busa
            H57       ...........C..........C.........-.......            1            Podolian
            H58       ..........T.....................-.......            1            Podolian
            H59       .G..............................-.......            1            Estonian Holst-Fr
            H60       ......C.....C...................-.......            1            Estonian Red
```

Figure 1 Sequence variation in 60 cattle D-loop sequences. Haplotype codes and haplogroups in parenthesis (zebu or taurine haplogroups T, T2 and T4) are given in the first column. If a haplogroup has not been indicated, the haplotype belongs to the taurine haplogroup T3. In addition, the frequencies of the haplotypes and the distribution among breeds are shown. The haplotype H3 sequence is identical to the taurine reference haplotype sequence (accession number V00654) and the sequence from H2 to the zebu reference sequence (L27733).

haplotypes. The haplotype termed H11 ($INRA124^{132\,bp}$-$INRA189^{98\,bp}$-$BM861^{158\,bp}$-$DYZ1^{366\,bp}$-BYM-$1^{258\,bp}$) was the most frequent and found in 262 bulls. The haplotype H23 ($INRA124^{132\,bp}$-$INRA189^{104\,bp}$-$BM861^{158\,bp}$-$DYZ1^{362\,bp}$-BYM-$1^{256\,bp}$) was observed in 38 bulls and the haplotype H16 ($INRA124^{132\,bp}$-$INRA189^{102\,bp}$-$BM861^{158\,bp}$-$DYZ1^{360\,bp}$-BYM-$1^{258\,bp}$) in 31 bulls. In all, 15 haplotypes (all five

zebu and 10 taurine haplotypes) were private (Table 1). Eleven private haplotypes were detected only once (42% of the total number of haplotypes). The Near Eastern sample lacked nine taurine haplotypes found elsewhere in the study area.

The haplotypic diversity calculated for the entire Y-data set was 0.816 ± 0.0155. The haplotypic diversity

Table 1 Frequencies of the Y-chromosomal haplotypes in the data of 405 bulls and the list of breeds sharing each haplotype

Haplotype	Allelic combination	Frequency	Breed(s)
H1 (zebu)	130-88-156-360-254	3	Iraqi
H2 (zebu)	130-88-156-360-256	1	Iraqi
H3 (zebu)	130-88-156-360-258	4	Bushuev
H4 (zebu)	130-88-158-360-254	1	Iraqi
H5 (zebu)	130-90-158-360-254	1	Iraqi
H6	132-82-158-362-256	6	Doela, Telemark
H7	132-90-158-360-256	6	Ukrainian Grey
H8	132-98-158-360-258	1	Ukrainian Whiteheaded
H9	132-98-158-364-258	1	Anatolian Black
H10	132-98-158-366-256	7	Finnish Holstein-Friesian, Polish Holstein-Friesian, Yaroslavl, Swedish Mountain, Lithuanian Grey
H11	132-98-158-366-258	262	All, except Jersey, Simmental, Ala-Tau, Doela, Telemark, Red Gorbatov, Yurino, Yakutian, Busa, Podolian, Anatolian Black, East Anatolian Red, Ukrainian Grey
H12	132-98-160-366-258	1	Iraqi
H13	132-100-158-362-254	1	Damascus
H14	132-100-158-362-256	3	Icelandic, Anatolian Black
H15	132-100-158-366-258	7	Western Fjord, Western Red Polled, Estonian Native, Ukrainian Whiteheaded
H16	132-102-158-360-258	31	Ala-Tau, Ukrainian Whiteheaded, Red Gorbatov, Yurino, Busa, Byelorussian Red, East Anatolian Red, Simmental, Suksun
H17	132-102-158-362-256	1	Anatolian Black
H18	132-102-158-362-258	6	Danish Red in Latvia, Lithuanian Red
H19	132-102-158-366-252	4	Eastern Red Polled, Fjall cattle, Swedish Red Polled
H20	132-104-158-360-256	1	South Anatolian Red
H21	132-104-158-360-258	1	Ukrainian Grey
H22	132-104-158-362-254	8	Eastern Finncattle
H23	132-104-158-362-256	38	Jersey, Podolian, Ala-Tau, Anatolian Black, East Anatolian Red, Yakutian
H24	132-106-158-360-256	4	Damascus, South Anatolian Red, Turkish Grey
H25	132-106-158-362-256	5	Anatolian Black, Damascus, East Anatolian Red, South Anatolian Red, Turkish Grey
H26	132-106-158-362-258	1	Turkish Grey

The allelic combinations of haplotypes are given in the marker order of INRA124, INRA189, BM861, DYZ1 and BYM-1. The haplotypes H1–H5 are zebu- and H6–H26 taurine specific.

in the Near Eastern taurine data (28 chromosomes) was 0.958 ± 0.020 and in the Eurasian taurine data (367 chromosomes) 0.786 ± 0.017. When inspecting genetic variation within breeds (at least five bulls tested), the haplotypic diversity varied from 0 to 1.0 (on average 0.625), with between one and four different haplotypes segregating within each breed (Supplementary Table S1). As expected, in the most diverse breeds in terms of haplotype numbers and haplotypic diversity, both zebu and taurine haplotypes were present.

We tested whether the frequencies of the three most common Y-haplotypes, H11, H23 and H16, in the Eurasian sample set of 367 taurine chromosomes deviated significantly from the Near Eastern-Anatolian sample set of 28 taurine chromosomes. The frequencies of haplotype H11 for the Eurasian and Near Eastern-Anatolian taurine cattle were 0.69 and 0.32, respectively. The χ^2 test indicated significantly different haplotype frequencies between these two regions ($\chi^2 = 16.3$, $df = 1$, $P < 0.001$), whereas no significant differences in the frequencies of haplotype H23 (the respective frequencies 0.09 and 0.14; $\chi^2 = 0.75$, $df = 1$, $P = 0.39$) or of the haplotype H16 (the respective frequencies 0.08 and 0.03; $\chi^2 = 0.76$, $df = 1$, $P = 0.38$) were detected.

MtDNA haplotype network and mismatch distribution

Phylogenetic relationships of 60 mtDNA haplotypes were explored with a RM network (Figure 2). As expected from the sequence alignment (Figure 1), two deeply divergent mtDNA lineages, zebu (H1 and H2 haplotypes) and taurine (the rest of the haplotypes), were

detected. The haplogroup T3 formed a star-like phylogeny of haplotypes, with one major haplotype H3. This pattern is typical for populations having experienced a population expansion (Slatkin and Hudson, 1991). In an NJ tree (Supplementary Figure S3), nodes separating the two zebu haplotypes (H1 and H2) from taurine haplotypes were supported by bootstrap value of 100%, and nodes clustering T4 (H11 and H12) and T2 (H7–H10) haplotypes from the rest of the taurine haplotypes, were supported by relatively high bootstrap values of 91 and 73%, respectively.

Analysis of mismatch distributions (Supplementary Figure S4) gave equivalent results to the RM network. A bimodal distribution was revealed showing the deep genetic divergence between the two cattle lineages, zebu and taurine. When each taurine haplogroup was considered separately, all mismatch distributions were bell-shaped and unimodal (not shown).

Y-chromosomal haplotype network and mismatch distribution

Figure 3 depicts a MJ network relating the 26 Y-haplotypes. Unlike the mt data, Y-haplotypes did not form a tightly clustered network of closely related haplotypes originating from one monophyletic group, nor a star-shaped pattern. Instead, the network of Y-haplotypes was a sum of at least one zebu and two taurine phylogenetic haplogroups.

A total of 17 Finnish and 13 Near Eastern-Anatolian bulls included in our study were typed for unique event markers (single nucleotide polymorphisms and indels) in
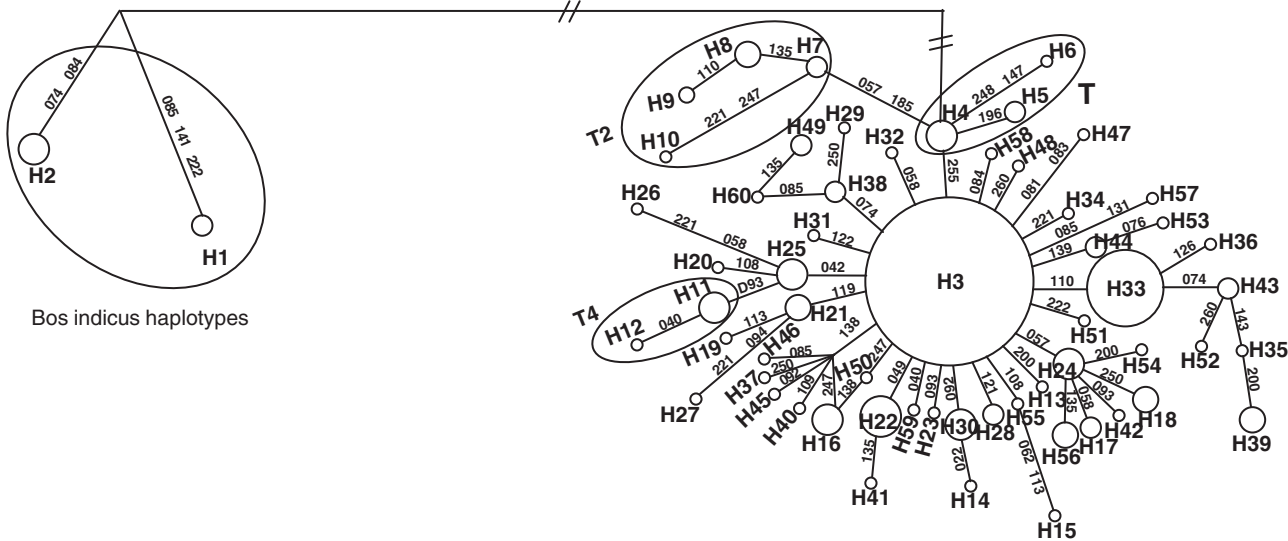
Figure 2 A RM network of mtDNA haplotypes. The sequence variations and codes of the haplotypes are as given in Figure 1. The haplogroups T, T2 and T4 and the *Bos indicus* haplotypes are marked. The rest of the haplotypes belong to the haplogroup T3. The size of the D-loop fragment is 255 bp between the nucleotide positions 16021 and 16275 in relation to the entire taurine mtDNA sequence (V00654). Circle areas are proportional to haplotype frequencies and branch lengths to number of substitutions between the sequences.
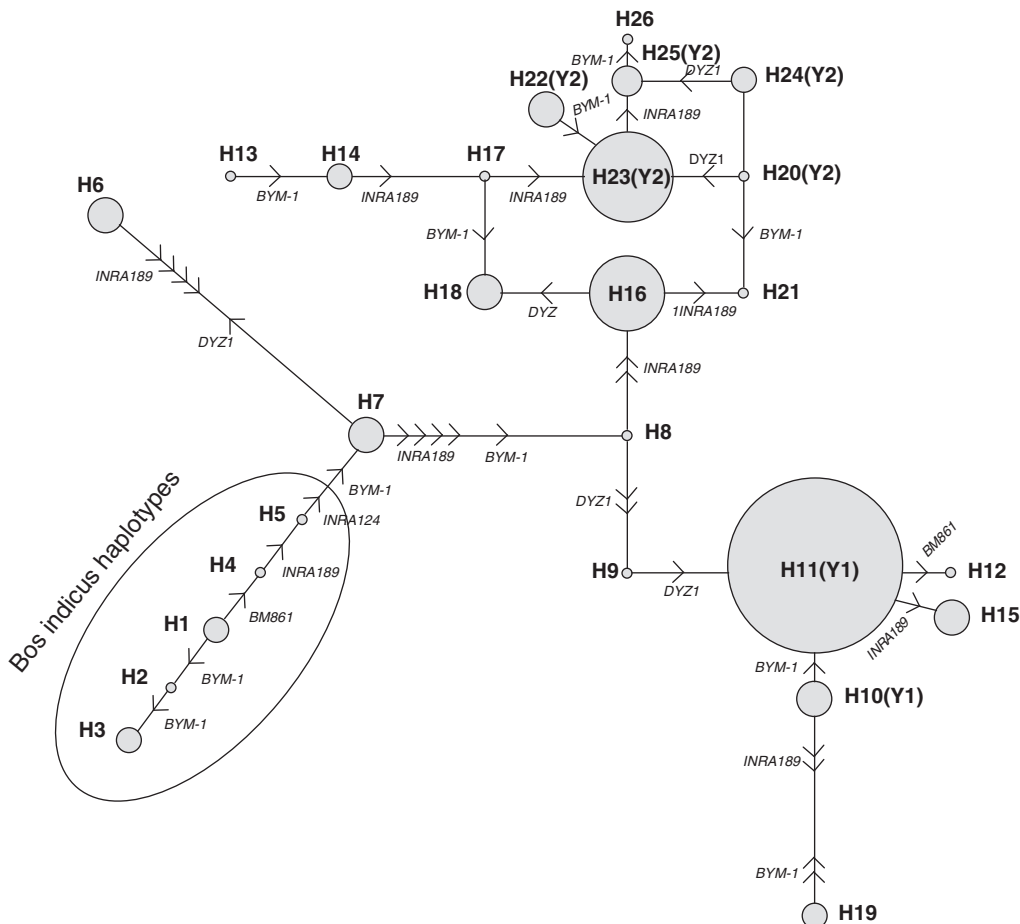
Figure 3 A MJ network relating the 26 Y-chromosomal microsatellites. The allelic combinations and codes of the haplotypes are given in Table 1. The haplogroups Y1 and Y2 are indicated if known for the haplotype. The *Bos indicus* haplotypes are shown. Arrows point in the direction of an increase in repeat number of 2 bp.

the study of Götherström *et al*. (2005), allowing us to assign their Y-microsatellite haplotypes as Y1 or Y2 groups. The microsatellite haplotypes of three Eastern Finncattle bulls, which displayed the haplotype H22, and the haplotypes of the Near Eastern bulls, which displayed the haplotypes H20, H23, H24 and H25 in our study, were assigned to the haplogroup Y2 according to Götherström *et al*. (2005), whereas bulls assigned to the haplogroup Y1 were also from Finland, and had haplotypes H10 and H11.

In addition, the Jersey, Swedish Mountain cattle, Swedish Red Polled and Simmental breeds were typed in both studies, but different bulls were examined. The Jersey breed in our data displayed the haplotype H23 and in the data of Götherström *et al*. (2005) the haplogroup Y2, whereas H10, H11 and Y1 were found in Swedish Mountain cattle, and H11, H19 and Y1 in Swedish Red Polled. The three Simmental bulls typed in our analysis had the haplotype H16, but in the study by Götherström *et al*. (2005), Simmental bulls were distributed both in Y1 and Y2. One male *B. indicus* sample typed by Götherström *et al*. (2005) displayed a phylogenetic lineage Y3. In a similar manner, the *INRA124* marker separated the zebu Y-haplotypes into their own branch in our network.

In our data set, bulls from Anatolia and the Near Eastern region carried the haplotypes H11 (haplogroup Y1), H12 and H9 (within one mutational step from the H11, very probably haplogroup Y1), and the H20, H23, H24 and H25 haplotypes (all these haplotypes could be assigned to the haplogroup Y2). The Near Eastern breeds showed also the H17, H14 and H13 haplotypes, which were found within 1–3 mutational steps from the H23 (Y2) and H16. It appears that present distribution of the microsatellite Y-haplotypes in the network agrees with the grouping of cattle Y-haplogroups on the basis of unique event markers. However, the haplotypes H6 and H7, found in two Norwegian native breeds (Doela and Telemark) and Ukrainian Grey, respectively, were clearly differentiated from the other taurine haplotypes due to their exceptional *INRA189* alleles. The allelic combination of the haplotype H6, when *INRA189* was excluded, was identical to haplotypes H14, H17, H23 (Y2) and H25 (Y2), and the allelic combination of haplotype H7 (excluding the *INRA189* marker) was identical to H20 (Y2) and H24 (Y2).

The frequency distributions of pairwise mismatches of Y-chromosomal microsatellite haplotype data are shown in Supplementary Figure S5. The mismatch distribution exhibited a bimodal distribution reflecting the existence of the two taurine lineages Y1 and Y2 in our microsatellite data, with a smooth 'third peak' observed in the mismatch distribution due to zebu-taurine differences. The average number of mismatch differences was 4.020.

Geographical patterning of mtDNA and Y-chromosomal diversity

Among the 38 breeds (including two 'pooled breeds' and GenBank data), the average pairwise *PhiPT* genetic distance for the mtDNA data was 0.171, and 40.4% of the pairwise comparisons (284/770) yielded *PhiPT* estimates that were significantly different from zero ($P < 0.05$). The average pairwise *PhiPT* genetic distance calculated from the Y-chromosomal data was 0.299 and

41.2% of the pairwise comparisons (426/1035) yielded *PhiPT* estimates that were significantly different from zero ($P < 0.05$). Three-dimensional PCO analysis was used to condense information of pairwise *PhiPT* genetic distances from mtDNA and Y-chromosomal data. For the mtDNA data, the first, second and third axes accounted for 42.4, 20.9 and 13.4% of variation, respectively (together 76.7%), and for the Y-chromosomal data 48.8, 29.1 and 8.9%, respectively (together 86.8%).

The three PCO axes were visualized on maps as a single RGB image for each data set (Figure 4). Clearly, variation in geographic patterns depends on a sampling density; for instance, large areas in Central Europe were not sampled, whereas a dense sampling was carried out in north Europe. The most distantly related breeds based on the mtDNA data (Figure 4a) are the ones that are shown in the colour located in the opposite corners or sides of the colour cube (for interpretation, see Supplementary Figure S1). Thus, breeds located in the red areas in the map (Finnish Holstein-Friesian [FFR]) are distant from areas in cyan (Bohus Polled [Bohus]). The map of Y-chromosomal haplotypic data (Figure 4b) indicates that red areas of the Central Asian population (Ala-Tau and Bushuev taurine breeds [Cent]) and Ukrainian Whitehead [UWH] are differentiated from blue areas of Jersey [JER] and Podolian cattle [Podo]. Similarly, Yaroslavl cattle [YAR] and other breeds marked in magenta (Eastern Finncattle [EFC], Ukrainian Grey [UKR] and the pooled Near Eastern population [NEAR]) are distant from Busa (along with Yurino [YUR], Red Gorbatov [REG] and Simmental [SMT]) marked in dark green. Strong colour differences between geographically proximate breeds are observed in north Europe. However, a 'maternal region' shown green on the map, is seen in South Scandinavia and the yellowish colour pattern of the mtDNA data shows a close affinity between the Finnish native cattle and North Russian breeds (Figure 4a). In the Y-chromosome data, a light green colour connects a wide geographic region, that is western Europe, South Scandinavia, the Baltic countries and North Russia.

Discussion

Our study of maternal and paternal genealogy of humpless taurine cattle (*Bos taurus*) includes mtDNA data of breeds from previously uncharacterized geographical regions, and the most comprehensive Y-chromosomal data from domestic cattle to date. The distribution and frequencies of taurine mtDNA haplogroups T, T2 and T3 in the Eurasian continent were congruent with those previously presented (Troy *et al*., 2001; Mannen *et al*., 2004; Lai *et al*., 2006). In our breed set, T3 was a predominant mtDNA haplogroup across the study region, whereas the haplogroups T and T2 were rare and T1 was not detected at all among the individuals analysed.

The haplogroup T4 was found only in Yakutian cattle from Sakha in northeastern Asia. This haplogroup has so far been detected only in East Asian taurine cattle (Mannen *et al*., 2004; Lai *et al*., 2006). The frequency of T4 was 0.21 in the Yakutian cattle (Figure 1). According to Achilli *et al*. (2008), T4 has an origin from the same genetic sources as the T3 haplotypes or from a genetically closely related population of aurochs. How-
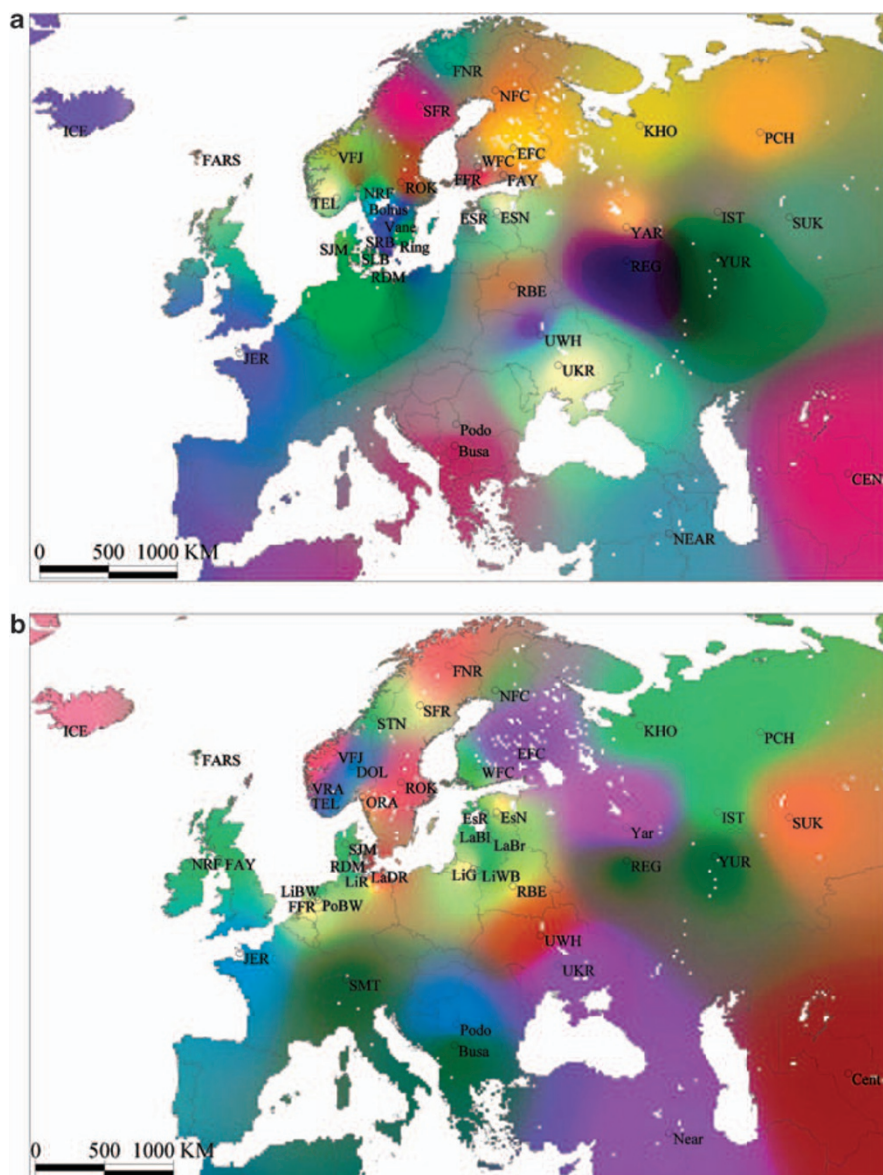
Figure 4 RGB images for the scores of three axes resulting form the PCO analyses of *PhiPT* values between population pairs (excluding the Yakutian cattle) for: (**a**) the mtDNA data and (**b**) the Y-chromosome data. The scores of axis 1 from the PCO analysis are shown in red, scores of axis 2 in green and scores of axis 3 in blue in such a way that the intensity of colour increases with increasing axis score. For interpretation of the images, see Supplementary Information. The breed codes are as follows: Bohus, Bohus Poll; BRE, Byelorussian Red; Busa, Busa; Cent, Central Asian breeds; DOL, Doela cattle; EFC, Eastern Finncattle; ESN, Estonian Native; ESR, Estonian Red; FARS, Faroe Islands cattle; FAY, Finnish Ayrshire; FFR, Finnish Holstein-Friesian; FNR, Fjall cattle; ICE, Icelandic cattle; IST, Istoben; JER, Jersey; KHO, Khlomogory; LaBl, Latvian Blue; LaBr, Latvian Brown; LaDR, Latvian Danish Red; LiBW, Lithuanian Black-and-White; LiG, Lithuanian Light Grey; LiR, Lithuanian Red; LiWB, Lithuanian White Backed; Near, Near Eastern-Anatolian breeds; NFC, Northern Finncattle; NRF, Norwegian Red cattle; ORA, Eastern Red Polled; PCH, Pechora type; PoBW, Polish Holstein-Friesian; Podo, Podolian cattle; RDM, Danish Red; REG, Red Gorbatov; Ring, Ringamala cattle; ROK, Swedish Red Polled; SFR, Swedish Mountain cattle; SJM, Jutland Breed; SLB, Swedish Holstein-Friesian; SMT, German Simmental; SRB, Swedish Red-and-White; STN, Troender cattle; SUK, Suksun; TEL, Telemark; UKR, Ukrainian Grey; UWH, Ukrainian Whitehead; Vane, Vane cattle; VFJ, Western Fjord cattle; VRA, Western Red Polled; WFC, Western Finncattle; Yar, Yaroslavl; Yur, Yurino.

ever, there is a need to conduct more intensive samplings in Central and East Asian regions to study the geographical distribution, origin and history of this haplogroup. The Yakutian cattle also show mtDNA haplotypes that can be found in European samples (Figure 1), suggesting that the Yakutian cattle have prehistoric maternal ancestries in domesticated Near Eastern cattle. The 255 bp mtDNA D-loop region sequenced here showed identity to the most common mtDNA haplotype H3 (Figure 1) in 14 Yakutian samples, and to the haplotype H21, which was also found in Northern Finncattle, in two Yakutian samples. Moreover, the singletons (H9, H19 and H20) of the haplogroups T2 and T3 detected in the Yakutian cattle were connected to non-singleton haplotypes found in European breeds (Figure 2), which also points towards a common ancestral population for the breeds (Crandall and Templeton 1993). The Y-chromosomal microsatellite

haplotype data, in turn, suggests a common prehistoric paternal ancestry for the Yakutian, Ala-Tau, Anatolian Black, East Anatolian Red, Jersey and Podolian breeds (Table 1).

Many Turano-Mongolian breeds have become extinct, like the Buryat cattle and the local Altay cattle in south Siberia, and some of the still existing breeds, for example the Japanese Black cattle and the Kazakh Whitehead, are threatened by extensive crossbreeding with international breeds (Felius, 1995). MtDNA studies have revealed a prehistorical genetic influence of zebu cattle in Mongolian cattle and several Chinese native breeds (Mannen et al., 2004; Lai et al., 2006). The Yakutian cattle, an endangered population of <600 breeding females (1200 animals in total), may therefore be considered as one of the few pure Turano-Mongolian type of breeds left globally. Currently, a multidisciplinary approach is being applied to explore the future conservation possibilities of the Yakutian cattle in the original breeding sites in the Eastern Siberian villages (Granberg et al., 2009).

In this study, the Ukrainian Whitehead and the Central Asian Ala-Tau breed displayed zebu-specific mtDNA haplotypes (Figure 1). The Near Eastern and East African regions, as well as Mongolia and China, have been recognized as areas where admixture of the two basic taxa of domestic cattle has taken place (Hanotte et al., 2002; Mannen et al., 2004; Lai et al., 2006; Edwards et al., 2007a). This study suggests that the Ukrainian and the Central Asian regions belong to hybrid zones where taurine-zebu crossbreds have existed. The admixtured nature of these breeds has not previously been reported (Dmitriev and Ernst, 1989; Felius, 1995). The indicus mtDNA haplotype found in the modern Ukrainian Whitehead cattle may descend from ancient Steppe cattle, which were upgraded with European bulls to establish the Ukrainian Whitehead breed (Dmitriev and Ernst, 1989). Similar kinds of longhorn and grey cattle are found in southeast and southern Europe, such as Maremmana, Hungarian Grey and Modicana, collectively termed as Podolian breeds (Felius, 1995). Studies of nuclear genetic markers have suggested that the genetic influence from zebu is evident in breeds of the Podolian group (Pieragostini et al., 2000; Cymbron et al., 2005). The detected genetic influence from zebu cattle in the Podolian cattle appears to originate, at least partly, from ancient Steppe cattle. According to Epstein and Mason (1984), longhorn grey-white cattle populated southern, southeastern and Central Europe from the Russian southern steppe regions more than thousand years ago. Moreover, we postulate that the globally famous Jersey cattle have an intrinsic origin in these ancient southern Russian steppe cattle, which is supported by our Y-chromosomal data indicating genetic affinity between the Jersey and the Serbian Podolian cattle. The ancient Steppe cattle may also have been one of the ancestral populations for the Yakutian cattle, as suggested by the current Y-chromosomal data.

Our spatial analyses show a weak phylogeographic structure in European taurine cattle (Figure 4). The Mantel tests (the Yakutian cattle data were excluded) indicated a positive but non-significant relationship between geographic distance and the degree of genetic differentiation detected between the breeds (mtDNA data: $R_{XY} = 0.084$, $P = 0.230$; Y-chromosomal data: $R_{XY} = 0.109$, $P = 0.109$). The current mtDNA and Y-chromosomal data are prone to be influenced by stochastic effects. With the autosomal marker data, it has been possible to detect groups of genetically related breeds (Tapio et al., 2006a) but, with the present uniparentally inherited markers, the groupings or geographical gradients are less evident due to the loss of mitochondrial and Y-chromosomal haplotypes in the course of cattle evolution and/or non-representative breed samples. The maps may be used to direct future samplings from areas where genetic data are absent, thus making connections between the current sampled areas, such as the Caucasian region and northwest Russia. In any case, the current spatial analyses produced interesting observations. They indicate distinct differences, especially among northern indigenous breeds, reflecting their different origins. The mtDNA data imply genetic affinity between Finnish and North Russian breeds (Figure 4a). Felius (1995) suggested that the polled Finnish and northern Scandinavian native breeds may originate from extinct north Russian polled cattle. Photographs of the Russian polled cattle from the early 20th century show that these animals had colour patterns, such as white or red-sided, found currently in Northern and Eastern Finncattle and other north European polled breeds (Kantanen et al., 2000). This may be similar to the eastern dispersal route for domestic sheep (Ovis aries) as discovered by Tapio et al. (2006b). This study of sheep mtDNA concluded that some mitochondrial lineages arrived in northern Europe from the Near East across Russia.

Our Y-data indicated that cattle Y-chromosomal microsatellite variation is structured in haplotype networks by haplogroups as also indicated in human studies (e.g. Bosch et al., 1999). Thirteen Finnish bulls, out of 17 typed previously by Götherström et al. (2005) for unique event markers were assigned to the haplogroup Y1. In this study, the allelic combination of the Y-chromosomal microsatellites showed that these Finnish bulls carry either the haplotype H10 or H11 (Table 1; Figure 3). Interestingly, also eight bulls of Near Eastern-Anatolian origin had the haplotype H11 (Damascus, South Anatolian Red and Turkish Grey) or the haplotypes H9 (Anatolian Black) and H12 (Iraqi), which were within one mutational step from the H11 (Table 1; Figure 3). Although no unique event markers have been typed in these Near Eastern-Anatolian bulls, we suggest that their haplotypes could be assigned to the taurine haplogroup Y1. This indicates that the haplogroup Y1 exists in the present-day native cattle in the region. The data of Götherström et al. (2005) indicated that Y1 was introduced into the gene pool of the taurine cattle as a result of a prehistorical admixture between taurine cattle and North European aurochs (B. primigenius) bulls, but our data do not rule out the possibility that the haplogroup Y1 has descended from the domesticated Near Eastern cattle.

Evidently, the Y1 has not been introduced more recently into the native Near Eastern-Anatolian cattle gene pool by crossing with European modern breeds, such as Holstein-Friesian bulls (Professor Okan Ertugrul, Ankara University, personal communication), whose Y-chromosomes were assigned to the haplogroup Y1 (Götherström et al., 2005). On the other hand, the map displaying genetic affinities between the populations does not reveal an extensive male-mediated introgres-

sion in Turkish Grey, East Anatolian Red, South Anatolian Red, Damascus, Anatolian Black and Iraqi breeds from West European breeds (Figure 4b). In addition, a previous autosomal microsatellite study (Cymbron et al., 2005) showed genetic distinctiveness between Near Eastern and European cattle breeds, excluding the possibility of a recent admixture effect.

MtDNA and Y-chromosome are haploid markers and their effective population sizes are consequently assumed to be equal if both sexes have contributed equally in breeding. In contrast, a higher mutation rate for the Y-chromosomal microsatellite haplotypic system can be expected than for the 255 bp fragment of the mtDNA D-loop. In the mtDNA D-loop, approximately 2.51×10^{-6}–3.77×10^{-6} mutations per generation are expected according to Bradley et al. (1996), if one cattle generation is assumed to be 4–6 years. According to human studies, microsatellite mutation rates on Y-chromosome are typically 2×10^{-3}–3×10^{-3} (Ellegren, 2000). Thus, more Y-chromosomal microsatellite haplotypes should have been found in our data than mtDNA haplotypes. The 57 taurine bulls of 16 different breeds that were analysed also for mtDNA D-loop sequence displayed nine Y-chromosomal haplotypes and 25 mtDNA haplotypes (results not shown). The lower number of Y-haplotypes is due to a smaller number of sires than dams contributing in the breeding process, leading to a lower male effective population size and a greater effect of genetic drift on Y-chromosomal diversity (Hellborg and Ellegren, 2004). Indeed, when the level of breed differentiation based on the present Y-chromosomal data was compared with that of 20 autosomal microsatellite data from the same set of breeds, and the four-times smaller effective population size on the Y-chromosome was taken into account, we found that genetic drift has produced a five-times wider divergence between the breeds for the Y-chromosome than for the autosomes. An intensive culling of breeding males was practiced during animal domestication in prehistoric times (Zeder and Hesse, 2000) and present-day artificial insemination bulls typically descend from only a few elite bulls. On the other hand, male-mediated cross-breeding of locally raised indigenous cattle breeds may have accelerated the loss of original Y-chromosomal haplotypes. Here, we have shown that the most frequent Y-chromosomal haplotype, H11 (Figure 3; Table 1), showed significantly higher frequency in European cattle than in the Near Eastern-Anatolian cattle. H11 appears to segregate in native breeds also (Table 1), but the elevated frequency of the H11 could also be due to the propagation of elite bulls of few West European commercial cattle breeds, such as Holstein-Friesian cattle, to other regions, for example, to northern Russia and Ukraine (Figure 4b; Li et al., 2007a). An irreversible loss of Y-chromosomal genetic diversity has occurred. We suggest that breeds showing less frequent Y-chromosomal haplotypes should be favoured in cattle gene resource conservation programs.

Acknowledgements

References

Achilli A, Olivieri A, Pellecchia M, Uboldi C, Colli L, Al-Zahery N et al. (2008). Mitochondrial genomes of extinct aurochs survive in domestic cattle. Curr Biol 18: R157–R158.
Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16: 37–48.
Bandelt HJ, Forster P, Sykes BC, Richards MB (1995). Mitochondrial portraits of human populations using median networks. Genetics 141: 743–753.
Bosch E, Calafelle F, Santos FR, Pérez-Lezaun A, Cornas D, Benchemsi N et al. (1999). Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. Am J Hum Genet 65: 1623–1638.
Bradley DG, Machugh DE, Cunningham P, Loftus RT (1996). Mitochondrial diversity and the origins of African and European cattle. Proc Natl Acad Sci USA 93: 5131–5135.
Bradley DG, Machugh DE, Loftus RT, Sow RS, Hoste CH, Cunningham EP (1994). Zebu-taurine variation in Y chromosomal DNA: a sensitive assay for genetic introgression in West African trypanotolerant cattle populations. Anim Genet 25: 7–12.
Briggs IC (1974). Machine contouring using minimum curvature. Geophysics 39: 39–48.
Crandall KA, Templeton AR (1993). Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134: 959–969.
Cymbron T, Freeman AR, Malheiro MI, Vigne JD, Bradley DG (2005). Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations. P Roy Soc B-Biol Sci 272: 1837–1843.
Cymbron T, Loftus RT, Malheiro MI, Bradley DG (1999). Mitochondrial sequence variation suggests an African influence in Portuguese cattle. P Roy Soc B-Biol Sci 266: 597–603.
Dmitriev NG, Ernst LK (1989). Animal Genetics Resources of the USSR. Food and Agriculture Organization of the United Nations: Rome, Italy.
Edwards CJ, Baird JF, Machugh DE (2007a). Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. Anim Genet 38: 520–524.
Edwards CJ, Bollongino R, Scheu A, Chamberlain A, Tresset A, Vigne JD et al. (2007b). Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. P Roy Soc B-Biol Sci 274: 1377–1385.
Edwards CJ, Gaillard C, Bradley DG, Machugh DE (2000). Y-specific microsatellite polymorphism in a range of bovid species. Anim Genet 31: 127–130.
Ellegren H (2000). Microsatellite mutations in the germline: implications for evolutionary inference. Trend Genet 16: 551–558.
Epstein H, Mason IL (1984). Cattle. In: Mason IL (eds) Evolution of Domesticated Animals. Longman: London, UK. pp 6–27.
Felius M (1995). Cattle Breeds-An Encyclopedia. Misset: Doetinchem, The Netherlands.

Forster P, Röhl A, Lünnemann P, Brinkmann C, Zerjal T, Tyler-Smith C et al. (2000). A short tandem repeat-based phylogeny for the human Y chromosome. Am J Hum Genet 67: 182–196.

Götherström A, Anderung C, Hellborg L, Elburg R, Smith C, Bradley DG et al. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. P Roy Soc B-Biol Sci 272: 2345–2350.

Goudet J (1995). FSTAT (vers. 1.2): a computer program to calculate F-statistics. J Heredity 86: 485–486.

Granberg L, Soini K, Kantanen J (2009). Sakha Ynaga—Cattle of the Yakuts. Annales Academiae Scientiarum Fennicae. Humaniora no. 355 Gummerrus: Vaajakoski, Finland.

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.

Hanotte O, Okomo M, Verjee Y, Rege E, Teale A (1997). A polymorphic Y chromosome microsatellite locus in cattle. Anim Genet 28: 318–319.

Hanotte O, Tawah CL, Bradley DG, Okomo M, Verjee Y, Ochieng J et al. (2000). Geographical distribution and frequency of a taurine Bos taurus and an indicine Bos indicus Y specific allele amongst sub-Saharan African cattle breeds. Mol Ecol 9: 387–396.

Hellborg L, Ellegren H (2004). Low levels of nucleotide diversity in mammalian Y chromosomes. Mol Biol Evol 21: 158–163.

Huff DR, Peakall R, Smouse PE (1993). RAPD variation within and among natural populations of outcrossing buffalograss Buchloe dactyloides (Nutt) Engelm. Theor Appl Genet 86: 927–934.

Kantanen J, Olsaker I, Brusgaard K, Eythorsdottir E, Holm L-E, Lien S et al. (2000). Frequencies of coat colours and horns in Nordic cattle breeds. Genet Sel Evol 32: 561–576.

Kimura M (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120.

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. Mol Phylogenet Evol 38: 146–154.

Li M-H, Tapio I, Vilkki J, Ivanova Z, Kiselyova T, Marzanov N et al. (2007a). Genetic structure of cattle populations (Bos taurus) in northern Eurasia and the neighboring Near Eastern regions: implications for breeding strategies and conservation. Mol Ecol 16: 3839–3853.

Li M-H, Zerabruk M, Vangen O, Olsaker I, Kantanen J (2007b). Reduced genetic structure of north Ethiopian cattle revealed by Y-chromosome analysis. Heredity 98: 214–221.

Loftus RT, Machugh DE, Bradley DG, Sharp PM, Cunningham P (1994). Evidence for two independent domestications of cattle. Proc Natl Acad Sci USA 91: 2757–2761.

Mannen H, Kohno M, Nagata Y, Tsuji S, Bradley DG, Yeo SJ et al. (2004). Independent mitochondrial origin and historical genetic differentiation in North Eastern Asian cattle. Mol Phylogenet Evol 32: 539–544.

Miller SA, Dykes DD, Polesky H F (1988). A simple salting procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16: 1215.

Nei M (1987). Molecular Evolutionary Genetics. Columbia University Press: New York.

Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288–295.

Pérez-Lezaun A, Calafeli F, Seielstad M, Mateu E, Comas D, Bosch E et al. (1997). Population genetics of Y-chromosome short tandem repeats in humans. J Mol Evol 45: 265–270.

Perret J, Shia Y-C, Fries R, Vassart G, Georges M (1990). A polymorphic satellite sequence maps to the pericentric region of the bovine Y chromosome. Genomics 6: 482–490.

Pieragostini E, Scaloni A, Rullo R, Di Luccia A (2000). Identical marker alleles in Podolian cattle (Bos taurus) and Indian zebu (Bos indicus). Comp Biochem Phys B 127: 1–9.

Rozas J, Rozas R (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15: 174–175.

Saitou N, Nei M (1987). The neighbor joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406–425.

Slatkin M, Hudson RR (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129: 555–562.

Smith WHF, Wessel P (1990). Gridding with continuous curvature splines in tension. Geophysics 55: 293–305.

Tapio I, Värv S, Bennewitz J, Maleviciute J, Fimland E, Grislis Z et al. (2006a). Prioritization for conservation of northern European cattle breeds based on analysis of microsatellite data. Conserv Biol 20: 1768–1779.

Tapio M, Marzanov N, Ozerov M, Ćinkulov M, Gonzarenko G, Kiselyova T et al. (2006b). Sheep mitochondrial DNA variation in European, Caucasian and Central Asian areas. Mol Biol Evol 23: 1776–1783.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DC (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882.

Troy CS, Machugh DE, Bailey JF, Magee DA, Loftus RT, Cunningham P et al. (2001). Genetic evidence for Near-Eastern origins of European cattle. Nature 410: 1088–1091.

Van De Peer Y, De Wachter R (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Biosci 10: 569–570.

Ward TJ, Skow LC, Gallagher DS, Schnabel RD, Nall CA, Kolenda CE et al. (2001). Differential introgression of uniparentally inherited markers in bison populations with hybrid ancestries. Anim Genet 32: 89–91.

Zadworny D, Kuhnlein U (1990). The identification of the kappa-casein genotype in Holstein dairy cattle using the polymerase chain reaction. Theor Appl Genet 80: 631–634.

Zeder MA, Hesse B (2000). The initial domestication of goats (Capra hircus) in the Zagros Mountains 10 000 years ago. Science 287: 2254–2257.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)