The Mitochondrial Control Region of Cervidae: Evolutionary Patterns and Phylogenetic Content

Emmanuel Douzery and Ettore Randi†*

*Laboratoire de Paléontologie, Paléobiologie et Phylogénie Institut des Sciences de l'Evolution, CNRS, Université Montpellier II, France; and †Istituto Nazionale per la Fauna Selvatica, Ozzano dell'Emilia, Italy

The mitochondrial control region (CR) sequence, also known as the D-loop, has been determined for six Cervidae (Artiodactyla, Ruminantia): the red and fallow deers (subfamily Cervinae), the brocket deer and two roe deers (subfamily Odocoileinae), and the Chinese water deer (Hydropotinae). These new sequences have been aligned with available cervid and bovid orthologues. Comparative analyses indicate that the 5′-peripheral domain exhibits a 75-bp length polymorphism near sequences associated with the termination of the H-strand replication. The New World Odocoileinae possess the longest cervid CR due to the presence of an additional 47-bp tandem repeat, located in the 3′-peripheral domain, downstream of the initiation site for H-strand replication (O_H) and the first conserved sequence block (CSB-1). This insertion represents a duplication spanning the O_H to CSB-1 region and constitutes an exclusive synapomorphy for New World Odocoileinae. Phylogenetic analyses of the complete CR support the paraphyly of antlered deers due to the nesting of the antlerless *Hydropotes* within Odocoileinae. *Capreolus* is the closest relative of *Hydropotes,* and the divergence of this Old World Odocoileinae clade may have occurred between 8.7 and 10.4 MYA. The conserved central domain of CR can be aligned across ungulates and indicates the Pecora monophyly, their close association with cetaceans, and the earlier emergence of suiformes.

Introduction

The control region (CR), the main noncoding region of metazoan mitochondrial DNA (mtDNA), is bounded in mammals by the transfer RNA genes t-RNA^{Pro} on the 5′ side of the light (L)-strand and t-RNA^{Phe} on the 3′ side of the heavy (H)-strand. Comparative analyses indicate that the CR is highly structured (Saccone, Pesole, and Sbisá 1991), with a conserved central region (CCR) flanked by two highly divergent peripheral domains (5′-left and 3′-right). The CR contains sequences which control the replication of the H-strand and the transcription of mtDNA (Clayton 1992). The promoters for H- and L-strand transcription initiation (HSP and LSP) are located in the right domain (fig. 1). After its elongation, the L-strand transcript is cleaved by the RNase MRP (for mitochondrial RNA processing) at conserved sequence blocks, (CSBs) 1, 2, and 3, with the last two being apparently fused among ruminant artiodactyls (Clayton 1992; Dairaghi and Clayton 1993; Ghivizzani et al. 1993). The 5′ end of the L-strand transcript subsequently primes the H-strand replication. The nascent H-strand syntheses initiate at the unidirectional origin O_H, and most of them end near a termination-associated sequence (TAS-1 or TAS-A) in the left domain (Doda, Wright, and Clayton 1981; MacKay et al. 1986).

Strong selective constraints seem to act on the CCR, which evolves slowly, mainly through point mutations and low rates of accumulation of short indels (Saccone, Pesole, and Sbisá 1991; Arnason, Gullberg, and Widegren 1993). In contrast, the peripheral domains quickly accumulate point mutations, indels, and variable numbers of tandem repeats (VNTRs). Individual heteroplasmy (Ghivizzani et al. 1993) and extensive length differences of the mtDNA molecule can be produced among species. In fact, the VNTR structure of short repeats (e.g., Arnason and Johnsson 1992) as well as long repeats (e.g., Wilkinson and Chapman 1991) can produce arrays several hundreds of base pairs (bp) long. Repetitive sequences (RSs) have been found in different regions of both peripheral domains of vertebrate mtDNA CRs, most of them occuring near the TAS (i.e., RS2 location) and between CSB-1 and CSB-2 (i.e., RS3 location; Hoelzel et al. 1994, fig. 1). VNTRs can be generated through different mechanisms of intra- or intermolecular recombination (Rand and Harrison 1989) or strand slippage during DNA replication (Levinson and Gutman 1987). After the initial generation of short repeats, asymmetrical crossing over or secondary structures of single-stranded DNA can accelerate the formation of longer VNTRs (Levinson and Gutman 1987).

Because of their high rate of nucleotide change in peripheral regions, CR sequences are widely used in population and evolutionary genetics of mammals. The 5′-peripheral domain has been especially studied in the artiodactyl family Bovidae: cattle (Loftus et al. 1994), sheep (Wood and Phua 1996), Grant's gazelles (Arctander et al. 1996*a*), and impalas (Arctander et al. 1996*b*). The whole CR provided phylogenetic information within placental orders (see the case of Cetacea; Arnason, Gullberg, and Widegren 1993). The more slowly evolving CCR has been used to reconstruct phylogenies of placental orders (Saccone, Pesole, and Sbisá 1991).

In this paper we explore the molecular evolution and phylogenetic content of the mitochondrial CR of the family Cervidae (order Artiodactyla, suborder Ruminantia, infraorder Pecora). The CRs compared in this study are from 10 cervid representatives belonging to the subfamilies (1) Cervinae (Old World genera): *Cervus elaphus* (red deer; two individuals), *Cervus nippon* (sika

Key words: mitochondrial DNA, control region, D-loop, phylogeny, Cervidae, tandem repeats.

Address for correspondence and reprints: Emmanuel Douzery, Laboratoire de Paléontologie, Paléobiologie et Phylogénie, Institut des Sciences de l'Evolution, UMR 5554 CNRS, Université Montpellier II, CC064, Place E. Bataillon, 34 095 Montpellier Cedex 05, France. E-mail: douzery@isem.univ-montp2.fr.

Mol. Biol. Evol. 14(11):1154–1166. 1997

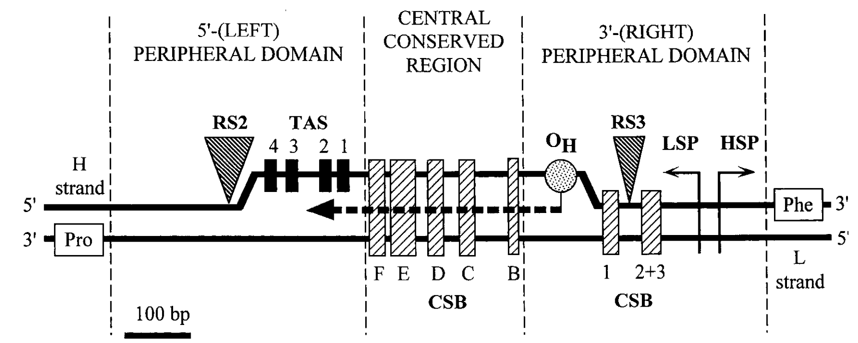


FIG. 1.—Schematic diagram of the organization of the mammalian mitochondrial DNA control region (CR) and its displacement loop structure (D-loop), with special reference to the Cervidae. The CR is bounded by the tRNA^{Pro} and tRNA^{Phe} genes. Comparative analyses across mammalian CRs evidenced three regions: a central conserved region (CCR) flanked by two highly divergent peripheral domains (5′ or left, and 3′ or right). The right domain contains the promoters for H- and L-strand transcription initiation (HSP and LSP). The L-strand transcript is cleaved at conserved sequence blocks (CSB-1 and CSB 2 + 3 among ruminant artiodactyls). Its 5′ end primes the H-strand replication, which subsequently initiates at the O_H origin. The nascent H-strand is represented by the dashed gray arrow: it displaces the parental H-strand creating the D-loop triplex structure. The daughter H-strand ends near a termination-associated sequence (TAS-1) in the left domain. Sequence similarities across the CRs of cervids yield to the distinction of three additional TAS motifs (TAS-2, TAS-3, and TAS-4) and to the recognition of five CSBs (B–F) in the CCR. Tandem repeats occur among mammalian CRs in both peripheral domains. In the case of Cervidae, the RS3 tandem repeat (47 bp long) is a synapomorphy of New World odocoileines (Odocoileini). The RS2 insertion is 75 bp long and convergently acquired between Odocoileini and the sika deer.

deer), *Dama dama* (fallow deer); (2) Hydropotinae: *Hydropotes inermis* (Chinese water deer); (3) Odocoileinae, with the tribes Odocoileini (New World genera): *Mazama* sp. (brocket deer), *Odocoileus hemionus* (mule deer) and *Odocoileus virginianus* (white-tailed deer); and Capreolini (Old World genera): *Capreolus capreolus* (European roe deer) and *Capreolus pygargus* (Siberian roe deer). The taxonomic arrangement is from Eisenberg (1981, p. 200) and Scott and Janis (1987). No representative was available for either of the subtribes Alcini (mooses) and Rangiferini (reindeers) or for the subfamily Muntiacinae (muntjacs). Data obtained in this study provide the opportunity to examine the pattern of accumulation of mutations in the mitochondrial major noncoding region by comparing sequences of increasingly divergent pecoran taxa. The phylogenetic content of the CR is therefore explored, in order to test for monophyly of the antlered subfamilies Cervinae and Odocoileinae, for the postulated basal position of the antlerless Hydropotinae within Cervidae (Groves and Grubb 1987; Miyamoto, Kraus, and Ryder 1990) and for affinities between Odocoileinae and Hydropotinae (Bouvrain, Geraads, and Jehenne 1989).

Materials and Methods

Tissue samples of *Dama dama* (fallow deer), *Capreolus capreolus* (European roe deer) and *C. pygargus* (Siberian roe deer) were collected, respectively, in the Castel Porziano Reserve population (Roma), in the western Italian Alps (Asiago), and by Aleksey Danilkin (Moscow) in the Amur region. Tissue samples of *Hydropotes inermis* were kindly provided by Oliver Ryder (San Diego Zoo) and maintained in the collection of tissues of the laboratory of Paleontology, Paleobiology, and Phylogeny in Montpellier (Catzeflis 1991).

Total DNA was extracted from 95% ethanol-preserved liver following standard procedures (Sambrook, Fritsch, and Maniatis 1989). DNAs for *Cervus elaphus* and *Mazama* sp. (it was not possible to distinguish between the red brocket *M. americana* and the brown brocket *M. gouazoupira*) were extracted from the same tissues used by Douzery, Lebreton, and Catzeflis (1995).

The entire roe deer and fallow deer mtDNA CRs were PCR amplified (Mullis et al. 1986) using the primers L-Pro 5′-CGTCAGTCTCACCATCAACCCCCAAAGC-3′ and H-Phe 5′-GGGAGACTCATCTAGGCATTTTCAGTG-3′ (Jager, Hecht, and Herzog 1992, unpublished data), which bind to positions 15740 and 420 of the bovine tRNA^{Pro} and tRNA^{Phe} (Anderson et al. 1982). Amplifications were performed with AmpliTaq DNA polymerase (Perkin Elmer), with 3mM $MgCl_2$ in the reaction buffer, and using the following thermal cycles in a 9600 Perkin Elmer machine: 94°C for 2 min; 94°C for 15 s, 55°C for 15 s, 72°C for 1 min (30 cycles); 72°C for 10 min. PCR products were purified by low-melting agarose gels. All sequences were obtained by double-stranded DNA cycle sequencing with ΔTaq Sequenase (Amersham), using the external L-Pro and H-Phe primers, with the internal L-362 5′-AATCACCATGCCGCGTGAAACC-3′ and H-493 5′-TGAGATGGCCCTGAAGAAAGAACC-3′ primers binding to the middle of the CCR of ungulates (unpublished data).

For the red deer, brocket deer, and Chinese water deer, a 3.0-kb segment of mtDNA spanning the contiguous cytochrome *b*, tRNA^{Thr}, complementary tRNA^{Pro}, CR, tRNA^{Phe}, and first half of 12S rDNA sequences was PCR amplified. The two primers used were L14724B (Irwin, Kocher, and Wilson 1991) and S3 5′-AACTGGGATTAGATACCCCACTATG-3′, annealing, respectively, in the flanking tRNA^{Glu} region and in the central conserved region of the 12S rDNA. Amplifications were performed with Taq DNA polymerase A.T.G.C. (Noisy le Grand, France), with 1.5 mM $MgCl_2$ in the reaction buffer, and using the following thermal cycles in a Crocodile II Appligene machine: 94°C for 5

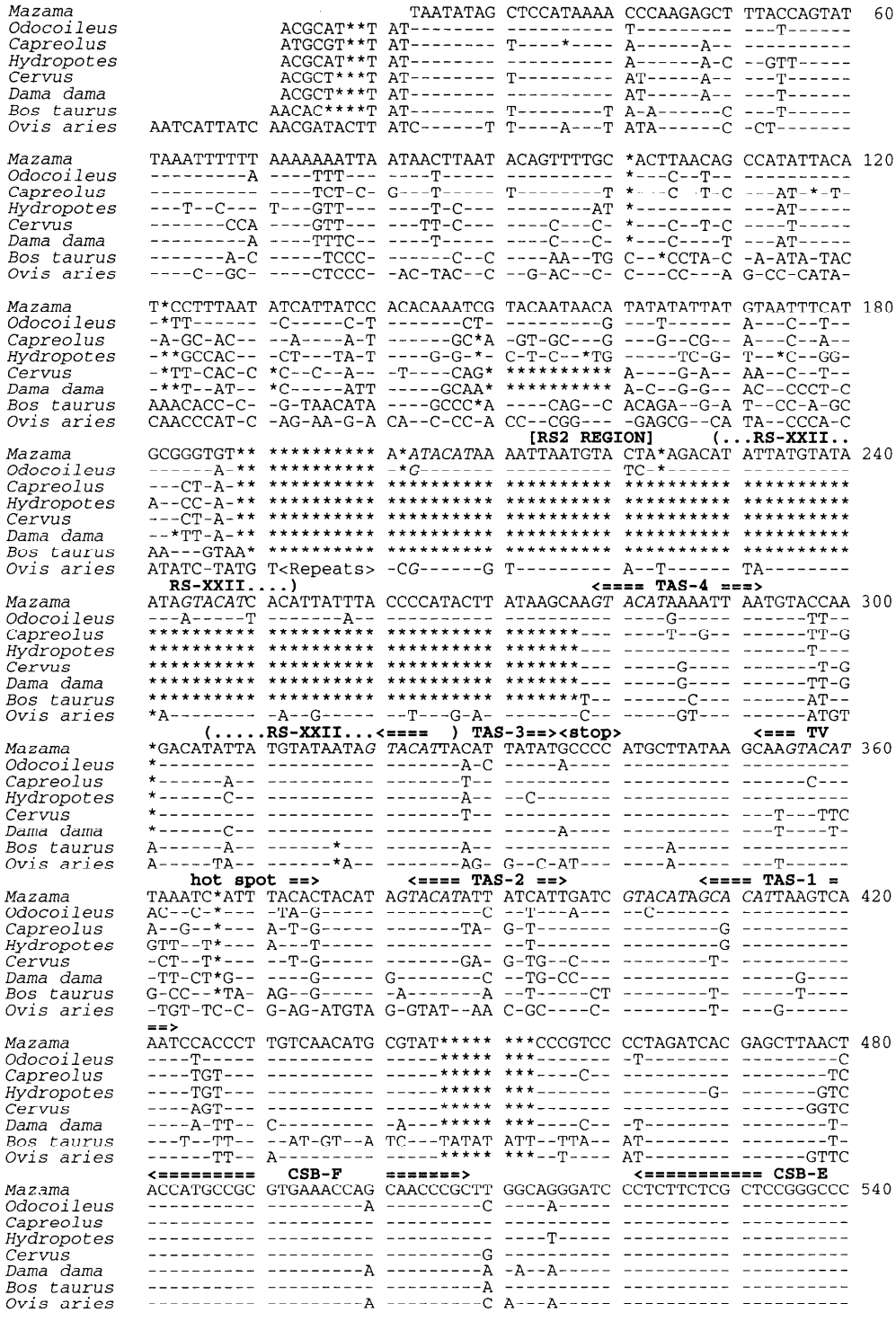
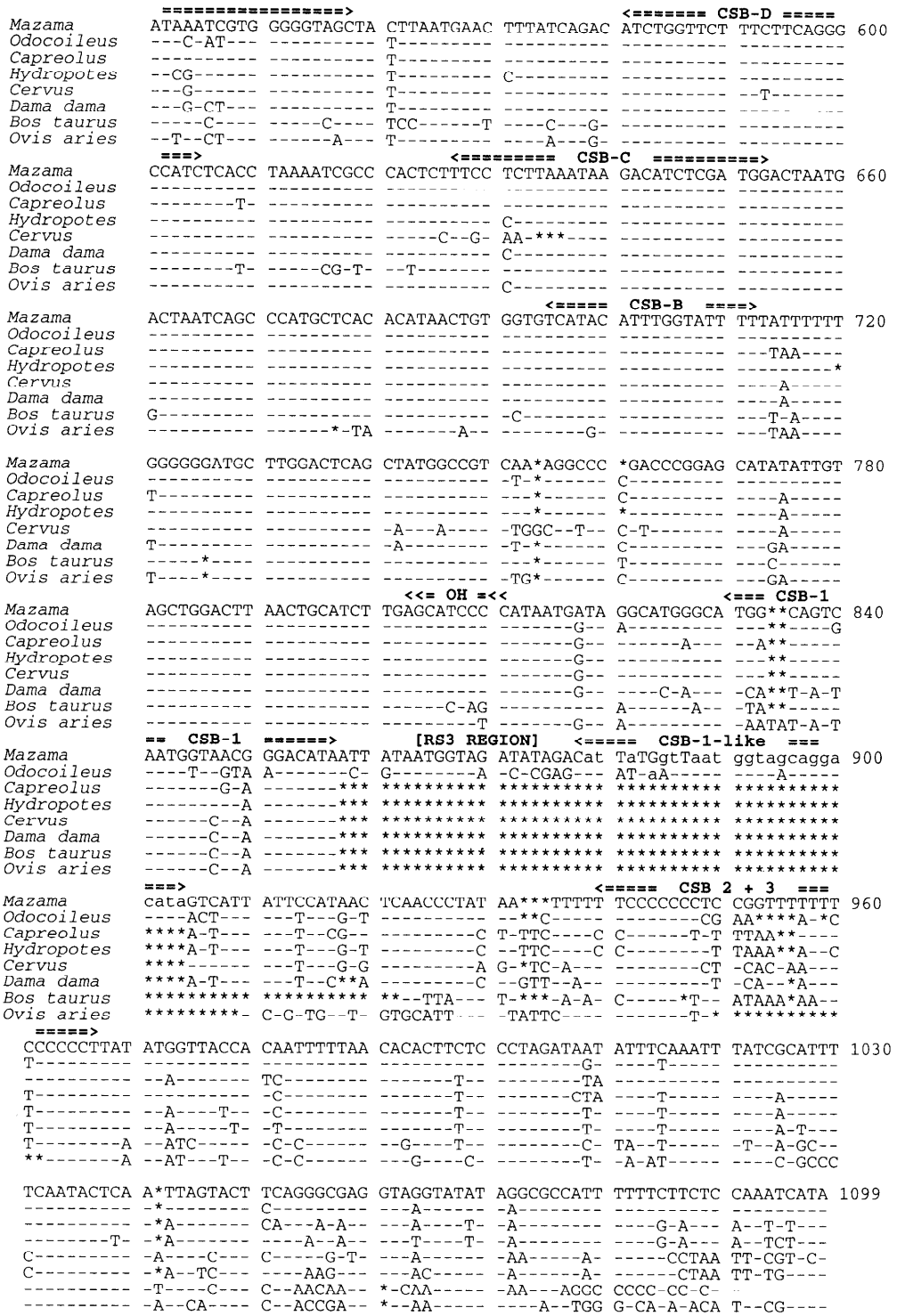


FIG. 2.—Alignment of the control region of six cervids and two outgroup bovids. By comparison to the *Mazama* sp. reference sequence, dashes indicate identity of nucleotide, and stars denote deletions. The termination-associated sequences (TAS-1–4, with TAS-1 corresponding to TAS-A in Madsen, Ghivizzani, and Hauswirth 1993), the conserved sequence blocks (CSBs B–F, 1, and 2 + 3, defined by Anderson et al. 1982; Hoelzel, Hancock, and Dovcr 1991; Ghivizzani et al. 1993), and the hot spot of transversions of *Capreolus* (TV hot spot, defined by Randi, Pierpaoli, and Danilkin, unpublished data) are indicated by <==>. The putative point of arrest of the D-loop synthesis is indicated as stop. The origin of replication of the H-strand (O_H) and the sense of replication are indicated by <= O_H =<. In the RS3 region, the insertion of the Odocoileini includes a partially duplicated CSB-1 segment denoted as CSB-1-like sequence, and the identical nucleotides are in small letters. The repeated GYRCAT motifs in the RS2 and TAS regions are indicated in italics. The RS-XXII motif and its duplication are indicated in brackets. The *Ovis aries* sequence is truncated in the RS2 region: the first three 75-nt-long repeats are indicated by <repeats> but are not aligned, and the numbering of the alignement does not take into account these segments (1,099 of the 1,264 positions are represented). One can note that the sheep CR is difficult to align in some areas where other cervid and bovid CR sequences are conserved. For practical sequencing reasons, the first 9 bp of the CR of *Mazama* is missing. For the last two blocks of 70 nt, the order of the species is exactly the same as previously. *Odocoileus* is for *O. hemionus, Capreolus* is for *C. capreolus, Hydropotes* is for *H. inermis,* and *Cervus* is for *C. elaphus.* The *Bos*



←

taurus sequence is from Anderson et al. (1982), the *Ovis aries* sequence is from Wood and Phua (1996), and the *Odocoileus hemionus* sequence is from Feng, Li, Rittenhouse, and Templeton (unpublished; EMBL accession number U12865). The new nucleotide sequences reported in this paper appear in the EMBL, GenBank and DDBJ nucleotide sequence databases under the accession numbers Y08207–Y08209, Z70317, and Z70318.

min; 94°C for 45 s, 55°C for 45 s, 72°C for 3 min 30 s (30 cycles); 72°C for 10 min. PCR products were cloned in the plasmid pGEM-T (Promega), and transformations were done in *Escherichia coli* strain JM109 competent cells. Three positive clones were isolated and the complete mitochondrial CR was dideoxy-sequenced on both strands (Sanger, Nicklen, and Coulson 1977) using the Pharmacia kit, with the internal direct L0 = 5′-CCCAAAGCTGAAATTCTACTTAAACTA-3′ and D1 = 5′-CATCTGGTTCTTACTTCAGG-3′ and reverse E1 = 5′-AATCAGCCCATGCTCACACATAAC-3′ and S0 = 5′-TCTAGGCATTTTCAGTGCCTTGCTTT-3′ primers.

The six new CR sequences from *Cervus elaphus, Dama dama, Mazama* sp., *Hydropotes inermis, Capreolus capreolus,* and *C. pygargus* were compared with orthologues already published or deposited in the data banks. The CR sequences of the following taxa were available: *Cervus elaphus* (another individual), *Cervus nippon, Odocoileus hemionus,* and *Odocoileus virginianus* for cervids (Feng, Li, Rittenhouse, and Templeton, unpublished paper; accession numbers U12867, U12868, U12865, U12869 in the EMBL data bank) and *Bos taurus* (cattle; V00654/J01394 by Anderson et al. 1982) and *Ovis aries* (sheep; Z35237 by Wood and Phua 1996) for the bovid outgroups. The CR sequences were aligned by eye and CLUSTAL W (Thompson, Higgins, and Gibson 1994). In highly variable regions, gaps were introduced when they saved at least two substitutions, and they were grouped to maximize the overall similarity. Long indels (75 and 47 bp) were excluded from subsequent phylogenetic analyses, and short indels were coded as fifth character states at each site where they occured. Five other ungulate CR sequences were aligned in their CCR: one artiodactyl suiforme (the pig, *Sus scrofa*: D17739 by Takeda et al. 1995), two cetaceans (a toothed-whale, *Orcinus orca*: M60409 by Hoelzel, Hancock, and Dover 1991; and a baleen-whale, *Balaenoptera physalus*: X61145 by Arnason, Gullberg, and Widegren 1991), two perissodactyls (a rhinocerotid, *Diceros bicornis*: L22010 by Jama et al. 1993; and an equid, *Equus caballus*: X79547 by Xu and Arnason 1994). The alignments were managed with the MUST package (Philippe 1993).

Presence and location of direct repeats and of possible secondary structures of single-strand DNA (hairpins) and transcribed primary RNA (cloverleaves) were analyzed with the method of Zuker and Stiegler (1981) using the programs REPEAT, HAIRPIN, and RNAFOLD in PCGENE (IntelliGenetics, Oxford). Phylogenetic reconstructions were conducted by the neighbor-joining (NJ) method (Saitou and Nei 1987) on percent divergences, by the maximum-parsimony (MP) method (PAUP 3.1.1.; Swofford 1993), and by the maximum-likelihood (ML) method (DNAML version 3.5c; Felsenstein 1993). Robustness of the phylogenies was assessed by the bootstrap method (Felsenstein 1985), with 1,000 resamplings followed by a distance (NJBOOT program, Philippe 1993) or an MP reconstruction (bootstrap option in PAUP 3.1.1.). Robustness was also computed by decay indices (DIs), i.e., the number of extra steps required to break the corresponding node (Bremer 1988), enforcing topological constraints with PAUP 3.1.1.

Results

Alignment and Conservation of TAS and CSB

The alignment of the 12 CRs is 1,099 nt long (fig. 2). The mean CR length is 968 ± 59 nt for the 10 cervids of this study. *Cervus elaphus* exhibits the shortest CR (less than 920 nt), whereas *Odocoileus* and *Mazama* possess the longest CRs (nearly 1,050 nt) due to the presence of two insertions in the RS2 and RS3 regions (fig. 1). *Mazama,* the two *Odocoileus* species and *Cervus nippon* possess a single 75-bp-long insertion in the RS2 region (fig. 1). This inserted segment is absent in the other Cervidae but is easily alignable with each of the four 75-bp-long tandem repeats of the bovid *Ovis aries* (positions 201–277 in fig. 2). The Odocoileini (*Mazama* and *Odocoileus*) possess a second single insertion segment, 47 bp long (positions 858–904 in fig. 2), located downstream of CSB-1 in the RS3 region (fig. 1).

The TAS-1 motif (Fig. 1, positions 408–423 in fig. 2) has been functionally associated with the termination of the D-loop (TAS-A in cattle: Madsen, Ghivizzani, and Hauswirth 1993). The 3′ end of the D-loop is located 70 nt upstream of TAS-1, i.e., downstream for the new H-strand, at a TCCCC element for pig and a GCCCC element for cattle (MacKay et al. 1986; Madsen, Ghivizzani, and Hauswirth 1993). The same GCCCC motif is present in the CR of Cervidae (positions 336–339 in fig. 2), except for *Dama* and *Odocoileus* (ACCCC motif).

Three 15-nt-long sequences with high homology to TAS-A of cattle and sheep (Madsen, Ghivizzani, and Hauswirth 1993; Wood and Phua 1996) can be identified in the 5′-peripheral domain of the cervid CR between the RS2 and CSB-F regions (TAS-2, TAS-3, and TAS-4, positions 382–397, 320–335, and 279–294 in fig. 2). The TAS-3 and TAS-4 cervid sequences correspond to the TAS-E and TAS-G identified in cattle by Madsen, Ghivizzani and Hauswirth (1993). The first six nucleotides of these three TAS-like domains are highly conserved in the CR of Cervidae and constitute a GYRCAT motif (G[C/T][A/G]CAT), which is duplicated immediately upstream of TAS-1 (positions 401–406 in fig. 2). Such a tandem structure of two GYRCAT motifs is again associated with TAS-2. Another GYRCAT block is found at the beginning of a transversional hot spot (positions 355–360 in fig. 2) but is less conserved among Cervidae.

The region spanning the TAS-4 to the transversional hot spot is fully duplicated in a single 75-bp-long RS2 segment in *Mazama* sp., *Odocoileus virginianus, O. hemionus,* and *Cervus nippon* (positions 201–277 and 278–353 in fig. 2). The two tandem repeats are characterized by the presence of a 22-nt-long motif, called RS-XXII (positions 230–251 and 306–327 in fig. 2), which has the ability to form a stable DNA single-strand hairpin (fig. 3).

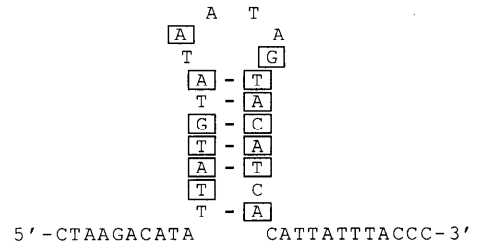


FIG. 3.—Potential secondary hairpin structure in RS-XXII region of the mitochondrial DNA of *Mazama* sp. This hairpin has a free energy value of −5.4 kJ/mol and corresponds to nucleotides 232–251 (fig. 2). Similar hairpins can be formed in the RS2 region of *Cervus nippon, Odocoileus virginianus,* and *O. hemionus.* Invariant positions across the RS-XXII motif of brocket deer, mule deer, white-tailed deer, sika deer, and sheep are boxed. They include six consecutive sites constituting the GYRCAT motif. Hairpins could represent critical regions of DNA slippage during replication and could generate tandem repeats.

Downstream of TAS-1, the CCR is highly conserved within the family Cervidae and extends from position 481 to position 720 (fig. 2). The CSBs are easily recognized (CSB-F to CSB-B; figs. 1 and 2) and show high similarities between the cervid and bovid taxa compared. The only exception is CSB-C because of the presence of a 2- (*Cervus nippon*) to 3-nt-long (*C. elaphus*) deletion characterizing the genus *Cervus* (positions 634–636 in fig. 2). In the CCR, most of the nucleotide substitutions are tightly clustered in a few regions (i.e., positions 510–515, 543–547, and 629–632 in fig. 2).

In the 3′-peripheral domain of the cervid CR, we found the putative initiation site for H-strand replication (O_H: fig. 2) by comparison with the two bovid sequences (cow and sheep: Wood and Phua 1996). This site is strictly conserved in all 10 Cervidae so far studied. The sequence corresponding to the CSB-1 of cow (Anderson et al. 1982) and sheep (Wood and Phua 1996) has been mapped 17 nt downstream of O_H (positions 830–847 in fig. 2). The 3′ end of CSB-1 is very conserved in Cervidae and has a sequence similar to the GYRCAT motif (fig. 4). This 3′-terminal GGACATA sequence of CSB-1 is exactly repeated in the terminal part of Odocoileini RS3 (brocket, mule, and white-tailed deer; fig. 4). Actually, the CSB-1 has been imperfectly copied in a CSB-1-like motif to partially generate the insert of 47 bp of *Mazama* and *Odocoileus* in the RS3 region (figs. 2 and 4). Moreover, there are strong similarities between the RS3 insertion and the region between O_H and CSB-1 for the three Odocoileini taxa (mean 74.1% similarity over 47 nt). The Odocoileini insertion therefore represents a duplicated element corresponding to the region spanning the first nucleotide after the O_H to the last of the CSB-1.

The pig (nonruminant Artiodactyla; Ghivizzani et al. 1993), the horse (Perissodactyla; Xu and Arnason 1994), and most mammals studied so far (Saccone, Pesole and Sbisá 1991) possess the two other separated conserved blocks, CSB-2 and CSB-3, located in the middle of the 3′-peripheral domain. However, cetaceans (Dillon and Wright 1993) and monotremes (Gemmel et al. 1996) lack the CSB-3. The cow (Anderson et al. 1982), the sheep (Wood and Phua 1996), and the present cervids apparently possess a single fused CSB 2 + 3 (positions 941–966 in fig. 2), which is immediatly flanked by sequences homologous to the L-strand transcriptional start sites (fig. 1). The fusion of CSB 2 + 3 (Ghivizzani et al. 1993) could therefore be a diagnostic character of Ruminantia artiodactyls, but the low nucleotide conservation of this feature is inconsistent with its functional importance.

Pattern of Transition and Transversion Accumulation

The *Cervus elaphus* CR was independently determined in two laboratories (our present study and Feng, Li, Rittenhouse, and Templeton, unpublished data). The two sequences are close: overall divergence is 3.6% with two single-base-pair indels. There are 28 transitions (TI) and only four transversions (TV). These differences likely reflect intraspecies polymorphism occuring at hypervariable sites. For example, site 510 exhibits the four nucleotide character states (fig. 2), and the two consecutive sites 543 and 544 are polymorphic (A-to-G and G-to-A transitions) within *Capreolus capreolus* and *C. pygargus* (42 individuals considered, 0.3%–2.2% divergence within species; unpublished data) as well as within the *Cervus elaphus* species.


```
                                          <===  CSB-1 and CSB-1-like  ===>
Mazama sp.                  CATAATGA TAGGCATGGG CATGG**CAG TCAATGGTAA CGGGACATA
Mazama RS3          ATT------G ---AT--A-A ---TA**TG- -T-------G -A-------
Odocoileus hem.     ---------G --A------- -----**--- -G----T--G TAA------
Odocoileus hem. RS3 ACTG-----G --AA--C-A- ---AT**T-A -T-------- -A-------
Odocoileus vir.     ---------- ---------- ----A**--- -G----C--G TAA------
Odocoileus vir. RS3 ACT------G --A-T----A ---AA**--- ------C--G TA-------
Capreolus capreolus                         ---A**--- ---------G -A-------
Capreolus pygargus                          ----**--- ---------- -A-------
Hydropotes inermis                          ----**--- ---------- -A-------
Cervus elaphus                              ----**--- --------C- -A-------
Cervus elaphus'                             ----**--- --------C- -A-------
Cervus nippon                               --T-**--- --------C- -A-------
Dama dama                                   --CA**T-A -T------C- -A-------
Bos taurus                                  --TA**--- --------C- -A-------
Ovis aries                                  --AATAT-A -T------C- -A-------
```

FIG. 4.—Alignment of the conserved sequence block 1 (CSB-1) for seven non-Odocoileini cervids and two bovid species. For the three members of the Odocoileini tribe, *Mazama, Odocoileus hemionus* (hem.), and *O. virginianus* (vir.), the region spanning the last nucleotide after the O_H toward the end of the CSB-1 is aligned, together with the inserted motifs of the RS3 region. The similarities between these motifs suggest that the RS3 insertions of Odocoileini result from the duplication of the end of O_H toward the end of CSB-1- region. Conventions of alignment and definition of the CSB-1- and CSB-1-like sequences are the same as in figure 2. The sequence resembling the GYRCAT motif is underlined.

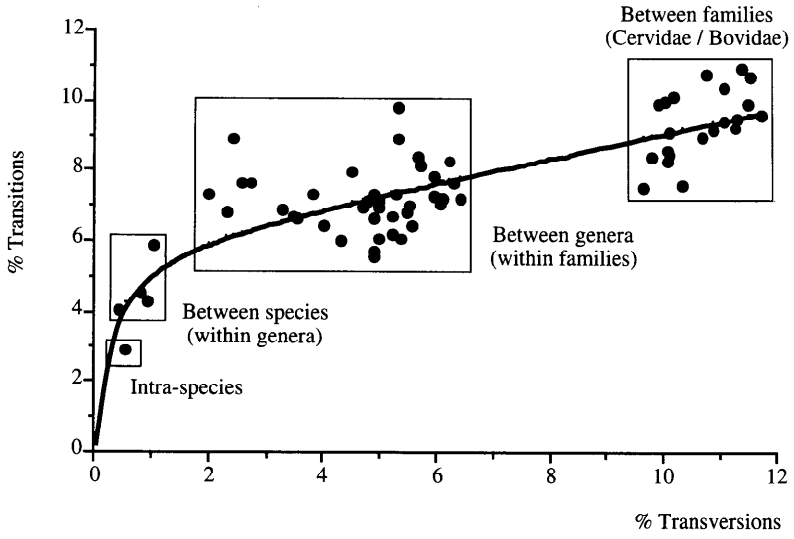


FIG. 5.—Evolution of the total percentage of transitions (TI) relative to the total percentage of transversions (TV) for 66 pairwise comparisons between the whole control region of 10 cervids and two bovids. The points corresponding to intraspecies, intragenera, intrafamily and interfamily comparisons are boxed. The nonlinear regression through these 66 points is a hyperbola type (this curve is descriptive but not predictive as data are not independent): TI = 5.93TV/(TV + 0.25) + 0.33TV (r^2 = 0.68). The curve can be approximated by the intersection of two linear regression plots. The first involves the five lowest transversional divergences, and the second involves the remaining 61 TV. This shows that the beginning of the saturation of TI relative to TV corresponds to the *Odocoileus hemionus/O. virginianus* divergence (5.9% TI and 1.1% TV).

The mean overall CR divergences between pairs of taxa of increasing taxonomic rank are as follows: 4.7%–6.7% within the genera *Capreolus, Odocoileus* and *Cervus*; 9.8% within the tribe Odocoileini; 11.9%–13.0% within the subfamilies Cervinae and [Odocoileinae + Hydropotinae]; 16.1% within the family Cervidae; and 27.4% between Bovidae and Cervidae. The percentage of TI plotted against percentage of TV for each pair of species is given in figure 5. For divergences greater than that of the pair *Odocoileus hemionus/O. virginianus* (5.9% TI and 1.1% TV), the slope of the curve decreases, likely reflecting the beginning of the saturation of TI relative to TV.

Phylogenetic Analyses

Because of the pattern of extensive indels in the mitochondrial CRs of cervids and the subsequent local ambiguities in the alignment of the 1,264 positions, we excluded from the phylogenetic reconstructions all sites where indels occur. An MP analysis conducted on the complete CR (846 positions with 211 informative TI and TV) produced a single most-parsimonious tree (539 steps long; consistency index [CI] = 0.60; retention index [RI] = 0.58). When the 102 informative TV were considered, a single most-parsimonious tree with identical topology was found (161 steps long; CI = 0.63; RI = 0.75). The observed topology remains stable whether NJ or ML methods are used, with all nodes supported by high bootstrap percentages (BPs) and decay indices (DIs) (fig. 6). The branching pattern depicts two clades. The first clusters the *Cervus* species and *Dama* within the Cervinae. The second was unexpected, as it places *Hydropotes,* the sole representative of the Hydropotinae subfamily, within the Odocoileinae. The latter group therefore joins the New World cervids (*Mazama* and the two *Odocoileus* species) together with a Chinese water deer/roe deer (*Hydropotes* + *Capreolus* species) clade.

To explore the phylogenetic content of the CR for comparisons between more distantly related taxa (Artiodactyla, Cetacea, Perissodactyla), the relatively conserved end of the 5′ domain, the CCR, and the beginning of the 3′ domain were aligned between cervids, bovids, suiforms (*Sus scrofa*), cetaceans (*Orcinus orca* and *Balaenoptera physalus*), and perissodactyls (*Equus caballus* and *Diceros bicornis*). The unambiguously alignable region spans 490 nt (corresponding to positions 315–788 in fig. 2), with 282 variable and 182 informative sites. The distance and MP approaches give a congruent phylogenetic picture (fig. 7). The Ruminantia (here represented by members of both bovid and cervid families) are monophyletic. Cetaceans (represented by one baleen whale and one toothed whale) are the sister group of ruminants. Moreover, the pig (Artiodactyla; suborder Suiformes) branches outside the [Ruminantia + Cetacea] clade. The lack of resolution within the Cervidae likely reflects the fact that the CCR evolves too slowly to bring phylogenetic signal at this intrafamily level.

Evolutionary Rates

Starting from the basal node (Cervidae) of the ML tree (fig. 6), we calculated a mean divergence of 9.6 ± 1.3% TI + TV leading to the terminal taxa. Assuming that the Cervidae family diverged 20 MYA (age of the oldest antlered deer; Ginsburg 1988), this gives a mean evolutionary rate equal to 0.48 ± 0.07% TI + TV/Myr/lineage. The removal of the saturating TI (see fig. 5) yielded a mean rate of 0.17 ± 0.04% TV only/Myr/lineage. These estimates represent means over the complete CR, and do not depict the heterogeneities of the substitution rates between the 5′ and 3′-peripheral do-

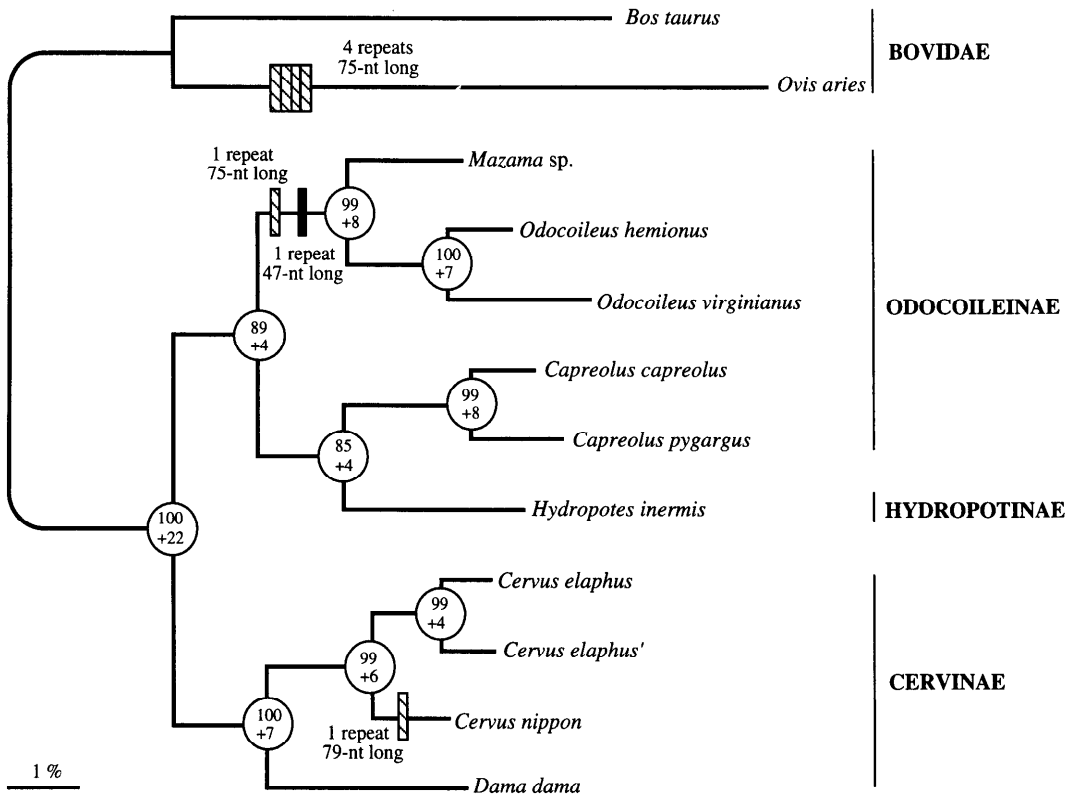


FIG. 6.—Phylogenetic tree describing the evolutionary relationships between 10 cervids and 2 outgroup bovids as deduced from the comparison of complete mitochondrial control region (CR) sequences excluding all sites where indels occur. The same topology is obtained whatever the kind of subsitutions used (transitions and transversions vs. transversions only) and whatever the method of tree reconstruction (distance vs. maximum parsimony vs. maximum likelihood). Here, the branch lengths are proportional to the expected number of changes as derived from the maximum-likelihood DNAML 3.5c program. At each node, the circles contain the bootstrap percentage (BP; after 1,000 resamplings) and the decay index (DI) computed on 102 informative transversions by the maximum-parsimony PAUP program. When the more saturating transitions are included with the transversions, the corresponding BPs and DIs appear to be equal or slightly lowered: 95/+5, 96/+8, 61/+1, 100/+17, 86/+6, 100/+29, 99/+10, 100/+13, and 77/+2 going from top to bottom nodes (from the *Mazama* and the two *Odocoileus* to the Cervinae nodes). The occurrence of the sequence repeats is also indicated in the most parsimonious fashion. The width of a rectangle is proportional to the repeat length. The streaked motif indicates that the repeats occur in the RS2 region, and the black motif indicates an occurence in the RS3 region. Topological constraints enforced to place *Hydropotes* as the sister group of all other cervids implies seven additional extra steps (transitions and transversions as well as transversions only). Douzery, Lebreton, and Catzeflis (1995) used the cervid model to test the generation time hypothesis, with a comparison of the scnDNA divergence between *Cervus elaphus* and the shorter-generation-time *Mazama* sp. The same species are included here, and no significant contrast in their CR evolutionary rates is observed.

mains and the CCR. The use of local molecular clocks, following the procedure of Bailey et al. (1991), provides ranges of estimated dates of divergence 0.4–2.5 Myr for the species *Cervus elaphus,* 1.4–4.3 Myr for the genus *Odocoileus,* 2.2–3.7 Myr for the genus *Capreolus,* 3.3–7.1 Myr for the genus *Cervus,* 6.8–10.2 Myr for the tribe Odocoileini, 8.7–10.4 Myr for the Capreolini/Hydropotinae cluster, 10.6–14.4 Myr for the subfamily Cervinae (as represented by *Cervus* and *Dama*), and 13.5–15.2 Myr for the Odocoileinae/Hydropotinae cluster (lower and upper values are deduced after the use of TV only or TI + TV, respectively). These results suggest contemporaneous origins for New World odocoileines and hydropotines + Old World odocoileines.

Discussion

Conserved Sequence Features of the Control Region of Cervidae

In the mitochondrial CRs of 10 species of Cervidae, we have identified sequences similar to all the important functional sites mapped, experimentally or comparatively, in other vertebrates (MacKay et al. 1986; King and Low 1987; Clayton 1991, 1992; Ghivizzani et al. 1993). The CCR of Cervidae, which has most of the CSBs described by Anderson et al. (1982), is very conserved among the studied Cervidae (fig. 2) and can be reliably aligned in a number of Artiodactyla, Cetacea, and Perissodactyla spanning more than 60 Myr of evolutionary divergence. On the contrary, the peripheral domains of the CR of Cervidae are variable in both nucleotide substitutions and length (fig. 2).

In the 5′-peripheral domain, there are short regions of highly variable sequences which alternate with short conserved sequences of motifs similar to the TASs of other vertebrates (Doda, Wright, and Clayton 1981; Stewart and Baker 1994; Gemmel et al. 1996). The length of the 5′-peripheral domain is variable in some Cervidae due to short indels (positions 121–130 and 148–159) and two tandem repeats (fig. 2). The region between TAS-4 and the beginning of the TV hot spot

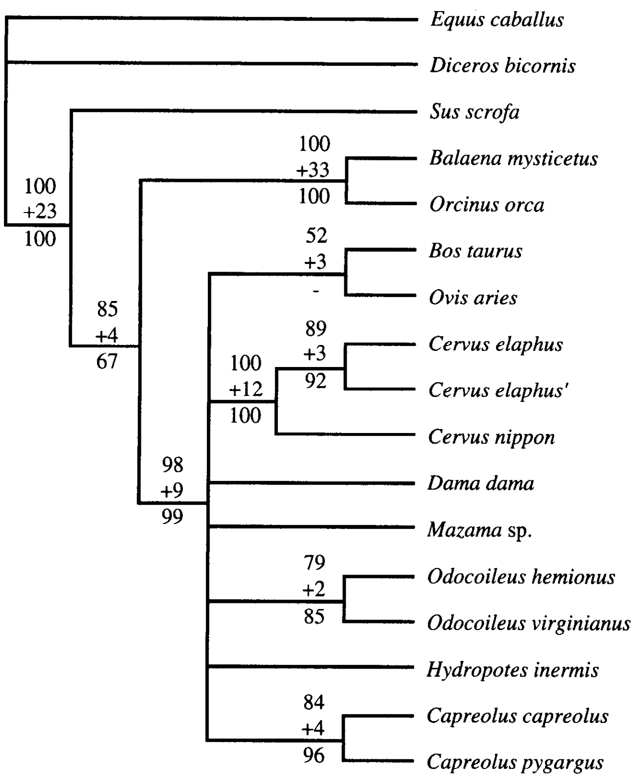


FIG. 7.—Majority-rule consensus tree after 1,000 bootstrap resamplings, describing the evolutionary relationships between 10 cervids, 2 bovids, 1 suiforme, 2 cetaceans, and 2 outgroup perissodactyls as deduced from the comparison of the conserved central region (and the relatively conserved end of the 5′-left domain and the beginning of the 3′-right domain) of their mitochondrial control regions (182 informative transitions, transversions, and indels). The same topology is obtained whatever the kind of method of tree reconstruction: distance (NJ) versus maximum parsimony (MP) versus maximum likelihood (ML). At each node, the bootstrap percentage (BP) and the decay index (DI) computed by the MP method are given above the branch, and the BP computed by the NJ method is given below (the hyphen indicates that the bovid grouping is not observed). One single most-parsimonious tree was found (length = 505; CI = 0.58), but only the nodes supported by more than 50% BP with the MP approach are represented. The monophyly of Cervidae is observed (DI = +1) but is not supported at the BP = 50% level.

(fig. 2) is of functional importance, with nucleotide sequences highly conserved in all the cervid species studied: TAS-4, TAS-3, and the GCCCC putative stop point of D-loop synthesis among Cervidae. In Odocoileini and sika deer, this 75-bp-long domain is tandemly duplicated in a single repeat which maps at RS2. This region contains VNTRs in a number of other mammalian species, for example vespertillionid bats (81-bp-long repeats: Wilkinson and Chapman 1991), shrews (78- to 80-bp-long repeats: Stewart and Baker 1994; Fumagalli et al. 1996), and sheep (75-bp-long repeats: Wood and Phua 1996). VNTRs at RS2 produce interspecific and intraspecific length differences and individual heteroplasmy (Hoelzel et al. 1994).

The RS2 repeat shows a short core sequence of six alternative purines and pyrimidines (GYRCAT), which is a putative partial copy of the TAS-1 (fig. 2), repeatedly inserted with flanking AT-rich strings in the region between TAS-1 and the 3′ D-loop end. Such tandemly repeated copies of functionally important sequences can form energetically stable single-stranded secondary structures (RS-XXII motif; figs. 2 and 3). The hairpins could constitute critical regions of DNA slippage during replication and could generate the tandem repetitions of RS-XXII. An identical duplicated motif near a TAS element has been found in a marsupial: a stretch of 25 nt is present in the opossum (*Didelphis virginiana*; Janke et al. 1994) and exhibits sequence similarities with RS-XXII (55%–59% identity with RS-XXII and the RS2 copy for marsupial/placental comparisons). The RS2 region has repeated copies of TAS and D-loop termination sequences, and this is likely a general character of VNTRs at these sites of the mammalian mtDNA (Anderson et al. 1982; MacKay et al. 1986; Wilkinson and Chapman 1991; Stewart and Baker 1994; Fumagalli et al. 1996; Wood and Phua 1996). These sequences apparently persist for long evolutionary periods: in lagomorphs, RS2 repeats are at least 6–8 Myr old (Biju-Duval et al. 1991), and in shrews, about 1 Myr old (Stewart and Baker 1994). The presence of multiple copies of functionally important sequences could enable the cell to use alternative pathways for regulation of mtDNA replication and transcription and can confer replicative advantages to mtDNA molecules with multiple copies of replication signals (Kuzminov 1996).

In the 3′-peripheral domain, the RS3 insertions of *Mazama* and *Odocoileus* appear to be characteristic of these two genera by comparison with the other cervid and bovid taxa (fig. 4). This insertion is here interpreted as a duplication of the region spanning the last nucleotide after the O_H toward the end of CSB-1. The sequence spanning CSB-1, O_H, and the other ≈80 nt toward the central CSB, can form a stable cloverleaf structure in single-stranded RNA chains. This secondary structure as well as the O_H and CSB-1 sites, should be important for controlling the initiation of H-strand DNA synthesis (Brown et al. 1986). This structure can be present in different vertebrate species (i.e., Stewart and Baker 1994; Wood and Phua 1996) and can be generated in Cervidae as well, although with different shapes and free energy values in the different species. Cloverleaf secondary structures of single-stranded polynucleotides can have important functional roles associated with the priming of transcription and replication of mtDNA (Clayton 1991). The Odocoileini pattern of a single 47-bp-long insertion in RS3 is unusual among mammals, because the other documented insertions in this region imply shorter fragments, ranging from 6 bp for carnivores (Hoelzel et al. 1994) to 20 bp for the rabbit (Mignotte et al., 1990), that are highly repeated (Mignotte et al. 1990; Ghivizzani et al. 1993; Hoelzel et al. 1994). The 47-bp-long insertion could represent a synapomorphy exclusively defining the tribe Odocoileini. To confirm this, comparison of the CRs of the other New World odocoileines (*Blastocerus, Hippocamelus, Ozotoceros* and *Pudu*) will be required.

Phylogenetic Content of the Control Region

The entire CR sequences, excluding the repeats and all the sites with indels, can be used to infer phyloge-

netic relationships among Cervidae. Transitions were shown to saturate sooner than transversions (fig. 5), as previously observed for the mitochondrial CRs of other placental mammals (e.g., Wills 1995), but the use of different classes of substitutions and different methods of tree reconstruction produced consistent results. The subfamilies Odocoileinae (including Hydropotinae) and Cervinae constitute two well-supported lineages. These clades are robust in terms of both bootstrap and decay index analyses (fig. 6: BP = 89% and 100% respectively; DI = +4 and +7 for a tree length of 161 steps).

The four representatives of the Cervinae (the two red deers, the sika deer, and the fallow deer) share a deletion of 11 bp at positions 150–160 (fig. 2) and cluster together. The Odocoileinae clade includes the New World genera *Mazama* and *Odocoileus,* and the Old World *Capreolus* unexpectedly linked with *Hydropotes* (Hydropotinae). Moreover the New World Odocoileinae have the apparently synapomorphic repeat at RS3, and they share with *C. nippon* the 75-bp-long repeat at RS2. While the RS3 insertion type is quite unusual for mammals, and therefore could represent a true synapomorphy for New World Odocoileinae (Fig. 6), VNTRs at RS2 are much more frequent. The phylogram in figure 6 implies that the presence of RS2 repeats both in sika deer and Odocoileinae is the result of a convergent evolutionary event.

The close relationship of *Hydropotes* with *Capreolus,* and therefore with the Odocoileinae, contrasts with current morphological classifications (Groves and Grubb 1987) and with other molecular data (Miyamoto, Kraus, and Ryder 1990; Kraus and Miyamoto 1991). From a morphological point of view, the main subdivision of Cervidae is based on a single character: absence (Hydropotinae) or presence (Odocoileinae + Cervinae) of antlers (Groves and Grubb 1987). Conversely, the hypothesis of an association of *Hydropotes* with the Odocoileinae is supported by the telemetacarpal condition and the large medial opening of the temporal canal (Bouvrain, Geraads, and Jehenne 1989). Such an association of Hydropotinae with Odocoileinae implies that antlers have been lost in Hydropotinae or convergently acquired in the other cervid subfamilies. Some morphological characters therefore seem to evolve with high homoplasy in ungulates (Scott and Janis 1993) and are difficult to use for determining reliable phylogenetic relationships.

Comparison of the mitochondrial 12S and 16S ribosomal (rRNA) genes supported either the basal position of *Hydropotes* relative to Cervinae, Muntiacinae, and Odocoileinae or the clustering of *Hydropotes* with *Odocoileus* (Miyamoto, Kraus, and Ryder 1990; Kraus and Miyamoto 1991). The reanalysis by the ML method of this data set (Kraus and Miyamoto 1991, figs. 3b, 4a, and 4b, excluding *Giraffa* and *Sus* because their 16S rRNA sequences are not available; TI/TV ratio = 3.0) strongly favors the grouping of *Hydropotes* with *Odocoileus.* The present data set of 12 pecoran CR sequences indicates that constraining the sister group relationship of *Hydropotes* to all other cervids involves seven additional steps (MP approach with TI and TV, as well as with TV only). This number of extra steps is equal to the DI supporting two well-established clades: the genus *Odocoileus* and the subfamily Cervinae (fig. 6). A basal position of *Hydropotes* relative to other cervids, and the subsequent monophyly of antlered deers, is therefore unlikely based on our CR data. Moreover, the CR sequence comparisons provide several exclusive synapomorphic substitutions or indels (CI = 1.0 in the most parsimonious tree) for a *Capreolus* + *Hydropotes* clade: C by G (position 124, fig. 2), C by G (position 147), insertion of one T (position 933), T by C (position 940), T by A (position 1010), A by T (position 1069), C by G (position 1085), and C by A (position 1087). The first two mutations appear in the fastest evolving CR region (the 5′-left domain: Wood and Phua 1996; unpublished data), whereas the six other occurs in a slower evolving area (the 3′-right domain), and none is diagnostic for the Capreolini + Hydropotinae clade in the CCR (the slowest evolving CR region; Saccone, Pesole, and Sbisá 1991). The above-mentioned changes in the 3′ domain are therefore less likely subject to homoplasy than those in the 5′ domain and could represent true synapomorphies. However, we cannot exclude the possibility that slow-evolving (and, reciprocally, fast-evolving) sites could be nested within fast-evolving (and reciprocally, slow-evolving) areas. Our complete CR sequences support the paraphyly of antlered deers due to the nesting of Hydropotinae within Odocoileinae, and are congruent with some morphological analyses (Bouvrain, Geraads, and Jehenne 1989), and with TV analyses of rRNA genes of mitochondria (Kraus and Miyamoto 1991), as well as with cytochrome *b* sequences (unpublished data).

The CCR expressed remarkable phylogenetic signal at deeper levels of genetic divergence and allowed us to reconstruct reliable phylogenetic relationships among representatives of Artiodactyla, Cetacea and Perissodactyla (fig. 7). Although the alignable portion of CCR is short (490 nt in this study), it seems to evolve in quite a regular fashion and could probably be used in phylogenetic studies together with other slow-evolving protein-coding mitochondrial genes. The close association of cetaceans with artiodactyls was suggested by comparison of nucleotide and amino acid sequences (Graur and Higgins 1994, but see Philippe and Douzery 1994; Gatesy 1997; Montgelard, Catzeflis, and Douzery 1997). Here, the CCR comparison suggests the sister-group relationship between ruminants (here represented by 12 pecoran taxa) and cetaceans (here represented by two taxa), with a more basal emergence of suiformes (here represented by only one taxon).

The use of TI + TV or TV only provides different estimates of the dates of divergence within the cervid family. The differences are introduced by the estimation of the ages of the first cladogenic event. Transversion estimates are always younger than TI + TV ones: respectively, 10.6 versus 14.4 Myr for the Cervinae and 13.5 versus 15.2 Myr for the Odocoileinae in a broad sense (Odocoileinae including Hydropotinae). After that, the intervals between speciation events within these Odocoileinae and Cervinae remain nearly identical whatever the kind of changes used. Thus, the initial discrepancies have repercussions on all subsequent datings of the more recent dichotomies.

The radiations of some odocoileine species appear to be Plio-Pleistocene events (1.4–4.3 Myr within the genus *Odocoileus* and 2.2–3.7 Myr within the genus *Capreolus,* as confirmed by the population genetics study of Randi, Pierpaoli, and Danilkin [unpublished data]). The divergences within Cervinae appear to be older events, having occured from the Miocene/Pliocene (3.3–7.1 Myr within the genus *Cervus*) to the Plio-Pleistocene (0.4–2.5 Myr within *Cervus elaphus*). One postulated drop in sea level occured 9 MYA and coincides with the dispersal of land mammals between Europe and North America (Opdyke 1990). This dating falls into the range of estimated ages of divergence for Odocoileini (6.8–10.2 Myr) and for Capreolini/Hydropotinae (8.7–10.4 Myr). An ancestral odocoileine stock showed a radiation in Eurasia 8.7–10.4 MYA, producing the palearctic Capreolini, today represented by the European and Siberian *Capreolus,* and the Asian Hydropotinae, represented by the unique surviving Chinese *Hydropotes inermis.* Simultaneously, immigrants could have invaded North America through the Bering Land Bridge 9 MYA starting from the same odocoileine Eurasian ancestral stock. Their diversification led to New World Odocoileinae, here represented by the genera *Mazama* and *Odocoileus* and characterized by the RS3 insertion in their mitochondrial CR.

Acknowledgments

We thank Aleksey Danilkin (Russian Academy of Science, Moscow), and Oliver Ryder (San Diego Zoo) for kindly providing us with tissues of Siberian roe deer, brocket deer (individual "Leona") and Chinese water deer. E.D. gratefully acknowledges François Catzeflis for the laboratory environment and the access to the collection of tissues. Jean-Yves Dubuisson is also thanked for his constant help in the management of computer programs. Massimo Pierpaoli and Nadia Mucci helped with laboratory analyses. The comments of Dr. Richard G. Harrison and two anonymous referees were greatly appreciated. Laboratory facilities were partially funded by the "Service Commun de Biosystématique de Montpellier." This work was supported by the "Groupement de Recherche et d'Etude des Génomes" (G.R.E.G.; décision no. 82/94 of July 19, 1994). This is contribution no. 97–065 of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554—CNRS).

LITERATURE CITED

ANDERSON, S., M. H. L. DE BRUIJN, A. R. COULSON, I. C. EPERON, F. SANGER, and I. G. YOUNG. 1982. Complete sequence of bovine mitochondrial DNA. Conserved feature of the mammalian mitochondrial genome. J. Mol. Evol. **156**:683–717.

ARCTANDER, P., P. W. KAT, R. A. AMAN, and H. R. SIEGISMUND. 1996*a*. Extreme genetic differences among populations of *Gazella granti,* Grant's gazelle, in Kenya. Heredity **76**:465–475.

ARCTANDER, P., P. W. KAT, B. T. SIMONSEN, and H. R. SIEGISMUND. 1996*b*. Population genetics of Kenyan impalas—consequences for conservation. Pp. 399–412 *in* R. K. WAYNE and T. B. SMITH, eds. Molecular genetics in conservation. Oxford University Press, Oxford.

ARNASON, U., A. GULLBERG, and B. WIDEGREN. 1991. The complete nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus.* J. Mol. Evol. **33**:556–568.

———. 1993. Cetacean mitochondrial DNA control region: sequences of all extant baleen whales and two sperm whale species. Mol. Biol. Evol. **10**:960–970.

ARNASON, U., and E. JOHNSSON. 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina.* J. Mol. Evol. **34**:493–505.

BAILEY, W. J., D. H. A. FITCH, D. A. TAGLE, J. CZELUSNIAK, J. L. SLIGHTOM, and M. GOODMAN. 1991. Molecular evolution of the ψη-globin gene locus: gibbon phylogeny and the hominoid slowdown. Mol. Biol. Evol. **8**:155–184.

BIJU-DUVAL, C., H. ENNAFAA, N. DENNEBOUY, M. MONNEROT, F. MIGNOTTE, R. C. SORIGUER, A. EL GAAÏED, A. EL HILI, and J.-C. MOUNOLOU. 1991. Mitochondrial DNA evolution in lagomorphs: origin of systematic heteroplasmy and organization of diversity in European rabbits. J. Mol. Evol. **33**:92–102.

BOUVRAIN, G., D. GERAADS, and Y. JEHENNE. 1989. New data relating to the classification of the Cervidae (Artiodactyla, Mammalia). Zool. Anz. **223**:82–90.

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. Evolution **42**:795–803.

BROWN, G. G., G. GADALETA, G. PEPE, C. SACCONE, and E. SBISÁ. 1986. Structural conservation and variation in the D-loop containing region of vertebrate mitochondrial DNA. J. Mol. Biol. **192**:503–511.

CATZEFLIS, F. M. 1991. Animal tissue collections for molecular genetics and systematics. Trends Ecol. Evol. **6**:168.

CLAYTON, D. A. 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. **7**:453–478.

———. 1992. Transcription and replication of animal mitochondrial DNAs. Pp. 217–232 *in* D. R. WOLSTENHOLME and K. W. JEON, eds. Mitochondrial genomes. International Review of Cytology, Vol. 141. Academic Press, San Diego, Calif.

DAIRAGHI, D. J., and D. A. CLAYTON. 1993. Bovine RNase MRP cleaves the divergent bovine mitochondrial RNA sequence at the displacement-loop region. J. Mol. Evol. **37**:338–346.

DILLON, M. C., and J. M. WRIGHT. 1993. Nucleotide sequence of the D-loop region of the sperm whale *(Physeter macrocephalus)* mitochondrial genome. Mol. Biol. Evol. **10**:296–305.

DODA, J. N., C. T. WRIGHT, and D. A. CLAYTON. 1981. Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. USA **78**:6116–6120.

DOUZERY, E., J.-D. LEBRETON, and F. M. CATZEFLIS. 1995. Testing the generation time hypothesis using DNA/DNA hybridization between artiodactyls. J. Evol. Biol. **8**:511–529.

EISENBERG, J. F. 1981. The mammalian radiations. An analysis of trends in evolution, adaptation and behavior. The University of Chicago Press, Chicago and London.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

———. 1993. PHYLIP (phylogeny inference package). Version 3.5c. Department of Genetics, University of Washington, Seattle.

FUMAGALLI, L., P. TABERLET, L. FAVRE, and J. HAUSSER. 1996. Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. Mol. Biol. Evol. **13**:31–46.

GATESY, J. 1997. More DNA support for a Cetacea/Hippopotamidae clade: the blood-clotting protein gene γ-fibrinogen. Mol. Biol. Evol. **14**:537–543.

GEMMELL, N. J., P. S. WESTERN, J. M. WATSON, and J. A. MARSHALL-GRAVES. 1996. Evolution of the mammalian mitochondrial control region—comparisons of control region sequences between monotreme and therian mammals. Mol. Biol. Evol. **13**:798–808.

GHIVIZZANI, S. C., S. L. D. MACKAY, C. S. MADSEN, P. J. LAIPIS, and W. W. HAUSWIRTH. 1993. Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. J. Mol. Evol. **37**:36–47.

GINSBURG, L. 1988. La faune des mammifères des sables miocènes du synclinal d'Esvres (Val-de-Loire). C. R. Acad. Sci. Paris II **307**:319–322.

GRAUR, D., and D. G. HIGGINS. 1994. Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. Mol. Biol. Evol. **11**:357–364.

GROVES, C. P., and P. GRUBB. 1987. Relationships of living deer. Pp. 21–59 *in* C. M. WEMMER, ed. Biology and management of the Cervidae. Smithsonian Institution Press, Washington, D.C.

HOELZEL, A. R., J. M. HANCOCK, and G. A. DOVER. 1991. Evolution of the cetacean mitochondrial D-loop region. Mol. Biol. Evol. **8**:475–493.

HOELZEL, A. R., J. V. LOPEZ, G. A. DOVER, and S. J. O'BRIEN. 1994. Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of Carnivores. J. Mol. Evol. **39**:191–199.

IRWIN, D. M., T. D. KOCHER, and A. C. WILSON. 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. **32**:128–144.

JAGER, F., W. HECHT, and A. HERZOG. 1992. Untersuchungen an mitochondrialer DNS (mtDNS) von hessischem Rehwild (*C. capreolus*). Z. Jagdwiss. **38**:26–33.

JAMA, M., Y. ZHANG, R. A. AMAN, and O. A. RYDER. 1993. Sequence of the mitochondrial control region, tRNAThr, tRNAPro and tRNAPhe genes from the black rhinoceros *Diceros bicornis*. Nucleic Acids Res. **21**:4392.

JANKE, A., G. FELDMAIER-FUCHS, W. K. THOMAS, A. VON HAESELER, and S. PÄÄBO. 1994. The marsupial mitochondrial genome and the evolution of placental mammals. Genetics **137**:243–256.

KING, T. C., and R. C. LOW. 1987. Mapping of control elements in the displacement loop region of mitochondrial DNA. J. Biol. Chem. **262**:6204–6213.

KRAUS, F., and M. M. MIYAMOTO. 1991. Rapid cladogenesis among the pecoran ruminants: evidence from mitochondrial DNA sequences. Syst. Zool. **40**:117–130.

KUZMINOV, A. 1996. Mutant fixation via plasmid dimerization and its relation to human diseases. Trends Genet. **12**:246–249.

LEVINSON, G., and G. A. GUTMAN. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. **4**:203–221.

LOFTUS, R. T., D. E. MACHUGH, D. G. BRADLEY, P. M. SHARP, and P. CUNNINGHAM. 1994. Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci. USA **91**: 2757–2761.

MACKAY, S. L. D., P. D. OLIVO, P. J. LAIPIS, and W. W. HAUSWIRTH. 1986. Template-directed arrest of mammalian mitochondrial DNA synthesis. Mol. Cell. Biol. **6**:1261–1267.

MADSEN, C. S., S. C. GHIVIZZANI, and W. M. HAUSWIRTH. 1993. Protein binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Mol. Cell. Biol. **13**:2162–2171.

MIGNOTTE, F., F. GUERIDE, A. M. CHAMPAGNE, and J. C. MOUNOLOU. 1990. Direct repeats in the noncoding region of rabbit mitochondrial DNA: involvment in the generation of intra- and inter-individual heterogeneity. Eur. J. Biochem. **194**:561–571.

MIYAMOTO, M. M., F. KRAUS, and O. A. RYDER. 1990. Phylogeny and evolution of antlered deer determined from mitochondrial DNA sequences. Proc. Natl. Acad. Sci. USA **87**:6127–6131.

MONTGELARD, C., F. M. CATZEFLIS, and E. DOUZERY. 1997. Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome *b* and 12S rRNA mitochondrial sequences. Mol. Biol. Evol. **14**:550–559.

MULLIS, K., F. FALOONA, S. SCHARF, R. SAIKI, G. HORN, and H. ERLICH. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. **51**:263–273.

OPDYKE, N. D. 1990. Magnetic stratigraphy of cenozoic terrestrial sediments and mammalian dispersal. J. Geol. **98**: 621–637.

PHILIPPE, H. 1993. MUST: a computer package of management utilities for sequences and trees. Nucleic Acids Res. **21**: 5264–5272.

PHILIPPE, H., and E. DOUZERY. 1994. The pitfalls of molecular phylogeny based on four species as illustrated by the Cetacea/Artiodactyla relationships. J. Mamm. Evol. **2**:133–152.

RAND, D. M., and R. G. HARRISON. 1989. Molecular population genetics of mtDNA size variation in crickets. Genetics **121**:551–569.

SACCONE, C., G. PESOLE, and E. SBISÁ. 1991. The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. J. Mol. Evol. **33**:83–91.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SAMBROOK, J., E. F. FRITSCH, and T. MANIATIS. 1989. Molecular cloning. Cold Spring Harbor Laboratory Press, New York.

SANGER, F., S. NICKLEN, and A. R. COULSON. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

SCOTT, K. M., and C. M. JANIS. 1987. Phylogenetic relationships of the Cervidae, and the case for a superfamily "Cervoidea." Pp. 3–20 *in* C. M. WEMMER, ed. Biology and management of the Cervidae. Smithsonian Institution Press, Washington, D.C.

———. 1993. Relationships of the Ruminantia (Artiodactyla) and an analysis of the characters used in ruminant taxonomy. Pp. 282–302 *in* F. S. SZALAY, M. J. NOVACEK, and M. C. MCKENNA, eds. Mammal phylogeny: placentals. Springer Verlag, New York.

STEWART, D. T., and A. J. BAKER. 1994. Patterns of sequence variation in the mitochondrial D-loop region of shrews. Mol. Biol. Evol. **11**:9–21.

SWOFFORD, D. L. 1993. PAUP 3.1.1: Phylogenetic analysis using parsimony. Illinois Natural History Survey, Champaign.

TAKEDA, K., A. ONISHI, N. ISHIDA, K. KAWAKAMI, M. KOMATSU, and S. INUMARU. 1995. SSCP analysis of pig mitochondrial DNA D-loop region polymorphism. Anim. Genet. **26**:321–326.

THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

WILKINSON, G. S., and A. M. CHAPMAN. 1991. Length and sequence variation in evening bat D-loop mtDNA. Genetics **128**:607–617.

WILLS, C. 1995. When did Eve live? An evolutionary detective story. Evolution **49**:593–607.

WOOD, N. J., and S. H. PHUA. 1996. Variation in the control region sequence of the sheep mitochondrial genome. Anim Genet. **27**:25–33.

XU, X., and U. ARNASON. 1994. The complete mitochondria DNA sequence of the horse, *Equus caballus*: extensive het eroplasmy of the control region. Gene **148**:357–362.

ZUKER, M., and P. STIEGLER. 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxil iary information. Nucleic Acids Res. **9**:133–148.

RICHARD G. HARRISON, reviewing editor

Accepted August 12, 1997