Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies

Hiendleder, Stefan; Kaupe, Bernhard; Wassmuth, Rudolf; Janke, Axel

Published in:
Royal Society of London. Proceedings B. Biological Sciences

DOI:

2002

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain.
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies

Stefan Hiendleder¹,²,³*, Bernhard Kaupe¹, Rudolf Wassmuth¹ and Axel Janke⁴

¹Department of Animal Breeding and Genetics, Justus-Liebig-University, Ludwigstr. 21B, D-35390 Giessen, Germany
²Central Biotechnical Unit, Strahlenzentrum of the Justus-Liebig-University, Leihgesterner Weg 217, D-35392 Giessen, Germany
³Department of Molecular Animal Breeding and Biotechnology, Ludwig-Maximilians University Munich, Hackerstraße 27, D-85764 Oberschleißheim, Germany
⁴Division of Evolutionary Molecular Systematics, Department of Genetics, University of Lund, Sölvegatan 29, S-22362 Lund, Sweden

Complete mitochondrial DNA (mtDNA) control regions (CR) were sequenced and analysed in order to investigate wild sheep taxonomy and the origin of domestic sheep (Ovis aries). The dataset for phylogenetic analyses includes 63 unique CR sequences from wild sheep of the mouflon (O. musimon, O. orientalis), urial (O. vignei), argali (O. ammon) and bighorn (O. canadensis) groups, and from domestic sheep of Asia, Europe and New Zealand. Domestic sheep occurred in two clearly separated branches with mouflon (O. musimon) mixed into one of the domestic sheep clusters. Genetic distances and molecular datings based on O. canadensis CR and mtDNA protein-coding sequences provide strong evidence for domesticaions from two mouflon subspecies. Other wild sheep sequences are in two additional well-separated branches. Ovis ammon collum and O. ammon nigrimontana are joined with a specimen from the transkaspian Ust-Urt plateau currently named O. vignei arkal. Ovis ammon amnon, O. ammon darziani and O. vignei bohchariansis represent a separate clade and the earliest divergence from the mouflon group. Therefore, O. musimon, O. vignei bohchariansis and Ust-Urt sheep are not members of a ‘moufloniform’ or O. orientalis species, but belong to different clades. Furthermore, Ust-Urt sheep could be a hybrid population or an O. ammon subspecies closely related to O. ammon nigrimontana.

Keywords: Ovis; mitochondrial DNA; domestication; molecular dating; taxonomy; conservation genetics

1. INTRODUCTION

Domestic animals have played an important role in human history, but the origins of most livestock species are not well understood. Mitochondrial (mt) sequences have been used to study the origin of cattle (Loftus et al. 1994), swine (Giuffra et al. 2000), horse (Vila et al. 2001) and goat (Luikart et al. 2001). For these species, the number of presumed wild progenitors is limited. For the domestic sheep (Ovis aries), however, a large number of wild and possibly ancestral species and subspecies exist (Ryder 1984). Accordingly, several Eurasian wild sheep of the highly polymorphic genus Ovis have been proposed as ancestors of domestic sheep or are believed to have contributed to specific breeds (Reed 1960; Zeuner 1963). The taxonomy of these wild sheep is confused and controversial (Geist 1991a), hindering unequivocal identification and classification for conservation management (Geist 1991b) of this important genetic resource for a major agricultural species.

Based on morphological data, numerous wild sheep classifications and revisions have been proposed during the last two centuries (summarized in Nadler et al. 1973a). A basic difference lies in the number of species recognized. Some authors follow the classification of Tsalkin (1951), proposing two (O. ammon, O. nivicola/canadensis) or even a single polymorphic species (O. ammon) for Eurasian and North American sheep. Other classifications follow Lydekker (1913) and Nasonov (1923) and discriminate between four to nine species. In a more recent classification based on chromosome numbers, n and geographical distribution of wild sheep, Nadler et al. (1973a) recognize four groups of sheep. These are mouflon (O. musimon/orientalis, 2n = 54), urial (O. vignei, 2n = 58), argali (O. ammon, 2n = 56) and amphiberian dall sheep (O. dalli, 2n = 54), bighorn sheep (O. canadensis, 2n = 54) and snow sheep (O. nivicola) (O. ammon). However, Siberian snow sheep were later shown to have a karyotype of 2n = 52 (Korobitsina et al. 1974). Geist (1991a) recognizes six species for the genus Ovis: the primitive vignei, the paedomorphic orientalis, the hypermorphic ammon and three pachycerine species, nivicola, dalli and canadensis. Despite the differences in chromosome number, different species of the genus Ovis can hybridize in captivity (Vorontsov et al. 1972; Nadler et al. 1973b) and natural habitats where

*Author for correspondence (S.Hiendleder@gen.vetmed.uni-muenchen.de).
ranges overlap (Nadler et al. 1971; Valdez et al. 1978), and produce fertile offspring. Accordingly, mouflon/urial hybrid zones with individuals displaying intermediate chromosome numbers of 55 to 57 are observed in northern (Nadler et al. 1971) and southeastern Iran (Valdez et al. 1978). These data have been interpreted in support of a single ‘moufloniform’ (O. orientalis) species, comprising mouflon and urial populations (Valdez et al. 1978; Valdez 1982). The International Union for the Conservation of Nature and Natural Resources (IUCN) has essentially adopted this classification of only three species (O. orientalis, O. ammon and O. nivicola) of Eurasian wild sheep (Shackleton et al. 1997).

Three major groups of Eurasian wild sheep, mouflon, urial and argali, have been proposed as ancestors of the domestic sheep or are believed to have contributed to specific breeds. Reed (1960) suggested that archaeological evidence pointed to Asian mouflon (O. orientalis) as the ancestor of domestic sheep. According to Zeuner (1963), however, the urial (O. vignei) was the first domesticate in the Aralo-Caspian basin and domestic forms subsequently spread throughout the Middle East and into Europe. Another line of domesticated sheep was believed to be derived from mouflon (O. musimon or O. orientalis) stock, which was brought into Europe and mixed with the urial derivatives. According to this view, the domestic sheep of southeast Asia are derived from urial, but argali (O. ammon) alleles are supposed to have been introduced repeatedly into these lines (Zeuner 1963).

In accordance with the domestication hypothesis of Zeuner (1963), Wood & Phua (1996) identified two major mitochondrial DNA (mtDNA)-types among New Zealand sheep based on sequence comparisons of the mtDNA control region (CR), but the origin or relationship to wild sheep populations of the mouflon (Ovis orientalis), urial (O. vignei) and argali (O. ammon) have been determined (Hiendleder et al. 1998a). Phylogenetic analyses on mtDNA restriction fragment length polymorphisms (RFLP) of Eurasian breeds (Hiendleder et al. 1998a) and CR sequence data of New Zealand sheep (Hiendleder et al. 1999) demonstrated two different maternal origins among modern sheep breeds. However, although the RFLP analyses showed a close relationship between some O. aries and mouflon (O. musimon) haplotypes, the data suggested an additional wild ancestor different from the proposed urial (O. vignei) and argali (O. ammon) sheep (Hiendleder et al. 1998a).

Complete CR sequences of O. aries reported to date are from European breeds (Zardoya et al. 1995; Zardoya et al. 1996) or breeds of European origin (Wood & Phua 1996) and just two complete CR sequences of wild sheep, mouflon (O. musimon) and urial (O. vignei boharinesis), have been determined (Hiendleder et al. 1998a). Phylogenetic analyses have only been performed on a subset of the available domestic sheep CR sequence data (Hiendleder et al. 1999). In the present study we have produced additional CR sequences that represent domestic sheep from west and central Asia, and eight wild sheep populations of the mouflon (O. musimon), urial (O. vignei), argali (O. ammon) and bighorn (Ovis canadensis) groups. The complete Ovis CR database has been used for phylogenetic analyses in order to investigate the origin and evolution of domestic sheep and to help clarify the taxonomy of the genus Ovis. New sequences from 2.3 kb of O. canadensis canadensis mtDNA protein-coding regions were used to calibrate divergence times among the different sheep lineages.

2. MATERIAL AND METHODS

DNA of wild and domestic sheep (table 1) was extracted from blood, skin or liver samples by standard procedures (Sambrook et al. 1989). Two primers, tRNA-Phe (5’-TCATCTAGG-CATTTTTCAGTG-3’) and tRNA-Pro (5’-CTCACCATCAA-CCCCCAAAGC-3’) were designed from the ovine reference sequence, O. aries [AF010406] (Hiendleder et al. 1998b), to amplify the complete mtDNA CR, using Pwo polymerase (Hybaid-AGS, Heidelberg, Germany). The PCR products were purified from agarose gels and cloned into pCR-TOPO-BluntII (Invitrogen, Groningen, The Netherlands). Both strands of plasmid inserts were sequenced by standard procedures on a LICOR 4200 automated sequencer (MWG Biotech, Ebersberg, Germany), Ovis ammon ammon and O. ammon darrwini CR was amplified using Taq polymerase (Hybaid-AGS), gel purified and cloned into pCR2.1-TOPO (Invitrogen). Complete CR sequences of three clones each, derived from independent PCR reactions, were determined using ’S sequencing reactions and an additional internal primer (DL3R 5’-TAGTGAGAT-GGCCCTGA-3’) according to standard procedures (Sambrook et al. 1989). In case of nucleotide (nt) variation among clones from the same individual, the consensus sequence was reported. In cases where the PCR produced several bands due to heteroplasmy (figure 1), the dominant fragment regardless of its size was cloned. CR nt sequence data reported in this paper have been submitted to GenBank with accession numbers AF242347, AF242348 and AY091486–AY091500. Users of these data are kindly requested to refer to the present publication.

Other complete CR sequences of the domestic sheep (L29055, Zardoya et al. (1995); Z35228-Z35268 and Z35293, Wood & Phua (1996); AF039578 and AF039579, Hiendleder et al. (1998a)) of the Rasa Aragonesa, Merino, Romney, Coopworth, Perendale × Romney and Merinolandschaf breeds, one O. musimon (AF039577, Hiendleder et al. (1998a)) and one O. vignei boharinesis (AF039580, Hiendleder et al. (1998a)) were obtained from GenBank.

The CR sequences were aligned manually. The number of repetitive sequence elements was reduced to the two repeats found in O. canadensis. In all 66 sequences, the 5’ and 3’ copies were retained. Maximum likelihood (ML) trees, distances and support values were calculated using the PAUP* program package v. 4.0 package (Strimmer & von Haeseler 1996). The Tamura–Nei (TN-93) model of nt evolution (Tamura & Nei 1993) with the assumption of rate heterogeneity among sites was used for the analysis of nt sequence data. In order to confirm the presence of major clusters identified by the ML analysis, the sequences were also analysed using maximum parsimony (Fitch 1971) and neighbour joining (Saitou & Nei 1987) as implemented in the PAUP* program package (Swoford 1998).

PCR primers (OatGL 5’-AGTACACTGACTTCCAAA-3’ and OatLH 5’-AAATTTTTTGGTTCTAAGA-3’) were used to amplify a 2.3 kb region (nt positions 9435–11713) of the mt genome of the bighorn sheep, O. canadensis canadensis. The primers were designed from the complete mtDNA sequence of the domestic sheep, O. aries [AF010406] (Hiendleder et al. 1998b). The PCR products were twice precipitated with ethanol and sequenced directly by cycle sequencing using fluorescent dye-labelled primers, and analysed on a LICOR 4000L automated sequencer (AH diagnostics, Skaftafell, Iceland) and MWG
Table 1. Origin of wild and domestic sheep sampled for mitochondrial DNA control region sequences.

<table>
<thead>
<tr>
<th>name*</th>
<th>breedb</th>
<th>number of specimens</th>
<th>country/region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovis canadensis canadensis</td>
<td>n.a.</td>
<td>1</td>
<td>Canada/Rocky Mountains</td>
</tr>
<tr>
<td>Ovis musimon</td>
<td>2</td>
<td>Germany/Hessen</td>
<td></td>
</tr>
<tr>
<td>Ovis vignei bochariensis</td>
<td>n.a.</td>
<td>2</td>
<td>Turkmenistan/southeast</td>
</tr>
<tr>
<td>Ovis vignei arkal</td>
<td>n.a.</td>
<td>1</td>
<td>Kazakhstan/Ust-Urt</td>
</tr>
<tr>
<td>Ovis ammon ammon</td>
<td>1</td>
<td>Mongolia/Altai</td>
<td></td>
</tr>
<tr>
<td>Ovis ammon darrvini</td>
<td>n.a.</td>
<td>1</td>
<td>Mongolia/Gobi-Altai</td>
</tr>
<tr>
<td>Ovis ammon nigrimontana</td>
<td>n.a.</td>
<td>2</td>
<td>Kazakhstan/Kara-Tau</td>
</tr>
<tr>
<td>Ovis ammon collium</td>
<td>n.a.</td>
<td>1</td>
<td>Kazakhstan/Karaganda</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Edilbey</td>
<td>1</td>
<td>Kazakhstan/Alma-Ata</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Astrachan</td>
<td>1</td>
<td>Kazakhstan/Tschimkent</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Gizzarr</td>
<td>1</td>
<td>Tadjikistan</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Awassi</td>
<td>1</td>
<td>Syria</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Kivircik</td>
<td>1</td>
<td>Turkey/Aegean</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Daglic</td>
<td>1</td>
<td>Turkey/west Turkey</td>
</tr>
<tr>
<td>Ovis aries</td>
<td>Akkaraman</td>
<td>1</td>
<td>Turkey/central Anatolia</td>
</tr>
</tbody>
</table>

* Nomenclature is according to Vorontsov et al. (1972).

b n.a., not applicable.

Figure 1. 1.3% agarose gel showing heteroplasmic length variation in sheep CR. PCR amplified complete CR were digested with BamHI to visualize an invariant 3' fragment and variable length 5' fragments containing from 4 to 7 copies (R4–R7) of the sheep 75/76 nt tandem-repeat motif (arrows). (1) Homoplasmic Ovis aries, (2) heteroplasmic O. aries, (3) heteroplasmic O. musimon, (4) heteroplasmic O. ammon collium and (5) heteroplasmic O. ammon nigrimontana.

Biotech, Munich Germany). This sequence has been deposited at GenBank with accession number AJ409303.

The protein-coding genes of NADH3, NADH4L and NADH4 of the bighorn sheep, O. canadensis (this study), cow, Bos taurus [J01394] (Anderson et al. 1982), domestic sheep, O. aries [AF010406] (Hiendleder et al. 1998b), pig, Sus scrofa [U17859] (Ursing & Arnason 1996), alpaca, Lama pacos [Y19184] (Ursing et al. 1998a), fin whale, Balaenoptera physalus [X61145] and blue whale, B. musculus [X72204] (Arnason & Gullberg 1993), the hippopotamus, Hippopotamus amphibius [AJ019057] (Ursing & Arnason 1998b), horse, Equus caballus [X79547] (Xu & Arnason 1994) and the donkey, E. asinus [X97337] (Xu et al. 1996) were aligned by eye. No gaps needed to be included. To this dataset the alignment of the partial Cyt b gene of O. canadensis [U17859] (Groves & Shields 1996) to the Cyt b of the species mentioned above was added. After removal of gaps produced by adding the partial Cyt b alignment, 999 amino-acid (aa) sites remained for analysis. Distance and support values were calculated by the TREE PUZZLE v. 4.0 program (Strimmer & von Haeseler 1996).

3. RESULTS

(a) General characteristics of ovine control regions

The lengths of the ovine mt CR sequences range from 1031 nt (O. canadensis canadensis) to 1333 nt (O. ammon nigrimontana). This length variability is caused by different copy numbers of a 75 nt tandem-repeat element in the Ovis CR (Hiendleder et al. 1998a,b). Repeat numbers vary from two in O. canadensis canadensis, to four in all but one O. aries, O. musimon, O. vignei bochariensis, five in O. aries Astrachan, O. vignei arkal, O. ammon collium, O. ammon ammon, and a maximum of six in both O. ammon nigrimontana. In some wild and domestic sheep samples, evidence for heteroplasmy (Hiendleder et al. 1998b), i.e. a variable number of repeat units within individuals, was observed (figure 1). In this case, the predominating band had been cloned, as it represents the majority of the mtDNA population within an individual. All O. aries sequences and the O. vignei bochariensis sequence obtained from GenBank had four copies of the tandem-repeat element, except for the O. musimon sequence, which contained five copies. The sequence variation among repeats of an individual was much smaller than between individuals. This allowed trimming of the number of repeats in all of the sequences to two, the lowest observed number of repeats in this dataset, and includes the sequence data of the repetitive elements for analysis.

The nt frequency in Ovis CRs standardized to two tandem-repeat elements is 31.9% adenine, 24.0% cytosine, 14.6% guanine and 29.5% thymine. These values are expected for mammalian CR sequences. Apart from repeat copy-number length differences, only three insertion/deletion events (indels) of more than a single nt were detected. All of the indels occurred within the first 180 nt of the CR. A four nt deletion was found in four O. aries sequences from GenBank, and an insertion of eight and a deletion of six nt are present in O. canadensis canadensis (figure 2). The tandem-repeat element exhibits...
an insertion in a subset of the newly produced sequences. In *O. ammon collium*, *O. ammon nigrimontana* and the specimen designated *O. vignei arkal*, a “T” insertion leads to a 76 nt sequence in all but the 5’-most repeat units (figure 3a).

Within each tandem repeat, mirror symmetrical and thus stem-loop structure-forming octamer sequences (Hiendleder et al. 1998b) are highly conserved. But the potential recognition site for H-strand arrest is variable among sheep. Nt variation in proposed ovine homologues (Wood & Phua 1996; Hiendleder et al. 1998b) of TAS-A and CSB 2 + 3 is low, but it is considerable in CSB 1 (figure 3a–d). By contrast, L- and H-strand promoters and the origin of H-strand replication are completely conserved (not shown).

### (b) Sheep phylogeny and molecular dating

The initial alignment contained 63 unique sequences with a length of 1045 nt (figure 2). Only 13 indel events were observed. Indels were excluded prior to phylogenetic analyses because the placement of indels is sometimes ambiguous and the evolution of gaps is not implemented in most molecular evolutionary models. After removal of gaps and ambiguous sites around gaps, 1008 nt with 205 variable positions (20.3%) in 61 unique sequences remained for analysis. All phylogenetic reconstruction methods found the same major clusters as shown in figure 4a,b. The trees are shown unrooted in order to illustrate the distinctness of the major clades. However, when CR sequences of the domestic goat, *Capra hircus* (Takada et al. 1997) and cow, *B. taurus indicus* (Loftus et al. 1994), were included as outgroups in the initial alignment, 862 characters remain for phylogenetic analyses after excluding gaps. The root of the sheep tree is then placed by all analysis methods along the branch leading to the bighorn sheep (*O. canadensis*). Wild and domestic sheep occurred in four clearly separated branches. Two of the branches contain distinct clusters of domestic sheep. The relationships within each of the clusters of the domestic sheep remain unresolved and thus form star-like patterns. One of these clusters encompasses all three representatives of the mouflon (*O. musimon*) Another branch groups *O. ammon collium*, *O. ammon nigrimontana* and a specimen from the transkaspian Ust-Urt plateau, as a clearly separated clade. *O. ammon ammon*, *O. ammon darwini* and *O. vignei bochariensis* occurred as the earliest split from the most distant bighorn (*O. canadensis canadensis*). However, the branching pattern remains unresolved.

The maximum distance among the CR sequences is ca. 11% and is observed between the bighorn and all other sheep. Distances are lower for other groups of sheep. Especially within the clusters of domestic sheep, only a few substitutions are observed. The low distances and the high transition/transversion ratio (table 2) indicate a limited amount of randomization in the sequence data, making them suitable for estimating divergence times. These can be calculated after the substitution rate among the sheep has been estimated. In order to calibrate the substitution rate, a well-established reference point like the A/C-60 can be used. The A/C-60 reference for calibrating substitution rates is based on the assumption that an artiodactyl family, the *Ruminantia*, and *Cetacea* diverged at 60 million years before the present (MyrBP) (Arnason & Gullberg 1996). This date is supported by the palaeontology of whales and has been corroborated by the use of other calibration points for the molecular clock within mammals (Arnason et al. 2000). However, CR sequences evolve too rapidly and can not be directly utilized for comparisons between sheep, cattle and whales. Instead, the deepest divergence among sheep has been dated on the basis of protein-coding sequences.

The ML distance between the domestic sheep (*O. aries*) and the bighorn (*O. canadensis*) protein-coding genes has been calculated to 0.01508 using the mtREV-24 model of aa sequence evolution. The number of absolute aa differences is 15. The average ML distance between the cow and the two sheep is 0.0804. The absolute difference is 76 aa substitutions. Based on the A/C-60 reference and the data from complete mt genomes, a cow–sheep split of 30 MyrBP has been estimated (Hiendleder et al. 1998). Thus, domestic sheep and bighorn sheep have then diverged at 30 MyrBP × (0.01508/0.0804), which equals 5.63 MyrBP. This value has then been used as a calibration point in order to calculate divergence times within the genus Ovis from distances of the CR sequences. Distances were calculated under the TN-93 model of sequence evolution. The transition/transversion parameter has been estimated by TREE-Puzzle from the dataset to 21.7 and the pyrimidine/purine transition parameter to 0.82. Furthermore, a gamma model (Yang 1994) of rate heterogeneity (I) with four classes of variable sites and one class of invariable sites (I), resulting in an alpha parameter of 0.86 as estimated by TREE-Puzzle, has been taken into account. Table 2 summarizes the distances and divergence times of well-separated lineages.

### 4. DISCUSSION

The taxonomy of wild sheep based on morphological traits such as horn shape and pelage characteristics is often confused (Geist 1991a). Hybrid zones between species (Nadler et al. 1971, 1973a; Valdez et al. 1978) further complicate the systematics of Eurasian wild sheep. Mitochondrial DNA sequence data are therefore expected to help clarify wild sheep taxonomy. Phylogenetic reconstructions placed haplotypes from the argali subspecies *O. ammon collium*, *O. ammon nigrimontana* and *O. ammon ammon*, *O. ammon darwini* in two well-separated branches (figure 4a). This is consistent with geographical origins of the subspecies (table 1) and current IUCN systematics and nomenclature for *O. ammon* ssp. (Shackleton et al. 1997). The Altai argali (*O. ammon ammon*) and the Gobi argali (*O. ammon darwini*) show much similarity, and it has been speculated that size differences could be due to differences in productivity of their ranges (Geist 1991a). Our molecular dating has estimated the time of divergence for *O. canadensis* from all other sheep of the present study at 5.63 MyrBP. Based on this estimate, divergence time for *O. ammon ammon* and *O. ammon darwini* haplotypes was calculated at equal to or more than 1.32 MyrBP (table 2). This is consistent with a subspecies status based on estimates of speciation duration (1.2–3.2 MyrBP) for mammals during the Pliocene/Pleistocene (Avise et al. 1998).

Genetic distance and divergence time for the Altai argali (*O. ammon ammon*) and Gobi argali (*O. ammon darwini*)
Table 2. Number of transitions and transversions observed between mtDNA control region sequences of major sheep lineages and derived distances and divergence times.

<table>
<thead>
<tr>
<th>comparison*</th>
<th>transitions (n ± s.e.)</th>
<th>transversions (n ± s.e.)</th>
<th>distance&lt;sup&gt;b&lt;/sup&gt; (TN93 + I + I)</th>
<th>divergence (MyrBP)</th>
<th>distance&lt;sup&gt;b&lt;/sup&gt; (TN93)</th>
<th>divergence (MyrBP)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ovis canadensis</em> canadensis vs. other sheep</td>
<td>86.2 ± 4.2</td>
<td>12.8 ± 0.5</td>
<td>&lt;0.205&gt;</td>
<td>&lt;5.63&gt;</td>
<td>&lt;0.109&gt;</td>
<td>&lt;5.63&gt;</td>
</tr>
<tr>
<td><em>Ovis ammon</em> collium/<em>Ovis ammon nigromontana</em>/<em>Ovis vignei</em> arkal vs. <em>Ovis aries</em></td>
<td>62.9 ± 5.7</td>
<td>3.1 ± 0.5</td>
<td>0.117 ± 0.007</td>
<td>3.21 ± 0.20</td>
<td>0.070 ± 0.004</td>
<td>3.62 ± 0.21</td>
</tr>
<tr>
<td><em>Ovis ammon</em> ammon/<em>Ovis ammon darwini</em>/<em>Ovis vignei</em> bochaniensis vs. <em>Ovis aries</em></td>
<td>55.8 ± 5.5</td>
<td>3.6 ± 0.9</td>
<td>0.103 ± 0.008</td>
<td>2.83 ± 0.22</td>
<td>0.063 ± 0.005</td>
<td>3.25 ± 0.26</td>
</tr>
<tr>
<td><em>Ovis ammon</em> ammon/<em>Ovis ammon</em> darwini arkal vs. <em>Ovis ammon</em> collium</td>
<td>31.6 ± 1.2</td>
<td>1.0 ± 0.0</td>
<td>0.056 ± 0.002</td>
<td>1.54 ± 0.05</td>
<td>0.033 ± 0.001</td>
<td>1.70 ± 0.05</td>
</tr>
<tr>
<td><em>Ovis ammon</em> ammon vs. <em>Ovis ammon</em> darwini</td>
<td>31.0 ± n.a.&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.0 ± n.a.&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.048 ± n.a.&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.32 ± n.a.&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.032 ± n.a.&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.65 ± n.a.&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td><em>Ovis aries</em> cluster A vs. <em>Ovis aries</em> cluster B</td>
<td>34.4 ± 2.5</td>
<td>0.5 ± 0.8</td>
<td>0.056 ± 0.003</td>
<td>1.54 ± 0.08</td>
<td>0.036 ± 0.002</td>
<td>1.86 ± 0.10</td>
</tr>
<tr>
<td><em>Ovis aries</em> cluster B '64' vs. rest cluster B</td>
<td>8.9 ± 2.2</td>
<td>1.4 ± 0.7</td>
<td>0.016 ± 0.002</td>
<td>0.43 ± 0.05</td>
<td>0.011 ± 0.002</td>
<td>0.57 ± 0.10</td>
</tr>
<tr>
<td><em>Ovis musimon</em> vs. rest <em>Ovis aries</em> cluster B '64'</td>
<td>9.0 ± 2.0</td>
<td>0.8 ± 0.9</td>
<td>0.014 ± 0.004</td>
<td>0.38 ± 0.11</td>
<td>0.009 ± 0.003</td>
<td>0.46 ± 0.15</td>
</tr>
<tr>
<td><em>Ovis aries</em> cluster A '88' vs. rest cluster A</td>
<td>9.6 ± 2.1</td>
<td>0.2 ± 0.6</td>
<td>0.011 ± 0.005</td>
<td>0.30 ± 0.14</td>
<td>0.010 ± 0.001</td>
<td>0.52 ± 0.05</td>
</tr>
<tr>
<td><em>Ovis aries</em> cluster A '73' vs. rest cluster A</td>
<td>5.8 ± 3.5</td>
<td>0.3 ± 0.6</td>
<td>0.010 ± 0.004</td>
<td>0.27 ± 0.11</td>
<td>0.008 ± 0.002</td>
<td>0.41 ± 0.10</td>
</tr>
</tbody>
</table>

* Subclusters within *Ovis aries* clusters A and B are labelled according to their bootstrap values as shown in figure 4b.
<sup>b</sup> Distances were calculated assuming a TN93 model of sequence evolution, plus a gamma (I) model of rate heterogeneity with four classes of variable sites, plus one class of invariable sites (I) and assuming rate homogeneity.
<sup>c</sup> n.a., not applicable.
haplotypes is very similar to that for two other subspecies haplotypes, northern Kazakhstan argali (O. ammon collium) and Kara Tau argali (O. ammon nigrimontana). According to Geist (1991a), northern Kazakhstan argali (O. ammon collium) is a subpopulation of the Tien Shan Argali (O. ammon karelini) and the Kara Tau argali (O. ammon nigrimontana), which is possibly a small desert derivative of that group. This hypothesis is supported by the distances and tree analysis of the mt CR haplotypes. Unexpectedly, the haplotype of the Arkal sheep specimen from the Ust-Urt region showed a close relationship with Kara Tau argali (O. ammon nigrimontana) (figure 4a; table 2). Until now, sheep of this region have been placed in the urial group and, accordingly, named O. vignei arkal. However, with regard to neck ruff, saddle patch and horn curl (Valdez 1982), Ust-Urt sheep show phenotypic similarities with O. vignei severzovi that occurs further to the southeast in the Kizil-Kum desert and Nura-Tau range (Shackleton et al. 1997). The status of O. vignei severzovi has long been controversial, and, based on a chromosome number of 2n = 56 (Bunch et al. 1998), it was recently reclassified as an argali, O. ammon severzovi. This raises

Figure 2. (First part of figure opposite) Complete alignment of variable nt positions in mt control region sequences of the genus Ovis. Dots show identity with the first sequence, deletions/insertions are represented by dashes. Nucleotide numbering is according to the complete mt genome of the domestic sheep (Hiendleder et al. 1998b). Insertions relative to the numbering of the reference sequence are indicated by alphabetical suffixes. Only two 75/76 nt tandem repeat sequences as found in O. canadensis have been retained. The presence or absence of two to four additional copies of the tandem repeat sequence is indicated by a corresponding number of asterisks/dashes.

### Table: Sheep Evolution

<table>
<thead>
<tr>
<th>Location</th>
<th>Species</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spain</td>
<td>O. aries</td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>O. aries</td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>O. aries</td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>O. aries</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>O. aries</td>
<td></td>
</tr>
<tr>
<td>Tadzhikistan</td>
<td>O. aries</td>
<td></td>
</tr>
<tr>
<td>O. ammon collium</td>
<td>O. ammon nigrimontana</td>
<td></td>
</tr>
<tr>
<td>O. canadensis</td>
<td>O. canadensis canadensis</td>
<td></td>
</tr>
</tbody>
</table>

### Diagram:
- The diagram illustrates the evolutionary relationships among different species of sheep across various regions.
- Each cell represents a genetic marker or mutation.
- The colors and symbols indicate the presence or absence of specific markers.
- The diagram is used to track the spread and characteristics of different sheep species across the listed regions.
the possibility that Ust-Urt sheep also belong to the argali group. Another explanation would be a hybrid origin of Ust-Urt sheep due to overlapping ranges of urial and argali, like the hybrid zones of mouflon and urial in Iran (Nadler et al. 1971; Valdez et al. 1978), as the former range of Ust-Urt sheep could have merged with the range of O. ammon severtzovi (Nadler et al. 1973a). Unfortunately, chromosome data of Ust-Urt sheep that could help to clarify this question are not available.

The Ust-Urt sheep haplotype is clearly separated from haplotypes obtained for the Bukhara urial (O. vignei bochariensis), a true urial with 2n = 58 (Nadler et al. 1973a), which is found in southeastern Turkmenistan, southern Uzbekistan and Tadjikistan. The bochariensis urial specimen is placed in a branch with O. ammon ammon, O. ammon darwini and O. canadensis canadensis. Nadler et al. (1973a) believed that urial sheep with 2n = 58 were the most primitive form, ancestral to all other sheep with lower chromosome numbers, considering that fusion rather than fission is the predominant mechanism for changes in diploid number in cattle, sheep and goats. The mtDNA-based phylogenetic reconstructions show that O. vignei bochariensis haplotypes are indeed basal to mouflon/domestic sheep and the O. ammon collium/O. ammon nigrimontana/Ust-Urt sheep branch. However, branching order in the O. canadensis canadensis/O. ammon ammon/O. ammon darwini/O. vignei bochariensis branch is not completely resolved. The molecular data on O. musimon, O. vignei bochariensis and Ust-Urt wild sheep clearly show that they cannot, as was done previously (Valdez et al. 1978; Valdez 1982; Shackleton et al. 1997), be lumped into a single ‘moufloniform’ (O. orientalis) branch of sheep. Instead, they are members of three distinct clades and, considering the estimated distances and divergence times, could be different species. Similarly, the CR distances between the two clusters with argali sheep haplotypes are comparable with those between well-established species like that of harbour and grey seal (Arnason et al. 1993) or pygmy and common chimp (Arnason et al. 1996). This and the resulting divergence times indicate that the two argali matrilines are in the process of speciation or already belong to different species.

In domestic animals, bi- or multiphyletic origins are not uncommon. Two distinct mtDNA lineages, an Asian and European clade, have been identified in cattle (Loftus et al. 1994; Bradley et al. 1996) and swine (Giulifra et al. 2000). In the dog (Tsuda et al. 1997; Vila et al. 1997), horse (Vila et al. 2001) and goat (Luikart et al. 2001), several clusters implying multiple integrations of matrilines have been described. The degree of sequence divergence between the two major mt lineages in bovines has led to the hypothesis of independent domestication from two different subspecies of Bos primigenius, giving rise to

---

**Figure 3.** (Opposite) Alignment of proposed regulatory elements (Wood & Phua 1996; Hiendleder et al. 1998a,b) in Ovis CR. (a) The 3’ tandem repeat sequence of 75 or 76 nt with putative D-loop strand termination site (dotted line box) and octamer sequences (solid boxes) that form the stem of predicted 39 nt secondary structures (–8.2 kcal mol⁻¹). The loop region of these structures is overlined. (b) TAS-A element. (c) CSB 1. (d) CSB 2 + 3. Nt numbering above sequence elements is according to the complete mt genome of the domestic sheep (Hiendleder et al. 1998b).

**Figure 4.** Phylogenetic relationships among wild and domestic forms of the genus Ovis. (a) Maximum likelihood tree of mtDNA CR sequences of all 61 unique sequences as reconstructed by TREE-PUZZLE v. 4.0. Branches with less than 50% quartet puzzling (QP) support have been collapsed. (b) Fourfold enlargement of O. aries and O. aries/O. musimon clusters A and B. QP support values are reported on the nodes. (Origin of O. aries: open circle, New Zealand; filled circle, Europe; filled square, Near East (Syria); filled diamond, Asia Minor (Turkey); open diamond, Central Asia (Kazakhstan, Tadjikistan); O. m., Ovis musimon.)
New Zealand O.aries (1963) that one line of domesticated sheep derived from Syria and Kazakhstan is only 64%. The position of that also encompasses (1999). The sheep mtDNA clades (Hiendleder have led to the proposal of European and Asian domestic sheep mtDNA lineages A and B (Hiendleder et al. 1998a). Domestic sheep samples from central Asia and the near East that have been investigated in the present study support this view, the near Eastern samples clustering with the European clade. The relatively high incidence of haplotype A in New Zealand breeds is presumably a relic of the fat-tailed sheep from India, which were the first sheep in Australasia (Ryder 1984), because a male-mediated upgrading of fat-tailed sheep with rams of European breeds leaves mtDNA unaffected. As in other major farm-animal species (MacHugh & Bradley 2001), the two well-separated O. aries mtDNA lineages A and B (figure 4a,b) detected among individuals from sheep breeds with different geographical origins suggest domestication from two distinct wild populations. Alternatively, domestic sheep could have originated from a single polymorphic ancestral population, where two major haplotypes became fixed through 'lineage sorting' (Avise et al. 1987). Although the latter possibility cannot be completely dismissed with the current dataset, the first alternative seems more likely, considering the estimated divergence times for clusters A and B, which are far in excess of the ca. 11 000 years (Ryder 1984) of domestication history, and the striking similarities in branching patterns observed in cattle, pig and water buffalo mtDNA (MacHugh & Bradley 2001). Estimates on the divergence time for clusters A and B (equal to or more than 1.54 MyrBP) are similar to the differences found between argali subspecies (table 2). The star-like branching patterns and short branches within the two clusters of domestic sheep. As tree topology (figure 4a), distances and divergence times (table 2) incompatible with any contribution to domestic sheep lineages. Furthermore, CR data from another sampled urial population (O. vignet cycloceros) are as much diverged from O. aries as O. vignet bochariensis (data not shown). This excludes urial (O. vignet ssp.) as a matriline source for cluster A of domestic sheep. As tree topology (figure 4a), distances and divergence times (table 2) also exclude argali (O. ammon ssp.) as maternal ancestors of O. aries, the origin of haplogroup A remains unknown. Considering the probable subspecies relationship of the founders of both clusters of domestic sheep lineages, mouflon (O. orientalis) populations of the eastern mouflon range are probable candidates.

We are grateful to N. Koshum, K. Mainz, K. Plakhov, Y. Plante, G. Trinkaus, M. Turarbekov and K. Vollmer who supplied or helped obtain wild and domestic sheep samples and thank U. Arnason for critical comments and helpful discussions. This work was kindly supported by a grant from the H.-Wilhelm Schaumann Stiftung zu Hamburg to S.H.

REFERENCES
Arnason, U. & Gullberg, A. 1993 Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. J. Mol. Evol. 37, 312–322.


Hendleder, S., Mainz, K., Plante, Y. & Lewalski, H. 1998a Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources. No evidence for contributions from Urial and Argali sheep. J. Heredity 89, 113–120.


Nasonov, N. 1923 Geographic selection dikh baranov starogo sveta (Distribution géographique des moutons sauvages du monde ancien). St Petersburg (Leningrad), Russia: Akad. Nauk SSR.


Ursing, B. M., Slack, K. E. & Arnason, U. 2000 Subordinal...
artiodactyl relationships in the light of phylogenetic analysis
of 12 mitochondrial protein-coding genes. Zool. Scripta 29,
83–88.
Valdez, R. 1982 The wild sheep of the world. Mesilla, NM: Wild
Sheep and Goat International.
Valdez, R., Nadler, C. F. & Bunch, T. D. 1978 Evolution of
Vigne, J.-D. 1999 The large ‘true’ Mediterranean islands as
a model for the Holocene human impact on the European
vertebrate fauna? Recent data and new reflections. In The
Holocene history of the European vertebrate fauna. Modern
Rahden/Westf, Germany: Verlag Marie Leidorf GmbH.
Vila, C., Savoleinen, P., Maldonado, J. E., Amorim, I. R.,
Rice, J. E., Honeycutt, R. L., Crandall, K. A., Lundberg,
J. & Wayne, R. K. 1997 Multiple and ancient origins of the
Vila, C., Leonard, J. A., Götherström, A., Marklund, S.,
Widespread origins of domestic horse lineages. Science 291,
474–477.
Vorontsov, N. N., Korobitsyna, K. V., Nadler, C. F., Hofman,
R., Sapozhnikov, G. N. & Gorchlov, J. K. 1972 Chromos-
somi dikich baranow i proiskhzhdenie domashnikh ovets.
Priroda, Moscow 3, 74–82.
27, 25–33.
Xu, X. & Arnason, U. 1994 The complete mitochondrial DNA
sequence of the horse, Equus caballus; extensive hetero-
Xu, X., Gullberg, A. & Arnason, U. 1996 The complete mito-
chondrial DNA (mtDNA) of the donkey and mtDNA com-
parisons among four closely related mammalian species-
Yang, Z. 1994 Estimating the pattern of nucleotide substi-
Zardoya, R., Villalta, M., Lopez-Perez, M. J., Garrido-
Pertierra, A., Montoya, J. & Bautista, J. M. 1995 Nucleotide
sequence of the sheep mitochondrial D-loop and its flanking

As this paper exceeds the maximum length normally permitted, the
authors have agreed to contribute to production costs.