Polish Journal of Veterinary Sciences Vol. 16, No. 2 (2013), 265–273

DOI 10.2478/pjvs-2013-0037

Original article

Complete mitochondrial genome of wild aurochs (*Bos primigenius*) reconstructed from ancient DNA

J. Zeyland[1], Ł. Wolko[1], J. Bocianowski[3], M. Szalata[1,2], R. Słomski[1,2], A.M. Dzieduszycki[4], M. Ryba[4], H. Przystałowska[1], D. Lipiński[1,2]

[1] Department of Biochemistry and Biotechnology, Poznan University of Life Sciences, Dojazd 11, 60-632 Poznan, Poland
[2] Institute of Human Genetics, Polish Academy of Sciences, Strzeszynska 32, 60-479 Poznan, Poland
[3] Department of Mathematical and Statistical Methods, Poznan University of Life Sciences, Wojska Polskiego 28, 60-637 Poznan, Poland
[4] The Polish Foundation for Restoration of the Aurochs, Radzyminska 94, 03-574 Warszawa, Poland

Abstract

Extinct aurochs *(Bos primigenius*), accepted as the ancestor of domestic cattle, was one of the largest wild animals inhabiting Europe, Asia and North Africa. The gradual process of aurochs extinction finished in Poland in 1627, were the last recorded aurochs, a female, died. Some aspects of cattle domestication history and the distribution of aurochs genetic material among modern cattle breeds still remain unclear. Analyses of ancient DNA (aDNA) from bone sample deliver new genetic information about extinct wild aurochs as well as modern cattle phylogeny. DNA was extracted from a fragment of aurochs fossil bone found in the Pisz Forest, Poland. The sample was radiocarbon-dated to about 1500 yBP. The aDNA was used for Whole Genome Amplification in order to form a DNA bank. Auroch mitochondrial DNA sequences were amplified using sets of 41 primers overlapping the whole mtDNA, cloned and sequenced. The sequence of the whole mitochondrial genome was reconstructed and deposed in GenBank [GenBank:JQ437479]. Based on the phylogenetic analyses of the *Bovine* mitochondrial genomes, a phylogenetic tree was created. As expected, the tree clearly shows that the mtDNA sequence of the analyzed PWA (Polish Wild Aurochs) individual belongs to haplogroup P. In the course of the comparative mtDNA analysis we identified 30 nucleotide marker positions for haplogroup P and nine unique PWA differences compared to the two remaining haplotype P representatives. Our analysis provides the next step to the reconstruction of the demographic history of this extinct but still exciting species.

Key words: Aurochs, phylogenetics, mtDNA

Correspondence to: J. Zeyland, e-mail jzeyland@yahoo.com, tel.: +48 618 487 202.

Introduction

Research based on the analyses of ancient DNA (aDNA) provides information on the demographic structure of extinct animal species. This way, researchers can study the evolution of these species and analyse the influence of climatic conditions on this process. Based on the results of comparative DNA analyses of existing animal species, four models of post-glacial recolonization of Europe were proposed. These models differ with regard to the share of populations from areas unaffected by climate changes – refugia in the reexpansion (Schmitt 2007). Analyses limited only to modern DNA can be misinterpreted and the geographical distribution of gene clusters may not provide all the information necessary to identify the methods of recolonization. Therefore, researchers have to use ancient DNA as an additional source of information.

The Tertiary remains of the aurochs point to India as the place where the species first appeared between 1.5 and 2 million years ago (Mona et al. 2010). In the Pleistocene, the aurochs had already colonized Eurasia, North Africa and Asia Minor. The oldest aurochs remains discovered in Europe date back 275 000 years. There are three subspecies of aurochs: the Indian aurochs (*Bos primigenius namadicus*), which is the ancestor of the zebu (*Bos indicus*), a domestic type of cattle; the North African aurochs (*Bos primigenius africanus*), whose domesticated descendants can be now found in Africa; and the Eurasian aurochs (*Bos primigenius primigenius*), whose domestication in the 6th century BC gave rise to the modern domestic cattle species in Europe. Comparative analysis of mitochondrial DNA (mtDNA) reveals considerable differences in haplotypes of modern breeds of European domestic cattle and zebu cattle, which confirms the theory of the existence of genetically and geographically isolated populations of aurochs (Bruford et al. 2003). Based on mtDNA sequences, European domestic cattle breeds have been classified mainly into macro-haplogroup T, composed of subgroups T, T1, T2, T3, T4 and T5. Zebu cattle, on the other hand, belong to a genetically distant macro-haplogroup I, which also includes geographically distinguishable subgroups I1 and I2 (Achilli et al. 2008, Chen et al. 2010). On the one hand, the results of aurochs aDNA analyses shed light on the demographic processes within the species and on the other, they partially reconstruct the domestication process of cattle (Beja-Pereira et al. 2006). However, it needs to be emphasized that aDNA analyses are conducted using highly degraded material and usually focus on a selected region of mtDNA. Most of the analyzed aurochs remains indicate that the animals from Northern and Central Europe, as well as a small group of Italian aurochs, belong to haplogroup P, which also includes the individual whose bone was found in Derbyshire, England [GenBank:GU985279] and modern Korean beef cattle [GenBank:DQ124389]. The expansion of haplogroup P is dated to 16-36 thousand years BP. It diverged from macro-haplogroup T about 71 ky, whose expansion took place during the domestication period in the Middle East (Achilli et al. 2008, Mona et al. 2010). The vast majority of aurochs from Italy are classified into haplogroup T, i.e. the macro-group, which also comprises European domestic cattle breeds (mainly subhaplogroup T3), East-Asian breeds (subhaplogroup T4) and modern African cattle (subhaplogroup T1) (Beja-Pereira et al. 2006). Within the analyzed population of modern domestic cattle breeds, 1.4% – mainly from Italy – does not correspond to haplotype T and was classified into haplogroup Q or R. Haplogroup Q, phylogenetically the closest to group T, diverged from it at about 48 ky. Haplogroup R, which diverged from T long before haplogroups P and Q appeared, was formed probably as a result of accidental cross-breeding of domestic bulls and European aurochs wild females, whose offspring were combined with captive herds (Achilli et al. 2009). In certain geographical zones, especially those where the aurochs survived for the longest time, the coexistence of wild individuals of this species and domestic cattle lasted for years, which was conducive not only to accidental cross-breeding of the animals but also allowed men to deliberately breed the animals in order to improve the features of the captive breeds (Bonfiglio et al. 2010). Based on DNA sequence analyses, the aurochs found in Eilsleben, Saxony (Germany) have been classified into an entirely new haplotype E. Results of the Bayesian phylogenetic analysis suggest that this haplotype first appeared nearly 53 000 years ago (Edwards et al. 2007). Extinction of the aurochs started in Africa about 2400 years BC. As for Europe, demographic changes within the species caused its extinction in France around the 10th century AD and in Germany at the turn of the 12th century. In the 17th century, the aurochs was found only in Mazovia, Poland. As a result of hunting, development of farming and domestic cattle grazing, which naturally narrowed the animals' habitat, and due to competition with other wild cattle breeds such as the European bison (*Bison bonasus*), the aurochs became extinct. The last recorded aurochs died in 1627 in the Jaktorów Forest.

In this paper, we present a complete mitochondrial genome sequence of an aurochs referred to as PWA (Polish Wild Aurochs) [GenBank:JQ437479], whose remains, dated to approx. 1500 yBP, were found in the Pisz Forest, Poland. To our knowledge,

this is the first recorded description of fossil material from an aurochs found in Poland and, at the same time, the first complete mtDNA sequence of the species from this country. Having analyzed the complete mtDNA genome sequence, we compared it with selected sequences in modern and ancient *Bovini* samples. The obtained results extend existing knowledge about the evolution and demographic structure of the aurochs and, to some extent, about modern domestic cattle breeds.

Materials and Methods

The study material used was a fragment of a pedicle bone of an aurochs with laboratory code PWA. The fossil material was found near the town of Pisz located in the Pisz Forest, Poland. The sample was radiocarbon dated to about 1500 yBP by an external radiocarbon dating laboratory. Morphological analysis, which confirmed the authenticity of the study material, was carried out in the Pisz Land Museum. All aDNA analyses were conducted in a dedicated laboratory. To avoid contamination, the outer layer of the analyzed fragment was abraded. The fragment was then washed with a solution containing 10% bleach. When dry, the sample was UV-irradiated for 1 hour. The 500-mg sample was ground in a mortar with the use of liquid nitrogen. Ancient DNA extraction was conducted according to the method of aDNA isolation from teeth and bones described by Rohland and Hofreiter (2007). During the isolation, a negative control was used to ensure that the material and reagents were not contaminated with modern DNA. In each set of experiments, a negative control represented by all the reagents except the bone powder was included. After aDNA isolation, DNA concentration in the sample was determined by measuring absorbance at 260 nm. Ancient DNA solution of 1 ng/µl was prepared. WGA was then conducted using the GenomePlex® Whole Genome Amplification (WGA) Kit according to the manufacturer's instructions (Sigma) in order to form aDNA bank for the PWA individual. In the experiment, a negative control represented by all the reagents except aDNA was included. Mitochondrial genome sequences of the PWA individual were amplified using sets of 41 primers overlapping the whole mtDNA. Primer sequences had previously been described (Zeyland et al. 2012). Polymerase chain reaction was conducted in a Veriti Thermal Cycler (Applied Biosystems) in 25 µl reactions containing 125 ng purified WGA product, 1x ReadyMix™ (Sigma Aldrich) and 12.5 µM of each primer. The PCR amplification profile consisted of an initial denaturation at 94°C for 5 min, followed by 30 cycles at 94°C for 45 s, 56°C for 45 s, 72°C for 90 s and a final extension at 72°C for 10 min. PCR products were purified by agarose gel electrophoresis and QIAquick isolation (Qiagen). In the experiments (PCR and purification), negative controls represented by all the reagents except aDNA were included. The PCR amplicons were cloned directly using StrataClone™ PCR Cloning Kit (Stratagene, Agilent Technologies) according to the manufacturer's instructions. *Escherichia coli* competent cells were transformed with recombinant plasmid vectors. Screening of white recombinant colonies was accomplished by PCR transferring the colonies into 25 µl reactions (1x ReadyMix™ and 12.5 µM of each M13 forward and reverse universal primers). The PCR profile was as described above. PCR products were identified by agarose gel electrophoresis. Positive clones were selected and DNA was sequenced bi-directionally using automated genetic analyzers (Applied Biosystems Prism). All fragments were sequenced three times. The obtained sequences were compared with the only complete mtDNA sequence from *B. primigenius* (CPC98) available at the time [GenBank:GU985279]. For those fragments that were found to contain changes, alternative primer pairs which included the changes were designed in order to either confirm or eliminate their presence. Cloning and sequencing were repeated for DNA fragments obtained during PCR using alternative primer pairs. In ancient mtDNA analysis, the recovered sequences may not be authentic but derived from external contamination or the nuclear genome. We regard our PWA individual mtDNA sequence as genuine because all procedures (aDNA extraction, whole genome amplification, polymerase chain reaction, cloning and sequencing) were conducted independently in two laboratories at the Institute of Human Genetics, Polish Academy of Sciences in Poznan and the Department of Biochemistry and Biotechnology, Poznan University of Life Sciences. All results for both laboratories were convergent. The phylogenetic analysis involved following 62 nucleotide sequences of mtDNA: three sequences of haplotype P (including the sequence of the analyzed PWA individual), seven sequences of each haplotype Q, seven sequences of each haplotype R, six sequences of I, 37 sequences of five haplotypes and suphaplotypes of macro-haplogroup T and one mtDNA sequence of wisent (*Bison banasus*) (Table 1). The sequences of whole 16,354 bp mitochondrial genome were used in the final dataset. Phylogenetic analysis was computed using MEGA 5.05. The sequences were aligned using the ClustalW method (Tamura et al. 2011). Based on the results of Bayesian Information Criterion (BIC), the Tamura and Nei substitution model with rate variation among

Table 1. Evolutionary diversity statistics of *Bos* mtDNA haplogroups.

Haplogroup	Number of mtDNAs	Distances within the groups	S.E.	Tv/Ti	Distances to haplogroup P	S.E.
P	2+PWA	0.0011	0.0002	0.00	–	–
T	37+1	0.0008	0.00008	0.10	0.004148	0.00045
Q	7	0.0006	0.0002	0.05	0.003862	0.00046
R	7	0.0011	0.0002	0.03	0.007387	0.00062
I	6	0.0011	0.0002	0.15	0.016272	0.00132
Overall	**61**	**0.0051**	**0.0003**	**0.06**	**–**	**–**

sites modeled with a gamma distribution (shape parameter 0.12) (TN93+G) was chosen for phylogenetic analysis. The maximum likelihood (ML) phylogenetic tree was constructed using MEGA 5.05 software, with 1000 bootstrap replicates for the reliability. The average evolutionary divergence within and between the haplogroups were estimated by a calculation of the number of base substitutions per site from averaging over all sequence. The molecular clock test was performed by comparing the ML value for the given topology with and without the molecular clock constraints under the TN93+G model. The null hypothesis of an equal evolutionary rate throughout the tree was rejected at a 5% significance level ($P < 0.0007$). The uncertainty of molecular clock dating was estimated using the Bayesian Markov chain Monte Carlo (MCMC) approach by computing with the BEAST 1.4.6 program (Drummond and Rambaut 2007), which offers two statistical distributions for describing changes in rate across a branch (Drummond et al., 2006); rates can be drawn independently from either a log-normal distribution (UCLN) or an exponential distribution (UCED). To find out which distribution fits the present data best, we initially fixed the tree topology to topology consensus (T_C). The data were partitioned, with each partition assigned a GTR + I + G substitution model. BEAST Markov chain Monte Carlo (MCMC) steps of 25×10^6 generations following a burn-in of 10^5 generations were performed for UCLN, UCED, and CLOCK models. We calculated the Bayes factor by comparing harmonic mean model likelihoods and used this to choose between models. For both non-autocorrelated models, we also calculated the covariance among branch rates, which indicated the degree of ancestor-descendant autocorrelation of rates across the tree. Using the optimal model, we then accommodated topology by removing the restriction of a fixed tree. Three replicate runs of 25×10^6 generations were performed to check for convergence of the MCMC. Mean parameter estimates and 95% highest posterior densities (HPDs) were determined by analyzing the combined BEAST tree files in TreeAnnotator 1.4.6. We refer to these results with topology flexible (T_F).

Results

Successful PCR amplifications were produced from WGA product of PWA individual for all 41 amplicons. All fragments were cloned and sequenced bi-directionally. All independent PCR amplicons were sequenced three times, to eliminate false polymorphisms occurring as a result of errors generated by polymerase. Sequencing of these fragments allowed for the reconstruction of the entire mitochondrial genome of the aurochs. The obtained sequences were verified in real time by comparison with the only complete mtDNA sequence from *B. primigenius* (CPC98) available at the time in GenBank [GenBank: GU985279]. Additional, alternative PCR amplicon cloning and sequencing were conducted in the case of nucleotide differences with the reference sequence. Any contamination of the aDNA pool with sequences of modern cattle was verified by comparing it with mtDNA sequences of haplogroups T, Q and R described in the GenBank database.

Phylogenetic analysis was carried out for 61 complete mtDNA sequences, which apart from the sequence of the analyzed PWA individual, included two sequences of the previously known haplotype P, 51 sequences of individuals representing three bovine haplogroups (Q, T and R) and six sequences from *B. indicus* classified as haplogroup I. Moreover, mtDNA sequence from *B. banasus* was used as an outgroup. 62 sequences were therefore analyzed in total. Haplogroup P characteristic of the extinct *B. primigenius* is now represented by only three complete mtDNA sequences (including PWA), which may affect the precision of the phylogenetic analyses. However, the ob-

tained level of external genetic diversity within haplogroup P does not diverge from the level observed within other analyzed haplogroups of modern cattle (Table 1). Bayesian analysis in Multidivtime delivered positive but very small values for the degree of autocorrelation of substitution rates across both topologies (Table 4). Finally, analysis of T_C using BEAST indicated that non-autocorrelated models of rate variation fit the data significantly better than a molecular clock (Bayes factor equal to 18.09, $p < 0.001$) (Table 5). Of the non-autocorrelated models, the log-normal distribution (UCLN) had a much better harmonic mean model likelihood than the exponential distribution (UCED), and relaxation (T_F) of a fixed topology further indicated fit. Using

Table 4. Model comparison for analyses relaxing the assumption of autocorrelation of rates across the tree. Harmonic mean model likelihoods were calculated from post-burnin Markov chain Monte Carlo samples generation in BEAST.

Model	Model likelihood	Covariance
T_C		
CLOCK	-35441	N/A
UCED	-33508	0.037 (-0.123, 0.182)
UCLN	-30994	0.081 (-0.033, 0.196)
T_F		
UCLN	-29738	0.063 (-0.097, 0.173)

For these model comparisons, the topology was fixed as T_C. The strict clock model serves as a base comparison. The tree T_F refers to analyses where topology is not fixed. Covariance indicates the degree of substitution rate autocorrelation between ancestor and descendent branches.

Using each of these uncorrelated models, the covariance of substitution rates between ancestor and descendent branches across the tree was not significantly different from zero. Based on the phylogenetic analyses of the sequences listed above, a phylogenetic tree was created. As expected, the tree clearly shows that the mtDNA sequence of the analyzed PWA individual belongs to haplogroup P (Fig. 1). Comparison between the sequence of haplogroup P with major *Bos* haplogroups demonstrated 80 variable nucleotide positions (Tables 2 and 3). Achilli et al. (2009) identified 37 mutations formed after haplotype P diverged from the common ancestor with haplogroup TQ. In the course of the comparative analysis of three haplogroup P representatives, we identified 30 marker positions for this haplogroup. Eleven out of the 30 mutations are disrupted by the reversion of the divergence process determined on the basis of the phylogenetic tree (4 – TQI and 7 – TQR). At the same time however, seven (64%) of the reversions are located within non-coding sequences. Five out of the 37 mutations (2171, 5681, 11468, 12738, 16247) mentioned by Achilli et al. (2009) are typical only for haplotype P found in modern cattle [GenBank:DQ124389]. The sequence of the analyzed PWA individual displays nine unique differences compared to the two haplotype P sequences used in the analysis. All of these specific mutations are classified as transitions. Individual CPC98 and a modern representative of haplotype P (Korean cattle) show 8 and 11 specific mutations respectively. The analyses identified 3 positions (4252, 4293, 16051), which differentiate the PWA individual from the other sequences in haplogroup P but correspond to TQRI bovine sequences. This means that the PWA sequence is more primary and at the same time phylogenetically closer to the common ancestor PQT. From among the nine differences specific for PWA, six are located within protein encoding genes and one in the tRNA-Met gene (m4252C>T). Three mutations are classified as nonsynonymous. Two of them are found in ND1 – 3188 T>C (Y30H) and 3302G>A (A68T), and one in CYTB – 15471C>T (L320F). Comparative analyses of amino acid sequences from the PWA individual and the sequences available in the GenBank database showed that the ND1 (Y30H) mutation detected in PWA had already been described in people. This human analogous, rare among European Leber's hereditary optic neuropathy (LHON; OMIM#535000) patients, mutation is responsible for the highly conserved tyrosine at amino acid 30 conversion to a histidine. This MTND1*LHON3394C allelic variant is found in 1% of the general population and is commonly associated with MTND6*LHON14484C in French Canadians. The MTND1*LHON3394C allelic variant alone serves an intermediate pathogenic role and probably is more dependent on environmental factors for its clinical repression (Brown et al. 1992). The ND1 (A68T) mutation is found in bovine haplotype R1 (Bonfiglio et al. 2010), whereas the CYTB (L320F) mutation has not been found in the database.

Discussion

Ancient DNA used as a study material enables researchers to analyze the history of given populations, which often leads to new, unexpected discoveries in the field of evolutionary demography. The possibility of conducting genomic aDNA analyses is frequently limited by insufficient quantity and quality of genetic material. Standard laboratory procedures,

Table 5. Variable positions in mtDNA of 3 representatives of *B. primigenius* haplotype P and the rest of *Bos* mtDNA haplogroups.

	1 0 6	1 6 6	1 6 9	1 7 3	1 9 0	2 1 5	2 2 2	2 4 9	3 0 0	3 0 1	1 1 2 8	1 4 1	2 1 4 5	2 1 7 1	2 5 5 8	2 5 8 5	3 1 8 8	3 3 0 2	3 3 7 9	3 5 5 0	4 2 2 3	4 2 9 6	4 6 7 6	5 1 5 6	5 5 0 1	5 6 8 1	5 7 4 3	5 8 9 0	5 8 9 9	6 1 6 0
PWA	T	A	G	A	T	*	C	C	A	T	G	A	G	G	A	C	C	A	C	A	T	C	G	A	T	T	C	T	G	C
CPC98	/	/	A	/	/	*	/	/	/	/	/	/	A	/	/	/	T	G	/	/	C	T	/	/	/	/	/	/	/	/
P	C	G	A	G	/	+C	/	/	/	/	/	/	/	A	/	/	T	G	/	/	C	T	/	/	/	C	/	/	/	/
T	/C	/G	/G	/G	C	*/+C	T	T	G	C	A	/A	/	/	G	T	T	G	T	G	/C	/	A	G	C	/	T	C	A	T
Q	C	/	/G	/	C	*	T	T	G	C	A	G	/	/	/	T	T	G	T	G	/	/	A	G	/	/	T	C	A	T
R	C	G	A	/G	C	+TCC	T	/	/	C	A	G	/	/	/	T	T	/	T	/	/	/	A	G	/	/	/	/	A	/
I	C	G	A	/G	C	+C	T	/	/	C	A	G	/	/	/	T	T	G	/	/	/	/	A	G	/	/	/	/	A	T

	6 4 3 6	6 4 5 1	7 3 5 6	7 9 5 2	7 9 9 4	8 2 3 6	8 3 5 8	8 3 7 0	1 0 1 2 6	1 0 6 1	1 1 0 0	1 1 1 4 0	1 1 4 6 8	1 1 7 4 0	1 2 0 6	2 3 4 7	2 4 4 3	2 4 6 8	2 5 6 9	2 6 7 5	2 7 3 5	2 7 7 8	3 0 5 0	3 5 8 5	3 9 1 9	4 0 1 6	4 5 8 9	4 8 7 0	4 1 3 3	5 3 4 4
PWA	A	C	G	T	G	C	T	C	T	G	T	G	G	C	C	C	C	C	C	G	T	C	C	G	C	G	A	T	G	A	T
CPC98	/	T	/	/	/	/	/	/	/	/	/	/	/	T	/	/	/	/	T	A	/	/	/	/	T	/	/	/	A	/	/
P	/	T	/	/	/	/	/	/	/	/	/	/	A	/	/	/	T	/	/	/	/	T	/	/	T	/	/	/	/	/	/
T	G	T	A	C	A	T	C	A	C	C	C	A	/	/	T	T	T	T	/	/	C	/	T	/A	T	A	G	G	/	G	C
Q	/G	T	A	C	A	T	C	/	C	C	/	A	/	/	T	/T	/	/	/	/	/	/	/	/	T	A	/	G	/	G	/
R	/	T	/	C	A	T	C	/	C	/	/	A	/	/	T	T	T	/	T	/	/	/	/	/	T	A	/	G	/	G	/
I	/	T	/	C	A	T	C	/	C	/	/	A	/	/	T	/	/	/	T	/	/	/	/	/	T	A	/	/	/	G	/

	1 5 4 7 1	1 5 6 2 7	1 5 6 7 3	1 5 9 5 1	1 5 9 5 3	1 5 9 9 4	1 6 0 1 9	1 6 0 4 9	1 6 0 5 1	1 6 0 5 8	1 6 0 7 4	1 6 0 8 5	1 6 1 2 2	1 6 1 4 1	1 6 2 3 1	1 6 2 4 7	1 6 2 5 5	1 6 2 6 4	1 6 3 0 1
PWA	T	G	C	C	G	G	T	T	T	T	C	C	C	T	T	C	C	A	C
CPC98	C	/	/	/	/	/	C	/	C	/	/	/	/	C	/	/	/	/	T
P	C	A	/	/	/	/	/	/	C	/	/	/	/	/	/	T	/	/	/
T	C	A	T	T	C	A	/	C	/	C	T	T	T	/	C	/	/T	G	/
Q	C	/	T	T	/	A	/	C	/	/C	T	T	T	/	C	/	/	G	/
R	C	/	T	/	/	A	/	C	/	C	T	/	/	/	/	/	T	/	T
I	C	/	T	/	/	/	/	/	/	/	/	/T	/	/C	C	T	T	G	T

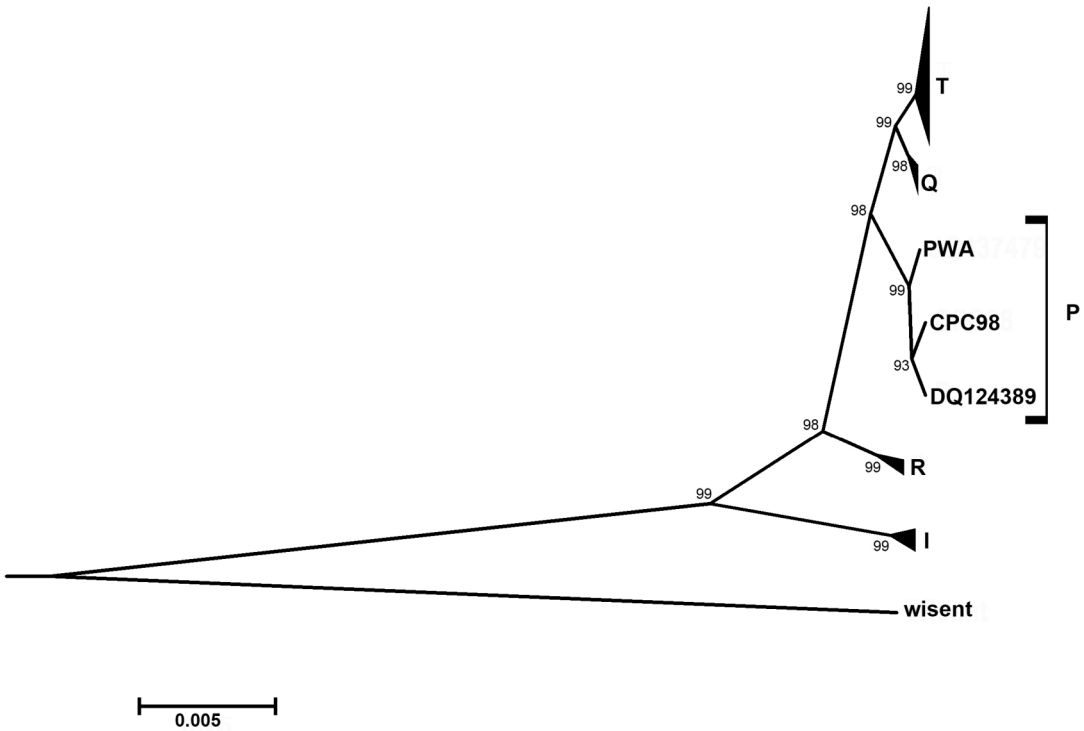

Fig. 1. Maximum likelihood (ML) phylogenetic tree of the representatives of all bovine mitochondrial haplogroups (I, P, Q, R, T) based on complete mitochondrial (mtDNA) genome sequences. This tree was rooted by wisent (*Bison bonasus*) mtDNA sequence as an outgroup. Bootstrap values (1000 replicates) are shown next to the branches. The number of the sequences within each of the haplogroups used in analysis is indicated by the breadth of the branch ends.

Table 2. Polymorphic positions within haplogroup P and to others major *Bos* haplogroups.

Haplotype [GenBank]	Number of mutations	Positions (n- noncoding region, *-nonsynonymus)
PWA [JQ437479]	9	169G(n), 3188C*, 3302A*, 4252T, 4293C, 6451C, 13539C, 15471T*, 16051T(n)
CPC98 [GU985279]	8	2145A, 11740T, 12469T, 12525A, 14580A, 16019C(n), 16141C(n), 16301T(n)
P [DQ124389]	11	106C(n), 166G(n), 173G(n), 221+C(n), 2171A, 5681C, 11468A, 12443T, 12738T, 15627A, 16247T(n)
TQRI	19	190T, 222C, 301T, 1128G, 1481A, 2585C, 4676G, 5156A, 5899G, 7952T, 7994G, 8236C, 8358T, 10126T, 11140G, 12016C, 13821G, 14873A, 15673C,
TQI	4	6160C, 16085C, 16231T, 16264A,
TQR	7	3379C, 12377C, 14129T, 15994G, 16049T, 16058T, 16074C,
TQ	11	249C, 300A, 3550A, 5743C, 5890T, 6436A, 7356G, 10691G, 15951C, 16122C, 16085C,
T	11	2558A, 5501T, 8370C, 11000A, 12468C, 12675T, 12750C, 13005G, 14036A, 15134T, 15953G

Table 3. Degree of autocorrelation in rates of molecular evolution by partition and tree T_C topology.

Genetic partition	Autocorrelation (95% confidence interval)
First positions	0.00204 (0.00176, 0.00439)
Second positions	0.00576 (0.00310, 0.00835)
Third positions	0.00501 (0.00338, 0.00835)

which include amplification, cloning and sequencing, prove effective only in the case of mitochondrial DNA. Therefore, obtaining complete mtDNA genome sequences of extinct animal species is largely determined by the quality of the analyzed samples. This paper presents the analysis of a complete mtDNA sequence. DNA was extracted from a fragment of a pedicle bone of an aurochs, found in the Pisz Forest, Poland. The sample dates to 1500 yBP.

The aurochs, an extinct wild ancestor of modern cattle breeds, was formerly widespread across Eurasia and northern Africa. The history of cattle domestication and the distribution of aurochs genetic material among modern cattle breeds remain unclear. What we know about these processes today is mainly based on the information obtained from the analyses of short fragments of mtDNA control regions. Based on the control region sequences, all Northern and Central European aurochs, as well as a smaller fraction of Italian aurochs were classified into haplogroup P, and the majority of Italian aurochs into haplogroup T (Mona et al. 2010). The vast majority of European domestic cattle breeds are also classified into haplogroup T. Only some of these belong to haplogroup P, typical of aurochs from North/Central Europe, or less frequently haplogroups R and Q (Beja-Pereira et al. 2006). So far, the only instance of haplotype E in *B. primigenius* has been identified as a result of mtDNA control region analysis based on fossil materials (Edwards et al. 2007). Genetic similarity between Italian aurochs and the European breeds determined with the analyses of short fragments of mtDNA control regions will certainly help us to understand the process of cattle domestication. However, it needs to be remembered that hypotheses based on such limited study material, i.e. selected fragments of mtDNA, are not conclusive as they may be burdened with error. Long simultaneous coexistence of domestic and wild breeds makes it difficult to collect information on the processes shaping the populations of domestic cattle and the aurochs. The first literature report describing a complete mtDNA sequence of an aurochs was published in 2010 (Edwards et al.). Individual CPC98, whose bone fragment dated to 6700 years, had been found in England and was classified into haplogroup P, which is consistent with the results of the analyses of mtDNA control region sequences for the majority of individuals analyzed within this species (Mona et al. 2010). The obtained complete sequence of mitochondrial genome of an aurochs from Poland (PWA) also belongs to haplogroup P. PWA is highly homologous to CPC98. Both sequences differed only with respect to 17 nucleotide positions. Three nonsynonymus mutations observed in the PWA sequence are very rare (ND1 Y30H) or unprecedented (CYTB L320F). Therefore, it needs to be remembered that the polymorphisms observed in ancient DNA are to some extent caused by the accumulation of mutations by post-mortem DNA modification in fossil remains by oxidative or hydrolytic modification of bases (Willerslev and Cooper 2005) or are generated at the stage of material amplification. The mitochondrial genome of the Polish and British individuals belongs to Neolithic haplogroup P, typical of the aurochs, whereas the genome of the animal described by the Italian researchers qualifies into pre-Neolithic homogeneous haplogroup T, typical of mtDNA sequence for the majority of European domestic cattle breeds (Lari et al. 2011). Haplogroups R and Q have been identified in an ancient cattle breed presently found in Italy. As the authors suggest, these groups may have originated from the hybridization of the aurochs with domestic cattle (Achilli et al. 2008, Beja-Pereira et al. 2006, Achilli et al. 2009, Bonfiglio et al. 2010). Thus, it may be speculated that the aurochs population in Northern/Central Europe differed from the one in Southern Europe, and that the latter could have considerably determined the process of cattle domestication on this continent (Achilli et al. 2009, Beja-Pereira et al. 2006, Mona et al. 2010). During the Pleistocene and between glaciations Northern and Central Europe was climatically unfavorable for the aurochs population, unlike Southern Europe, which proved to be a more suitable habitat for this species. Climatic and environmental changes affected genetic variation. Genetic variation observed in Italy, compared to the population in Northern/Central Europe, is considerably higher, which confirms the previous thesis and indicates that the Southern refugium could have been a reservoir of genetic variation for the species. The observed correlation between the region of ancient material recovery and the haplogroup of a given individual suggest that the aurochs from Northern/Central Europe could have become isolated from those from the South of the continent. Sequences specific for group P observed in the aurochs from Northern/Central Europe are extremely rare among the Italian individuals. Adverse conditions during the Pleistocene glaciations could have led to the reduction of the Northern population of *B. primigenius*, whereas the founder effect and the lack of gene flow caused haplotype unification within the Northern/Central population. Even interglacial periods did not affect the demographic expansion of the aurochs to any substantial degree, leaving the species with rather low genetic variation (Mona et al. 2010). Haplogroups P and T differ also with regard to their demographic history. Macro-haplogroup T, which originated in the Middle East, underwent a Neolithic genetic bottleneck and subsequent rapid expansion as a consequence of domestication and cross-breeding of the individual. The history of haplogroup P, on the other hand, does not show evidence of genetic bottleneck (Troy et al. 2001, Achilli et al. 2008, Edwards et al. 2010). Genetic differences between northern/central European (haplotype P) and Italian aurochs (haplotype T) were fully confirmed by the PWA ancient mtDNA sequence analysis conducted in this study. Polish Wild Aurochs was found in Pisz Forest, the

second largest forest complex in Europe, located in the northern/central part of the continent. We identified haplotype P, which is not surprising and it confirms the current theory on classification of European aurochses. Due to the insufficient amount of excavated material and the fact that aurochses can no longer be found among contemporary animals, we should be very careful about forming theories on aurochs demography. We may hope that materials found in the future will help in understanding the history of this animal.

References

Achilli A, Bonfiglio S, Olivieri A, Malusg A, Pala M, Hooshiar Kashani B, Perego UA, Ajmone-Marsan P, Liotta L, Semino O, Bandelt HJ, Ferretti L, Torroni A (**2009**) The multifaceted origin of taurine cattle reflected by the mitochondrial genome. PLoS One 4: e5753.

Achilli A, Olivieri A, Pellecchia M, Uboldi C, Colli L, Al-Zahery N, Accetturo M, Pala M, Hooshiar Kashani B, Perego UA, Battaglia V, Fornarino S, Kalamati J, Houshmand M, Negrini R, Semino O, Richards M, Macaulay V, Ferretti L, Bandelt HJ, Ajmone-Marsan P, Torroni A (**2008**) Mitochondrial genomes of extinct aurochs survive in domestic cattle. Curr Biol 18: R157-R158.

Anderson S, de Bruijn MH, Coulson AR, Eperon IC, Sanger F, Young IG (**1982**) Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. J Mol Biol 156: 683-717.

Beja-Pereira A, Caramelli D, Lalueza-Fox C, Vernesi C, Ferrand N, Casoli A, Goyache F, Royo LJ, Conti S, Lari M, Martini A, Ouragh L, Magid A, Atash A, Zsolnai A, Boscato P, Triantaphylidis C, Ploumi K, Sineo L, Mallegni F, Taberlet P, Erhardt G, Sampietro L, Bertranpetit J, Barbujani G, Luikart G, Bertorelle G (**2006**) The origin of European cattle: evidence from modern and ancient DNA. Proc Natl Acad Sci U S A 103: 8113-8118.

Bonfiglio S, Achilli A, Olivieri A, Negrini R, Colli L, Liotta L, Ajmone-Marsan P, Torroni A, Ferretti L (**2010**) The enigmatic origin of bovine mtDNA haplogroup R: sporadic interbreeding or an independent event of *Bos primigenius* domestication in Italy? PLoS One 5: e15760.

Brown MD, Voljavec AS, Lott M T, Torroni A, Yang CC, Wallace DC (**1992**) Mitochondrial DNA complex I and III mutations associated with Leber's hereditary optic neuropathy. Genetics 130: 163-173.

Bruford MW, Bradley DG, Luikart G (**2003**) DNA markers reveal the complexity of livestock domestication. Nat Rev Genet 4: 900-910.

Chen S, Lin BZ, Baig M, Mitra B, Lopes RJ, Santos AM, Magee DA, Azevedo M, Tarroso P, Sasazaki S, Ostrowski S, Mahgoub O, Chaudhuri TK, Zhang YP, Costa V, Royo LJ, Goyache F, Luikart G, Boivin N, Fuller DQ, Mannen H, Bradley DG, Beja-Pereira A (**2010**) Zebu cattle are an exclusive legacy of the South Asia neolithic. Mol Biol Evol 27: 1-6.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4:e88.

Drummond AJ, Rambaut A (**2007**) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

Edwards CJ, Bollongino R, Scheu A, Chamberlain A, Tresset A, Vigne JD, Baird JF, Larson G, Ho SY, Heupink TH, Shapiro B, Freeman AR, Thomas MG, Arbogast RM, Arndt B, Bartosiewicz L, Benecke N, Budja M, Chaix L, Choyke AM, Coqueugniot E, Döhle HJ, Göldner H, Hartz S, Helmer D, Herzig B, Hongo H, Mashkour M, Ozdogan M, Pucher E, Roth G, Schade-Lindig S, Schmölcke U, Schulting RJ, Stephan E, Uerpmann HP, Vörös I, Voytek B, Bradley DG, Burger J (**2007**) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. Proc Biol Sci 274: 1377-1385.

Edwards CJ, Magee DA, Park SD, McGettigan PA, Lohan AJ, Murphy A, Finlay EK, Shapiro B, Chamberlain AT, Richards MB, Bradley DG, Loftus BJ, MacHugh DE (**2010**) A complete mitochondrial genome sequence from a mesolithic wild aurochs (*Bos primigenius*). PLoS One 5: e9255.

Lari M, Rizzi E, Mona S, Corti G, Catalano G, Chen K, Vernesi C, Larson G, Boscato P, De Bellis G, Cooper A, Caramelli D, Bertorelle G (**2011**) The complete mitochondrial genome of an 11,450-year-old aurochsen (*Bos primigenius*) from Central Italy. BMC Evol Biol 11: 32.

Mona S, Catalano G, Lari M, Larson G, Boscato P, Casoli A, Sineo L, Di Patti C, Pecchioli E, Caramelli D, Bertorelle G (**2010**) Population dynamic of the extinct European aurochs: genetic evidence of a north-south differentiation pattern and no evidence of post-glacial expansion. BMC Evol Biol 10: 83.

Rohland N, Hofreiter M (**2007**) Ancient DNA extraction from bones and teeth. Nat Protoc 2: 1756-1762.

Schmitt T (**2007**) Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. Front Zool 4: 11.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (**2011**) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.

Troy CS, MacHugh DE, Bailey JF, Magee DA, Loftus RT, Cunningham P, Chamberlain AT, Sykes BC, Bradley DG (**2001**) Genetic evidence for Near-Eastern origins of European cattle. Nature 410: 1088-1091.

Willerslev E, Cooper A (**2005**) Ancient DNA. Proc Biol Sci 272: 3-16.

Zeyland J, Wolko L, Lipiński D, Woźniak A, Nowak A, Szalata M, Bocianowski J, Słomski R (**2012**) Tracking of wisent-bison-yak mitochondrial evolution. J Appl Genet 53: 317-322.