Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 28 (2003) 485–499
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Ancient DNA analysis reveals woolly rhino evolutionary relationships☆

Ludovic Orlando,[a] Jennifer A. Leonard,[b] Aurélie Thenot,[a] Vincent Laudet,[c] Claude Guerin,[d] and Catherine Hänni[a,*]

[a] *CNRS UMR 5534, Centre de Génétique Moléculaire et Cellulaire, Université Claude Bernard Lyon 1, 16 rue R. Dubois, Bâtiment G. Mendel, 69622 Villeurbanne Cedex, France*
[b] *Department of Organismic Biology, Ecology and Evolution, University of California, Los Angeles, CA 90095-1606, USA*
[c] *CNRS UMR 5665, Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France*
[d] *CNRS UMR 'Paléoenvironnements et Paléobiosphère', Université Claude Bernard Lyon1, 69622 Villeurbanne Cedex, France*

Received 14 June 2002; revised 19 November 2002

Abstract

With ancient DNA technology, DNA sequences have been added to the list of characters available to infer the phyletic position of extinct species in evolutionary trees. We have sequenced the entire *12S rRNA* and partial cytochrome *b* (*cyt b*) genes of one 60–70,000-year-old sample, and partial *12S rRNA* and *cyt b* sequences of two 40–45,000-year-old samples of the extinct woolly rhinoceros (*Coelodonta antiquitatis*). Based on these two mitochondrial markers, phylogenetic analyses show that *C. antiquitatis* is most closely related to one of the three extant Asian rhinoceros species, *Dicerorhinus sumatrensis*. Calculations based on a molecular clock suggest that the lineage leading to *C. antiquitatis* and *D. sumatrensis* diverged in the Oligocene, 21–26 MYA. Both results agree with morphological models deduced from palaeontological data. Nuclear inserts of mitochondrial DNA were identified in the ancient specimens. These data should encourage the use of nuclear DNA in future ancient DNA studies. It also further establishes that the degraded nature of ancient DNA does not completely protect ancient DNA studies based on mitochondrial data from the problems associated with nuclear inserts.

1. Introduction

Phylogenetic relationships among extant species can be assumed thanks to morphological and genetical data. However, the literature is full of cases of incongruence between phylogenetic trees and divergence dates of extant species estimated using either the former or the latter approach (Cooper and Fortey, 1998). This conflict has recently expanded to affect palaeontological remains, as it has become possible to retrieve DNA from the sub-fossils remains of extinct species. In some cases, such as for cave bears, ancient DNA analysis has supported one of a variety of hypotheses based on morphology (Hänni et al., 1994; Loreille et al., 2001). But, in other cases, such as the phylogenetic link between extinct moas and extant ratites, the analysis of ancient DNA sequences led to completely unexpected conclusions (Sorenson et al., 1999).

Beside the horses, the rhinoceroses (Rhinocerotidae, Perissodactyla) were one of the most flourishing mammals family in the past. At least 17 different Miocene species have been recognised in Europe (Guérin, 1980a,b), including the woolly rhinoceros. This is one of the most fascinating prehistoric species, and is found depicted in cave paintings and engravings discovered in Font-de-Gaume, Rouffignac, and especially Chauvet caves (France) (Guérin, 1989; Prothero et al., 1989). The woolly rhinoceros was highly adapted to grazing the temperate and tundra grasslands. Based on the fossil record, the woolly rhino evolved in Asia, where it

☆ Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AY178623–AY178633.

* Corresponding author. Fax: +33-472440555.
E-mail address: hanni@univ-lyon1.fr (C. Hänni).

1055-7903/$ - see front matter © 2003 Published by Elsevier Inc.
doi:10.1016/S1055-7903(03)00023-X

appeared in the early Pleistocene, and reached Europe at the beginning of the Riss glaciation event (250 KYA[1]). Unlike the woolly mammoth, it never reached America. Its last representatives disappeared at the end of the Pleistocene, about 10 KYA (Thew et al., 2000).

Five species of rhinoceroses are living today. The black (*Diceros bicornis*) and the white (*Ceratotherium simum*) rhinoceroses are restricted to Africa (Dorst and Dandelot, 1972; Halthenorth and Diller, 1985). In Asia, the two-horned Sumatran rhinoceros (*Dicerorhinus sumatrensis*) and the one-horned Javan rhinoceros (*Rhinoceros sondaicus*) are found in Indonesia. The great unicorn rhinoceros (*Rhinoceros unicornis*) is found only in the marshy grasslands of India (Bengal and Nepal). Because the woolly rhinoceros was endemic to Eurasia, it has been hypothesised that it is more closely related to extant Asian, as opposed to extant African rhinoceroses (Kurten, 1968). However, it is not known to which of the three extant Asian rhinoceros it is the most closely related. With this study, we attempt to answer this long-standing question with a molecular phylogeny based on ancient *12S rRNA* and cytochrome *b* (*cyt b*) sequences.

2. Materials and methods

2.1. Samples

Seven samples of *Coelodonta antiquitatis* tooth roots were obtained from three different layers in the Scladina cave (Sclayn, Belgium, Otte et al., 1998): SC30100, SCI26/A129, SC82210, and SC81205 from layer C1A (40–45,000 years BP), SC6600 and SC7400 from layer C2A (60–70,000 years BP), and SC16400 from layer C5 (90–130,000 years BP). They were selected because: (i) other samples from Scladina cave have previously yielded authentic ancient DNA sequences (Loreille et al., 2001; Orlando et al., 2002) and (ii) because of the good state of protein preservation, supported by δ^{15}N isotopic ratios conducted on collagen amino-acids ($0.4 < \%\delta^{15}$N < 2.6, H. Bocherens, pers. com.).

2.2. Extraction

Extraction and amplification from tooth root extracts were performed in separate rooms with specific equipment, dedicated to ancient DNA work (Hänni et al., 1994; Vila et al., 2001). Each sample was extracted independently, except SCI26/A129 and SC81205. A 200-year-old ancient human skull (1.3 g) was coextracted with SC16400 as a control. Tooth roots (0.5–1.3 g) were reduced to powder with a hammer in a sterile enclosed plastic bag. Decalcification of the powder was conducted at the same time as protein digestion by an overnight incubation in 10 ml of 0.5 M EDTA (pH 8.5), 0.5% *N*-lauryl-sarcosyl, and 1 mg/ml proteinase K at 55 °C. Three steps of centrifugation of the supernatant in a mixture of phenol/chloroform/isoamyl alcohol (25:24:1, 1000–1200 rpm, 8–15 min) separated DNA from a generally highly rich protein fraction. The supernatant was then concentrated by Centricon 30 dialysis according to the manufacturer (Amicon) and finally, DNA was retrieved in 80–10 µl water from each column.

2.3. PCR amplification

Using our previously published data (Tougard et al., 2001), we designed several rhino specific primers (8 for the *12S rRNA* gene and 10 for the *cyt b* gene) suitable for ancient DNA analysis. These primers were designed to amplify short overlapping fragments (150–222 bp) because of the degraded nature of ancient DNA (Table 1). Depending on the sample, 1–5 µl of extract was used in a 100 µl PCR reaction. Ten units of Perkin Elmer Gold *Taq* polymerase was used per reaction. BSA was added to the PCR (Roche, 1 mg/ml), and $MgCl_2$ concentration varied between 2 and 3 mM. DNA was amplified with 50 cycles of denaturation (92 °C, 60 s), annealing (45–50 °C, 60 s) and elongation (72 °C, 45 s) in an Eppendorf PCR Mastergradient apparatus. To minimise the amount of ancient extract used, one fragment of *cyt b* and one fragment of *12S rRNA* were sometimes amplified in a multiplex PCR using two pairs of primers (782L–1006H and 14614L–14809H, Fig. 1; 14156L–14331H and 14844L–15047H). Three independent blanks were carried out for each set of PCR experiment as described in Loreille et al. (2001). An extraction blank insured no exogenous contamination occurred. A PCR blank that remained open during the preparation time of PCR proved no aerosol contaminated our tubes in the PCR room. The blank that remained open when ancient extracts were added into the PCR tubes tested the same absence of environmental contamination in the room (Fig. 1). Subcloning was pursued only on PCR sets displaying all blanks negative.

2.4. Cloning and sequencing

To construct a consensus sequence devoid of *Taq* polymerase errors and/or post-mortem DNA decay, all the PCR amplification products were subcloned using Topo TA cloning kit according to the manufacturer instructions (Invitrogen, The Netherlands). Two to eight clones per amplification product were sequenced in both strand directions on a Megabace[1000] capillary automatic sequencer (Amersham) following plasmid preparation (QIAprep spin miniprep kit, Qiagen).

[1] *Abreviations used:* mtDNA, mitochondrial DNA; *cyt b*, cytochrome *b*; numt, nuclear insert of mtDNA; BP, Before Present; KYA, thousands years ago; MYA, millions years ago.

Table 1
Bootstrap supports for nodes A and B

Gene	Sites	Phylogenetic method			
		K2	TN	MP	ML
cyt b	$668^{(\mathrm{all},\ 281)}$	94/92	95/92	83/84	88/85
	$445^{(\mathrm{1st+3rd},\ 258)}$	86/88	87/83	73/77	84/71
	$223^{(\mathrm{3rd},\ 190)}$	70/50	69/–	57/–	51/–
12S rRNA	$916^{(\mathrm{all},\ 291)}$	96/100	95/100	96/100	95/100
	$886^{(\mathrm{HVR\ excl.},\ 263)}$	96/100	95/100	93/99	88/98
cyt b + 12S rRNA	$1584^{(\mathrm{all},\ 572)}$	100/100	100/100	99/100	100/100
	$1554^{(\mathrm{HVR\ excl.},\ 544)}$	100/100	100/100	99/100	100/99
	$1361^{(\mathrm{all},\ 549)}$	100/100	99/100	99/100	100/100
	$1331^{(\mathrm{HVR\ excl.},\ 521)}$	100/100	99/100	97/100	96/100
	$1137^{(\mathrm{all},\ 3rd)}$	100/100	99/100	96/99	100/99
	$1107^{(\mathrm{HVR\ excl.},\ 3rd)}$	99/100	100/100	95/99	93/99

The sites used in the phylogenetic analyses and the number of parsimony informative sites are, respectively, reported in parentheses. Above slash: bootstrap support for node A in Fig. 4 (*Coelodonta antiquitatis* and *Dicerorhinus sumatrensis* monophyly). Below slash: bootstrap support for node B in Fig. 4 (Rhinoceroses monophyly). K2, Kimura two parameters corrected distance; TN, Tajima and Nei distance; MP, Maximum of Parsimony; ML, Maximum of Likelihood; all, complete 12S rRNA; HVR excl., hypervariable region of the 12S rRNA excluded; 1st, first codon position; 3rd, third codon position.

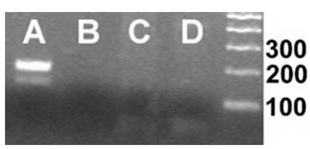

Fig. 1. Electrophoresis gel of multiplex PCR products (782L–1006H primers for the *12S rRNA* gene and 14614L–14809H primers for the *cyt b* gene). (A) SC7400 extract. (B) Mock extract. (C) PCR blank. (D) Room blank. The size of the ladder fragments are reported in bp.

2.5. Sequences analysis

12S rRNA and *cyt b* sequences from other species were retrieved from GenBank: *Artibeus jamaicensis* AF061340, AF061340; *Pteropus scapulatus* AF321050, AF321050; *Chalinolobus tuberculatus* AF321051, AF321051; *Tayassu tajacu* X86944, X56296; *Bos taurus* V00654, V00654; *Phoca vitulina* X63726, X63726; *Ursus arctos* Y08519, X82308; *Herpestes auropunctatus* Y08506, X94926; *Panthera tigris* Y08504, X82301; *Equus caballus* X79547, D32190; *Equus grevyi* X86943, X56282; *Tapirus* sp. AF038012, AF056030; *Diceros bicornis* AJ245721, X56283; *Ceratotherium simum* Y07726, Y07726; *Rhinoceros unicornis* X97336, X97336; *Rhinoceros sondaicus* AJ245724, AJ245725; *Dicerorhinus sumatrensis* AJ245722, AJ245723. Note that *Tapirus* sp. refers to *Tapirus pinchaque 12S rRNA* and *Tapirus terrestris cyt b* sequence. Sequences were aligned manually using the Seaview software (Galtier et al., 1996). For the *12S rRNA*, secondary structure constraints in stem and loop regions have been respected according to Douzery and Catzeflis (1995). All the phylogenies were constructed with the Phylo_win program (Galtier et al., 1996). To insure results were not dependent on the method employed, distances were calculated with Kimura's two parameters model or Tajima and Nei's correction. Parsimony and likelihood algorithms were also used. The effects of different substitution rates in the *12S rRNA* stems or loops were also tested either by taking into account global substitutions along the gene (although excluding one hypervariable region between positions 932 and 975 in Fig. 2), or by considering only transversions in loops and both transitions and transversions in stems. For the *cyt b* data, phylogenies based on third or first and third codon position substitutions were constructed (Irwin et al., 1991). In each case, the robustness of the branching was evaluated by doing 1000 bootstrap replicates, except for likelihood analyses where only 100 were performed (Table 1).

3. Results

3.1. Authentication criteria

Among the seven fossils analysed, three yielded ancient DNA (SC7400, SC81205, and SC30100, Figs. 2 and 3). SC16400 was extracted at the same time as the ancient human skull. Neither of the two attempts to amplify human DNA from the SC16400 sample using HyperVariable Region I primers (16037L 5′ AAGCAG ATTTGGGTACCAC 3′–16160H 5′ TGTACTACAGG TGGTCAAGT 3′) succeeded in the rhino extract as opposed to the human extract; conversely, no PCR product was obtained when using rhino primers on the human extract, suggesting no cross-contamination between samples occurred during the extraction (data not shown). The same holds true for SC81205 that yielded ancient products whereas no DNA was retrieved with

SCI26/A129, the sample that was extracted in the same session.

The complete *12S rRNA* gene (8 fragments, 975 bp) and 7 *cyt b* fragments (688 bp) were amplified from sample SC7400 (Figs. 2 and 3). Each fragment from sample SC7400 was amplified two to five times and cloned. Overlapping regions were always identical, except between the consensus sequences of the fragments 23L–210H and 128L–325H of the *12S rRNA* gene. However, three clones from three independent PCR amplifications of the 128L–325H fragment led to the same sequence than the 24 clones from 4 independent PCR amplifications of the fragment 23L–210H (noted # in Fig. 2). Thus, in the overlapping region, the sequence from the latter consensus was taken as authentic.

SC30100 and SC81205 gave partial sequences of *12S rRNA* gene (Figs. 2 and 3). From the sample SC81205, we obtained the fragments 128L–325H, 275L–488H, and 782L–1004H of the *12S rRNA* gene. Their sequence was deduced from the consensus of multiple clones of two independent PCR amplifications (Figs. 2 and 3). In the *12S rRNA* gene, the sample SC81205 showed 0–2.0% substitutions from the sequence of the sample SC7400. The sample SC30100 yielded the fragment 14742L–14915H of the *cyt b* gene, and 128L–325H and 275L–488H of the *12S rRNA* gene. The final sequences of sample SC30100 were deduced from the clones of one amplification product; the *cyt b* sequence perfectly matches the sequence determined for the sample SC7400 whereas the *12S rRNA* sequence exhibits one substitution (Figs. 2 and 3).

Authenticity of the sequences we obtained is supported by the following observations: (i) partial sequences for samples SC7400 and SC81205 were independently replicated with identical results in the UCLA laboratory (fragment 275L–488H), (ii) base composition is similar to that measured in extant rhinos (Xu and Arnason, 1997, 38.1% A, 23.8% C, 16.4% G, 22.8% T for the *12S rRNA* L-strand, and 38.1% A, 48.4% C, 2.7% G, 10.6% T for *cyt b* 3rd codon position), (iii) substitution profiles are biased toward transitions (among the Rhinocerotidae, we found mean TS/TV *ratio* for *12S rRNA* = 6.3 and *cyt b* 3rd codon position = 5.3, both results, respectively, in agreement with 6.0 and 3.3 for ratios between the white and the Indian rhino published in Xu and Arnason, 1997), (iv) substitutions between ancient and all the sequences from extant rhinos on the 2nd position of codon are rare (less than 2% of global substitutions, Irwin et al., 1991), and (v) there were no stop codons observed in the *cyt b* coding sequence.

3.2. Amplification of divergent endogenous DNA

Interestingly, sequences of some clones obtained from several fragments were recognised as non-bona fide mtDNA sequences. Some of them could be interpreted as numts for the following reasons (Figs. 2 and 3).

For SC7400 sample (60–70 KYA), two clones (1.5 and 1.f, Fig. 3) from one amplification of the fragment 14844L–15047H of *cyt b* appeared very divergent from the sequences of extant rhinoceroses, accumulating 25 substitutions when compared to the *C. simum* reference and only a weak bias toward transition (TS/TV = 2.6). A 4 bp deletion was necessary to align it to the sequences from extant species (positions 743–746 in Fig. 3). In fragment 15191L–15361H of *cyt b*, the sequences of 4 clones (1.3, 1.6, 1.7, and 1.9) from one PCR amplification diverged significantly from the consensus deduced from 4 independent amplifications (13 transitions and 5 transitions, Fig. 3).

For the sample SC81205, 9 clones (1.2, 1.4, 1.9, 1.b, 1.d, 1.e, 1.f, 1.j, and 1.n) from one amplification product of the *12S rRNA* fragment 128L–325H accumulated 18 transitions, 5 transversions, and 8 indels when compared to the consensus sequence of sample SC7400 (Fig. 2).

Using the BLAST program, a *C. simum* and a *D. bicornis* pseudogene (Accession No. AF142096 and AJ404858) appear as more similar to the consensus sequences from the 4 or the 9 clones (respectively, identity percentages = 90 and 97%, *e*-values = 9×10^{-34} and 2×10^{-71}, blast scores = 149 and 255) than rhinoceros mitochondrial genes (respectively, *e*-value $< 6 \times 10^{-32}$ and 5×10^{-26}, blast scores < 143 and 104). Furthermore, less substitutions are observed with these pseudogenes than with rhinoceroses mtDNA genes (respectively 4.8% versus 7.6–13.2%, and 2.2% versus 13–15.9%).

Extensively damaged PCR fragments were also recognised as non-bona fide mtDNA sequences. In the *12S rRNA* gene, two different rhino-like sequences were obtained for SC7400 fragment 275L–488H (noted SC7400 275L–488H* DAM. in Fig. 2). The mtDNA fragment was recognised as it was: (i) amplified more often, (ii) more closely related to modern rhinos, and (iii) amplified in both laboratories (Lyon and UCLA); the other less frequently amplified fragment accumulated 4 transitions (2 T → C and 2 C → T), 1 transversion (A → C), and 1 indel.

For SC81205 sample (40–45 KYA), two amplifications of the *12S rRNA* 782L–1004H fragment yielded a sequence very similar to the one from sample SC7400 (observed divergence = 0.5%). A third amplification product displayed 11.7% substitutions when compared to the sequence from sample SC7400 (clones 2.5 and 2.9 in Fig. 2): 19 transitions (9 C → T, 3 T → C, 4 G → A, and 3 A → G), 2 transversions, and 3 indels (Fig. 2). Among these, 8 transitions (2 A → G, 2 G → A, 3 C → T, and 1 A → T) and 1 indel are never observed among extant rhino mitochondrial sequences. Moreover, in the 49 bp overlapping 5′ of the fragment, the amplification product exhibits 6 transitions (4 C → T,

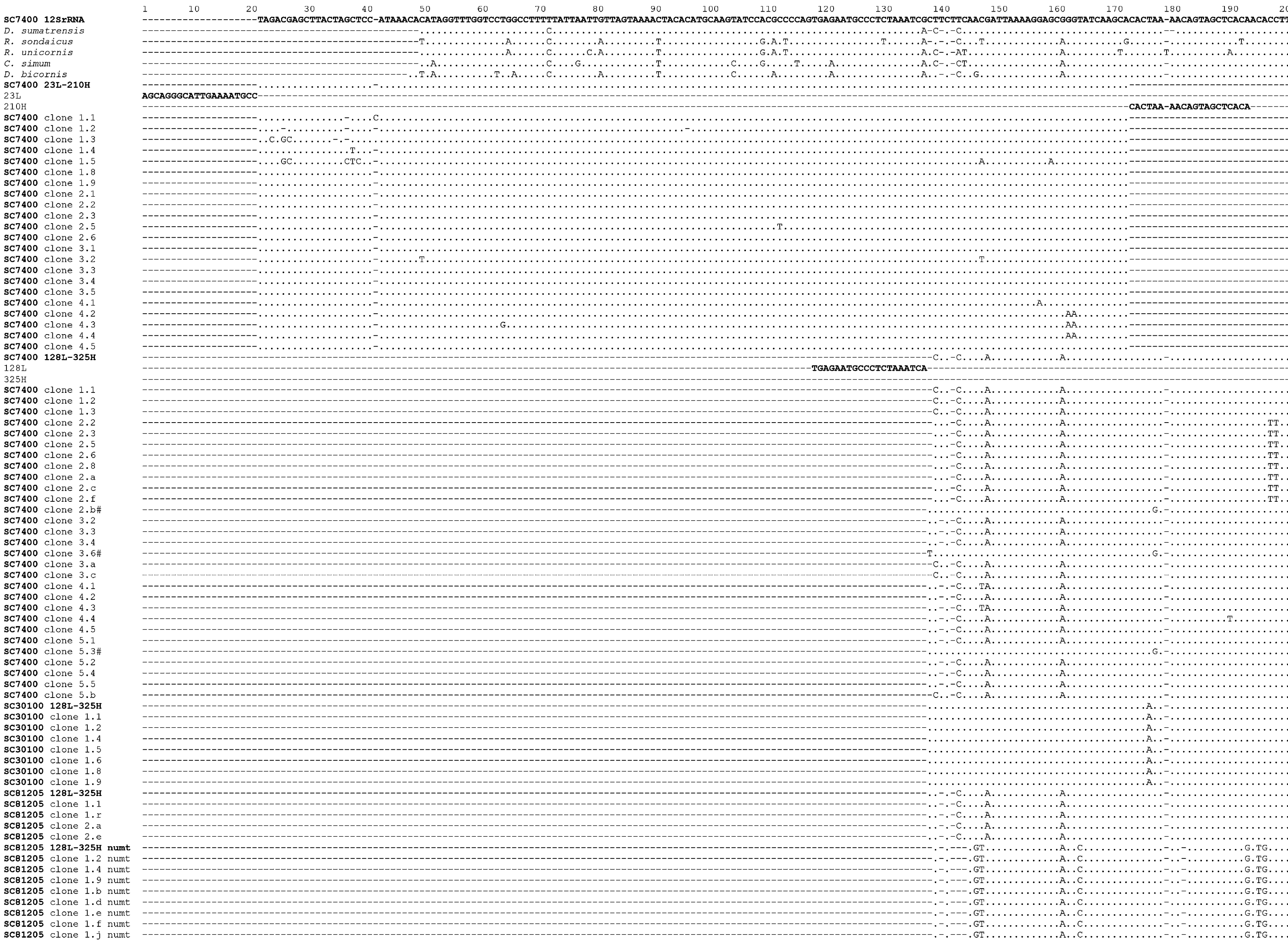

Fig. 2. Alignment of the clones from the *12S rRNA* PCR fragments. Identity to the *12S rRNA* sequence of sample SC7400 is indicated by dots. Amplifications performed in the UCLA laboratory are reported with asterisks (*). Primer names refer to their position along the complete mtDNA genome of *Ceratotherium simum* (Accession No. Y07726, Xu and Arnason, 1997). The three clones of three independent amplifications of the 128L–325H fragment that exhibit perfect match in the overlapping part of the fragment 23L–210H are indicated with #. For each sample, clone W.Z stands for the Zth clone of the Wth amplification of the reported fragment. DAM., highly degraded fragments; numt, nuclear insert of mitochondrial DNA; Y, pyrimidine.

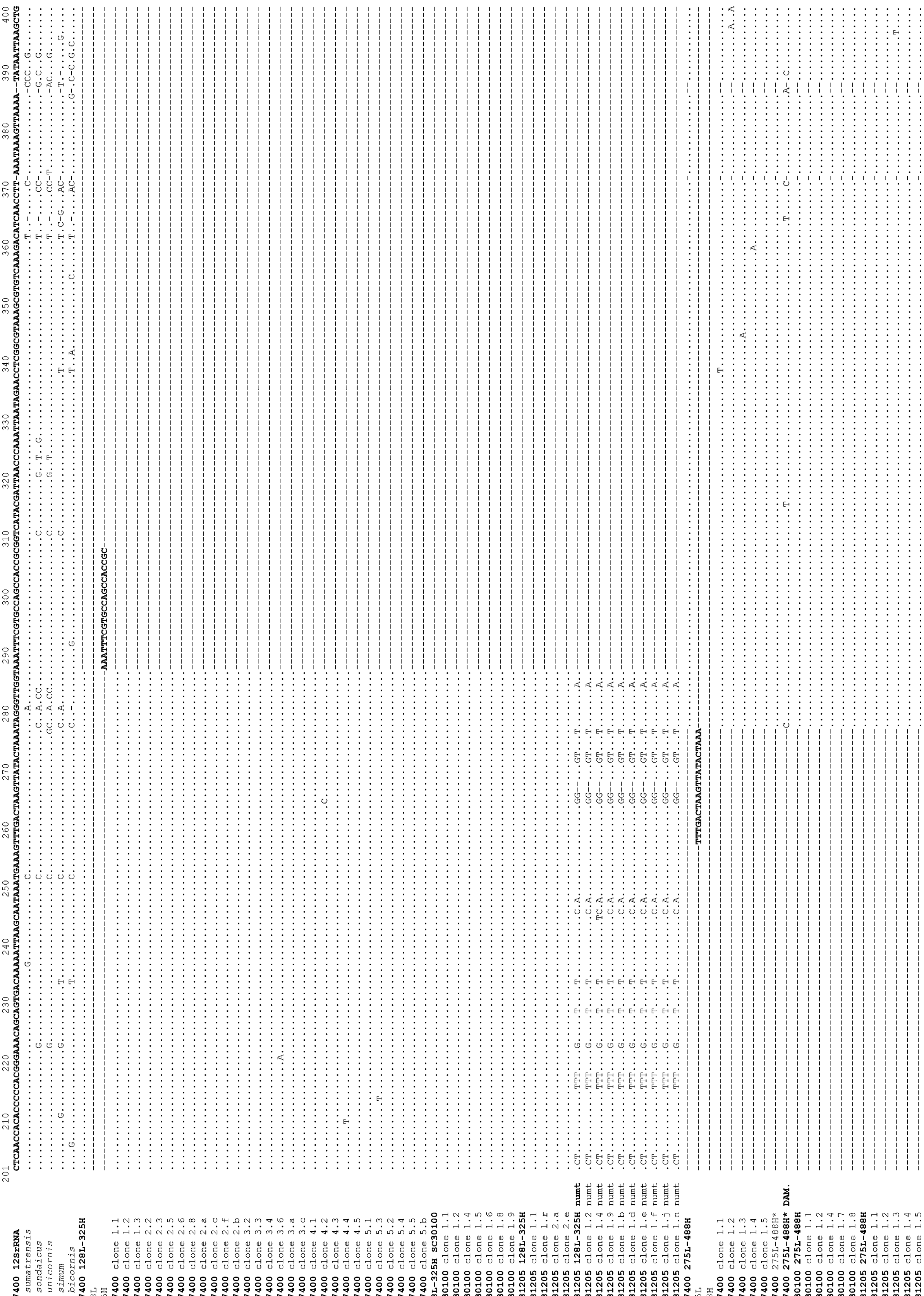

Fig. 2. (continued)

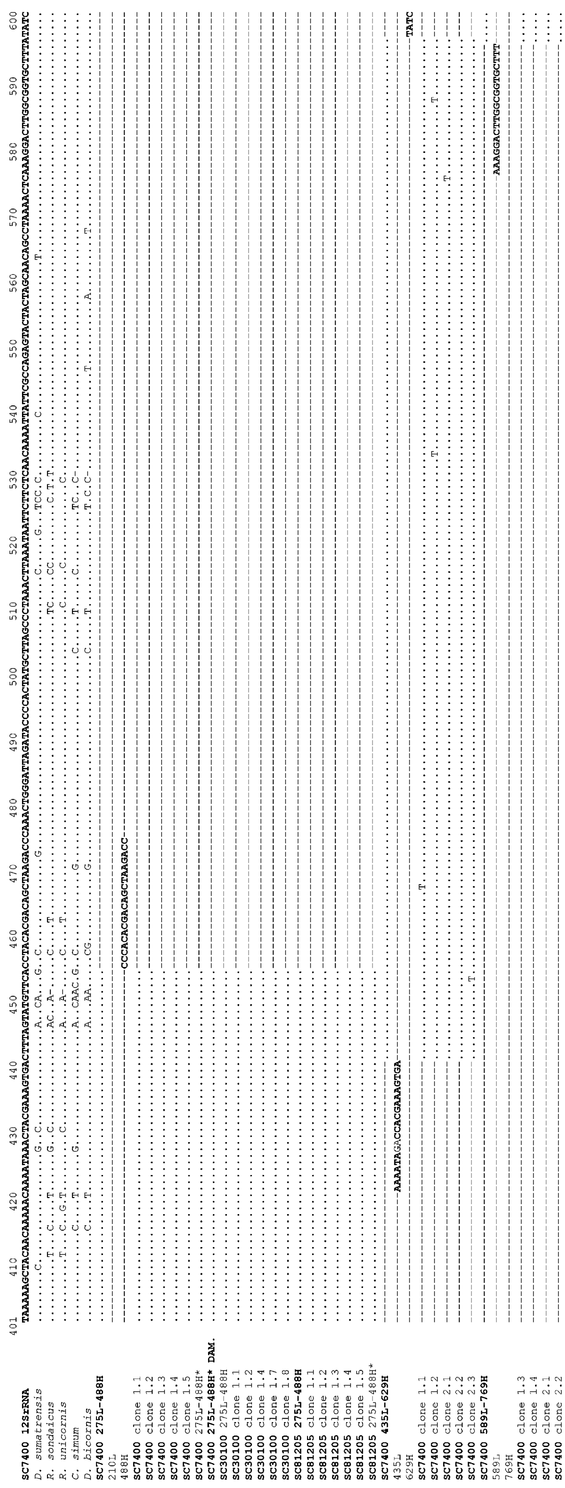

Fig. 2. (*continued*)

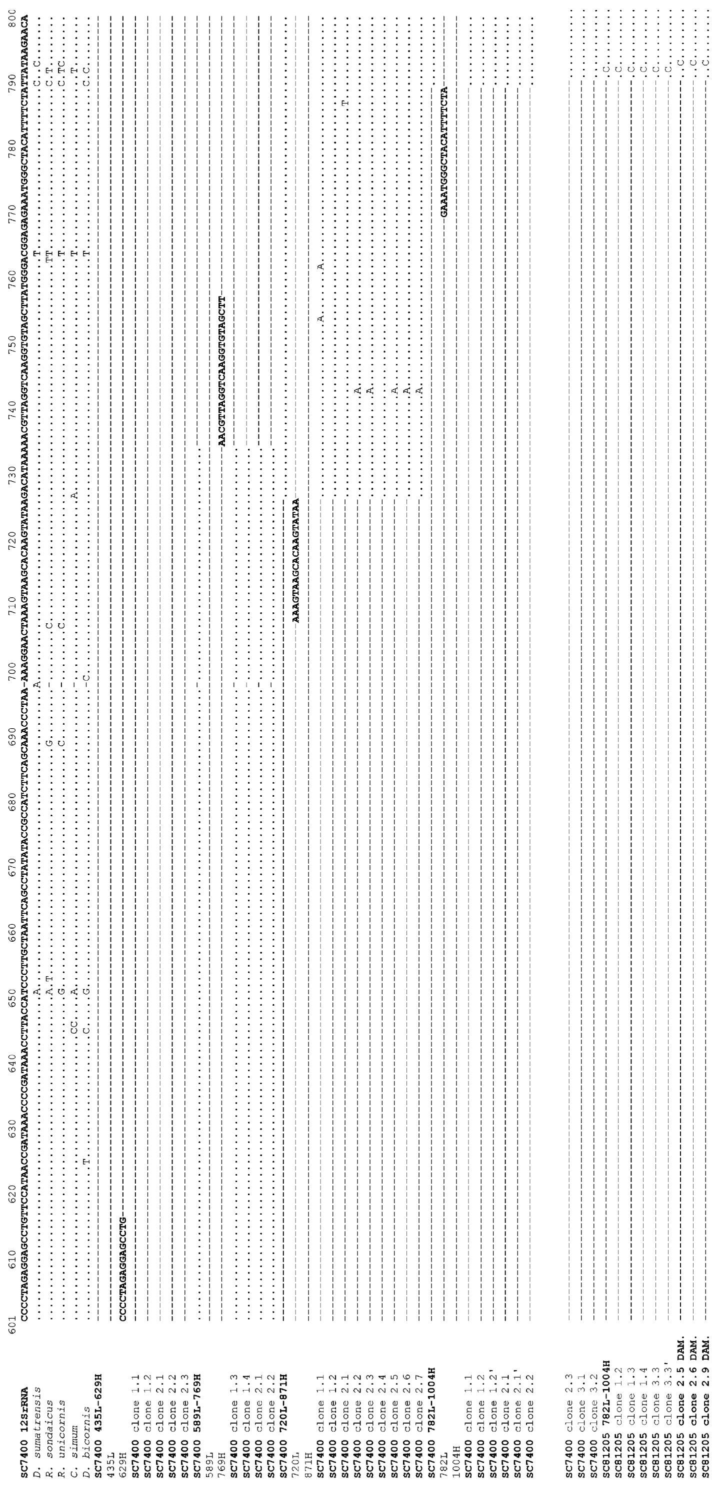

Fig. 2. (*continued*)

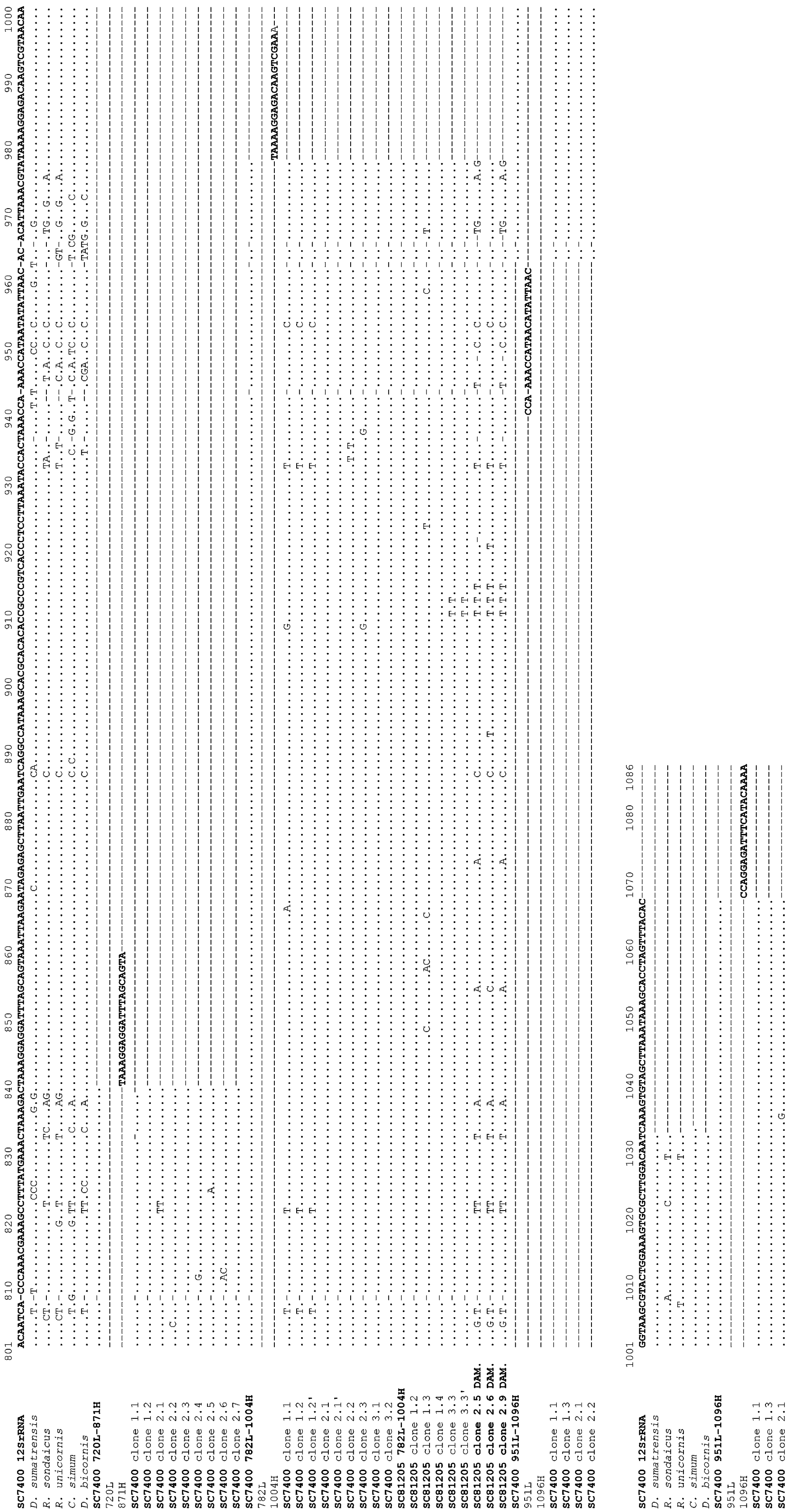

Fig. 2. (*continued*)

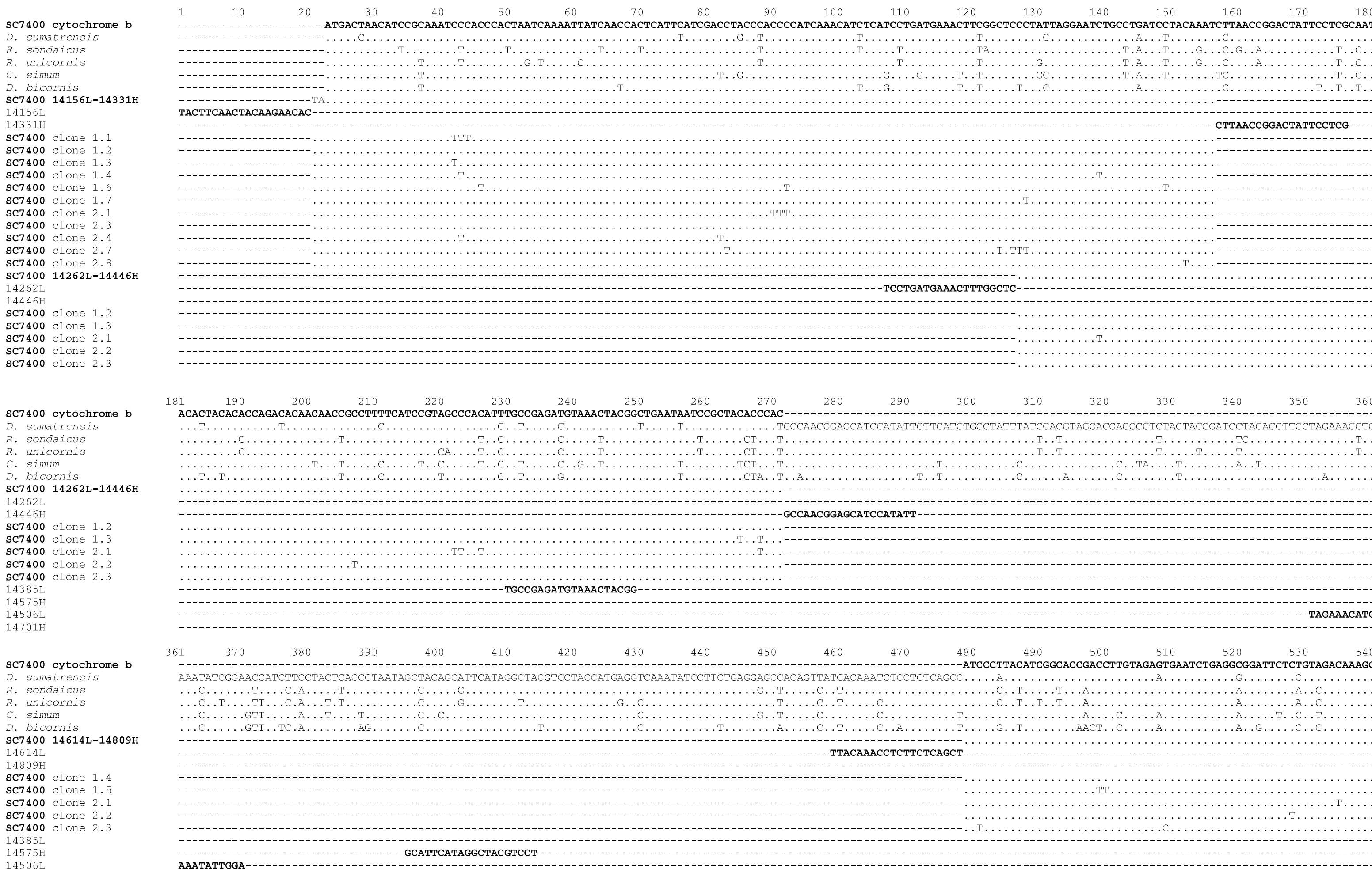

Fig. 3. Alignment of the clones from *cyt b* PCR fragments. Identity to the cytochrome *b* sequence of sample SC7400 is indicated by dots. Amplifications performed in the UCLA laboratory are reported with asterisks (*). Primer names refer to their position along the complete mtDNA genome of *Ceratotherium simum* (Accession No. Y07726, Xu and Arnason, 1997). For each sample, clone W.Z stands for the Zth clone of the Wth amplification of the reported fragment. numt, nuclear insert of mitochondrial DNA; Y, pyrimidine.

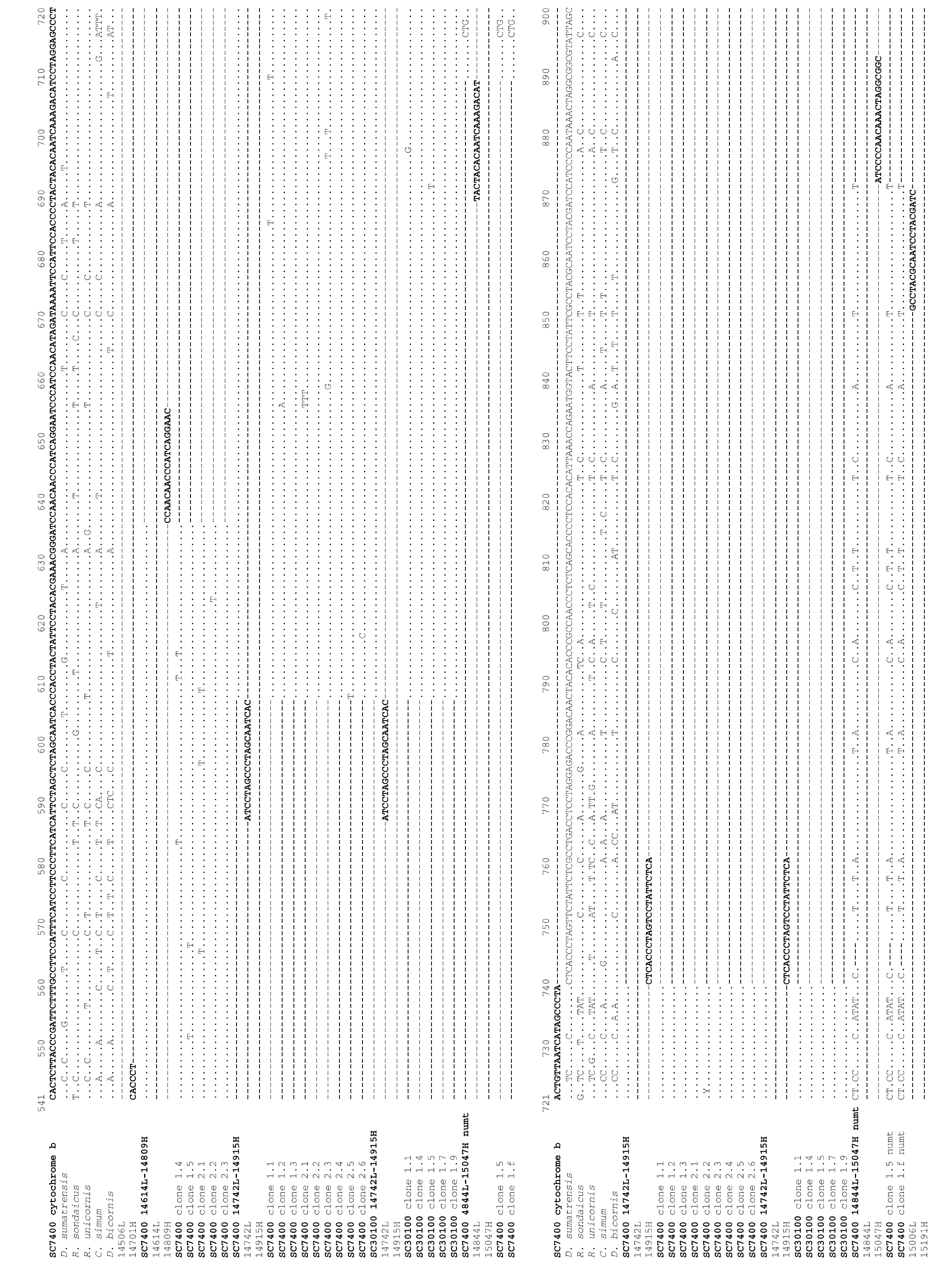

Fig. 3. (*continued*)

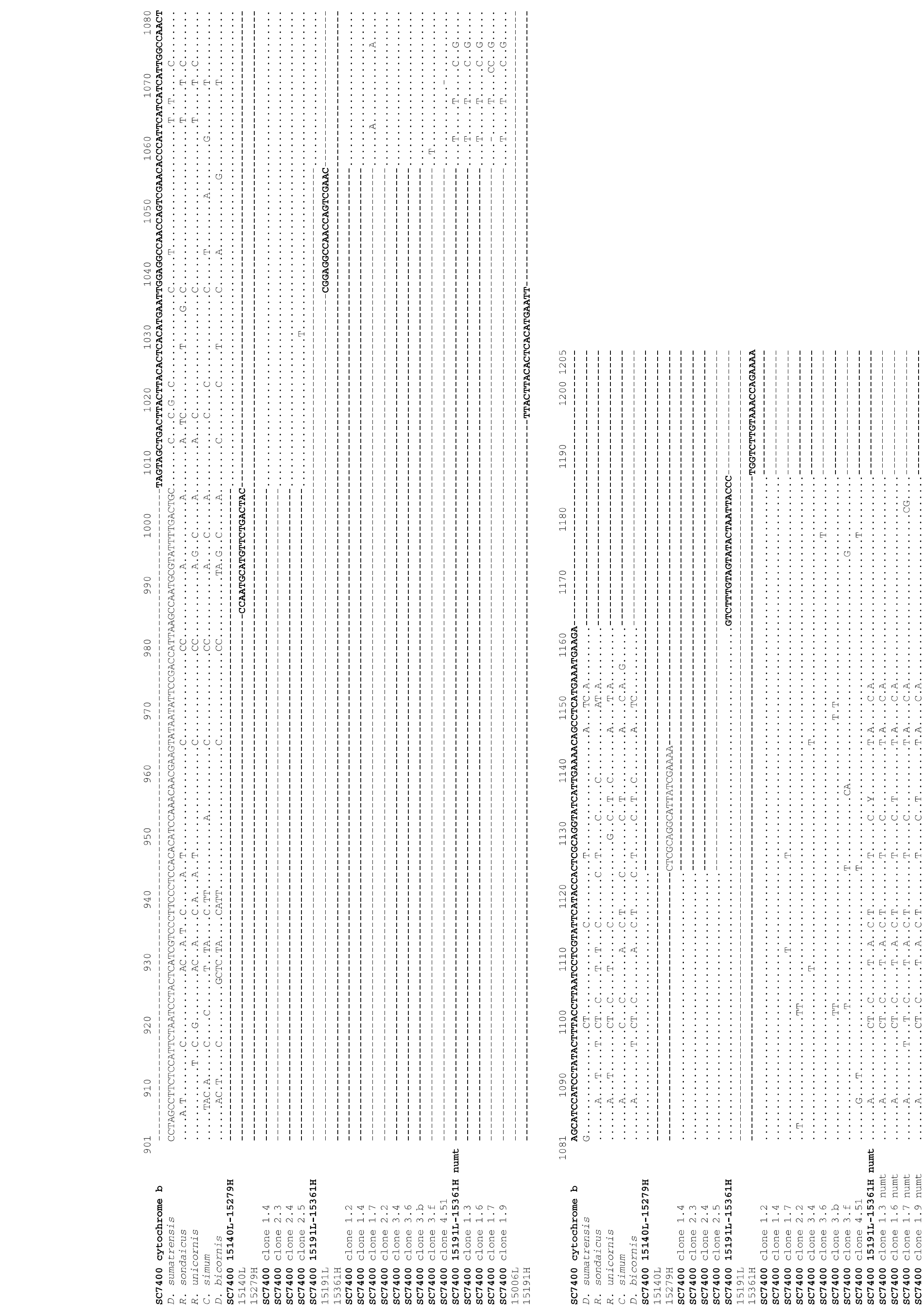

Fig. 3. (continued)

1 A → G, and 1 G → A) and 1 transversion (A → C) with regards to the sequence from sample SC7400. Likewise, in the 14 bp overlapping 951L–1096H fragment, 4 transitions (1 C → T, 2 A → G, and 1 G → A) and 1 indel are needed to align the fragment from SC81205 with the sequence from sample SC7400. Such substitutions profiles, strongly biased towards GC → AT transitions, could correspond more to amplification of degraded templates, rather than numts (Hofreiter et al., 2001).

3.3. Phylogenetic analyses

In single gene phylogenies (*12S rRNA* or *cyt b*), *C. antiquitatis* always fits into a strongly supported monophyletic clade, consisting of all extant rhinoceros sequences (Table 1). *D. sumatrensis* appears to be the extant species most closely related to the extinct woolly rhino, but bootstrap support varies between 51 and 100% depending on the method employed.

By combining the two gene sequences into a single analysis, the bootstrap support within Rhinocerotidae increases. We plotted pairwise substitutions between sequences of *12S rRNA* as a function of those observed in *cyt b* (data not shown). A significant linear correlation indicates that the two genes reflect the same evolutionary history, and therefore can be combined into a single analysis (Montgelard et al., 1997). In all phylogenetic analyses, *C. antiquitatis* appears inside the highly supported monophyletic Rhinocerotidae family clade (bootstrap value between 99 and 100%, and Table 1). Whichever phylogenetic method was used, the sister status of *C. antiquitatis* and *D. sumatrensis* received very strong bootstrap support (Fig. 4 and Table 1).

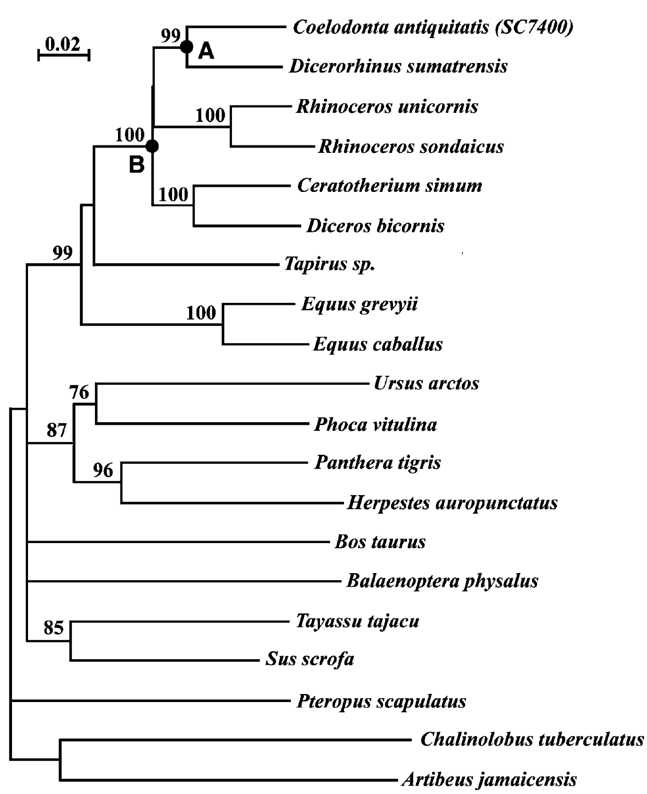

Fig. 4. Phylogenetic position of *Coelodonta antiquitatis* among perissodactyls. The tree was rooted with non-cetferungulate sequences (Arnason et al., 2002). Tajima and Nei corrected distance tree deduced from combined *12S rRNA* and *cyt b* data. The analysis excludes indels, the *12S rRNA* terminal hypervariable region and non-conservative susbtitutions in *cyt b*. The numbers at the nodes refer to percent bootstrap support (1000 replicates). Only bootstrap values greater than 70 are indicated.

4. Discussion

4.1. Phylogenetic position of the woolly rhino

Based on morphological criteria, most of the rhino species from the Pleistocene, including *C. antiquitatis*, are suspected to be closely related to the living Sumatra rhino *D. sumatrensis* (Kurten, 1968). The genus *Dicerorhinus* is typified by several morphological characters: long-nasal bones which are not fused at their rostral end, a partially ossified nasal division and the development of an anterior dentition absent in other rhinoceros species (Guérin, 1980a,b; Guérin, 1989). By many morphological criteria, the genus *Coelodonta* appears to be an extreme example of the *Dicerorhinus* lineage: nasal division completely ossified, low head bearing, very hypsodont jugal teeth, heavy and thick limbs, all already present in *Brandtorhinus* individuals, a sub-genera of the *Dicerorhinus*. Based on these characters, *C. antiquitatis* was grouped inside the *Dicerorhininae* lineage, including several others extinct *Dicerorhinus* species (Guérin, 1980a). Kurten also suggested that all extinct Pleistocene rhinos were related to *Dicerorhinus*, except the gigantic *Elasmotherium* (Kurten, 1968). Using molecular data, the present study strongly supports the hypothesis that the woolly rhino is a member of the *Dicerorhinus* lineage. The convergence of molecular and morphological data for the woolly rhino is in contrast to the case of the woolly mammoth for which ancient DNA studies have not been able to resolve phylogenetic relationships between extant and extinct taxa (Hagelberg et al., 1994; Noro et al., 1998; Ozawa et al., 1997). We infer that this lineage appeared 21–26 MYA during the Oligocene, based on the fossil record demonstrating the Cetartiodactyla radiation at 60 MYA or using the split between equids and ceratomorphs (tapirs and rhinoceroses) at 56 MYA (Kimura-2 corrected distances, Garland et al., 1993; Xu et al., 1996; Xu and Arnason, 1997). Both calculations fit the 23 MYA origin determined by paleontologists, based on the earliest appearance of the most primitive known representative *Dicerorhinus leakeyi* 19.5 MYA. Our divergence calculation should however be taken as preliminary since the same calibration fails to recover the about 2 MYA for the date of emergence of the *Equus* genus (Eisenmann and Baylac, 2000; Oakenfull et al., 2000). Taken as a whole, the woolly rhino offers a case of convergence between morphology and molecules, both for phylogenetic relationships and divergence dates.

4.2. Divergent endogenous DNA sequences

Mitochondrial inserts in the nuclear genome were retrieved both in the *cyt b* gene (fragments 14844L–15047H and 15191L–15361H from sample SC7400) and in the *12S rRNA* gene (fragment 128L–325H from sample SC81205). We thus have retrieved 40–70,000-year-old nuclear DNA. Greenwood et al. also described the recovery of a 180 bp nuclear fragment of 28S rDNA in a 33,000-year-old cave bear bone specimen (Greenwood et al., 1999) and the amplification of a 160 bp fragment of nuclear endogenous retrovirus in a 26,000-year-old mammoth sample preserved in the permafrost (Greenwood et al., 2001). Thus in the future, it will be important to consider the possibility of recovering nuclear DNA even from very old specimens.

Two *12S rRNA* gene PCR products, amplified from SC7400 with 275L–488H primers or from SC81205 with 720L–871H primers, also exhibit sequences very divergent from the consensus. Respectively, 50% (2/4) and 68% (13/19) of the transitions exhibited with SC7400 consensus sequences are of G/C $\rightarrow$ A/T type. Such substitutions could have appeared during the amplification step due to the very deaminated nature of pyrimidines in ancient templates (Höss et al., 1996) as has been recently demonstrated on bone and teeth samples that vary in age between 25,000 to over 50,000 years (Hofreiter et al., 2001). Our results reinforce the view that ancient consensus sequences must be deduced from the cloned sequences of several amplification products to avoid misleading substitutions due to post-mortem degradation events (Greenwood et al., 2001; Hofreiter et al., 2001; Loreille et al., 2001).

Most ancient DNA studies are based on mitochondrial DNA because of the higher copy number and availability of sequences from extant species for comparison. As a consequence of DNA decay, nuclear fragments have been infrequently targeted as they are thought to be less conserved in fossils. Here, we highlight that even studying very-old samples does not prevent numt recovery which could lead, without careful sequence attention, to biased phylogenetic analyses.

Acknowledgments

We thank H. Bocherens, M. Otte, D. Bonjean, and M. Patou-Mathis for the generous gift of bone samples, L. Granger and C. Lerondel for technical assistance in sequencing, E. Douzery and M. Robinson-Réchavi for critical reading of the manuscript, and R.-K. Wayne for laboratory facilities. We warmly thank J.J. Jaeger and F.M. Catzeflis for constant support in the race for rhino DNA recovery. The following organisations supported this work: CNRS (APN), MENRT (ACI), Université Claude Bernard (BQR), and NSF. L.O. receives a fellowship from Ecole Normale Supérieure de Lyon.

References

Arnason, U., Adegoke, J.A., Bodin, K., Born, E.W., Esa, Y.B., Gullberg, A., Nilsson, M., Short, R.V., Xu, X., Janke, A., 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proc. Natl. Acad. Sci. USA 99, 8151–8156.

Cooper, A., Fortey, R., 1998. Evolutionary explosions and the phylogenetic fuse. Trends Ecol. Evol. 13, 151–155.

Dorst, J., Dandelot, P., 1972. Guide des grands Mammifères d'Afrique. Delachaux & Niestlé edit., Neuchâtel.

Douzery, E., Catzeflis, F.M., 1995. Molecular evolution of the mitochondrial *12S rRNA* in Ungulata (Mammalia). J. Mol. Evol. 41, 622–636.

Eisenmann, V., Baylac, M., 2000. Extant and fossil *Equus* (Mammalia, Perissodactyla) skulls: a morphometric definition of the subgenus *Equus*. Zool. Scripta 29, 89–100.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.

Garland, T.J., Dickerman, A.W., Janis, C.M., Jones, J.A., 1993. Phylogenetic analysis of covariance by computer simulation. Syst. Biol. 42, 265–292.

Greenwood, A.D., Capelli, C., Possnert, G., Pääbo, S., 1999. Nuclear DNA sequences from late Pleistocene megafauna. Mol. Biol. Evol. 16, 1466–1473.

Greenwood, A.-D., Lee, F., Capelli, C., De Salle, R., Tikhonov, A., Marx, P., Mac Phee, R.D., 2001. Evolution of endogenous retrovirus-like elements of the woolly mammoth (*Mammuthus primigenius*) and its relatives. Mol. Biol. Evol. 18, 840–847.

Guérin, C., 1980a. Les rhinocéros (Mammalia, Perissodactyla) du Miocène terminal au Pléistocène supérieur en Europe occidentale: comparaison avec les espèces actuelles. Docum. Lab. Géol. Lyon 79, 785–1185.

Guérin, C., 1980b. A propos des rhinocéros (Mammalia, Perissodactyla), néogènes et quaternaires d'Afrique: essai de synthèse sur les espéces et sur les gisements. Proc. 8th Panaf. Congress. Prehist. Quat. Stud. 58–63.

Guérin, C., 1989. La famille des Rhinocerotidae (Mammalia, Perissodactyla): systématique, histoire, évolution, paléoécologie. Cranium 2, 3–14.

Hagelberg, E., Thomas, M.G., Cook Jr., C.E., Sher, A.V., Baryshnikov, G.F., Lister, A.M., 1994. DNA from ancient mammoth bones. Nature 370, 333–334.

Halthenorth, T., Diller, H., 1985. Mammifères d'Afrique et de Madagascar. Delachaux & Niestlé edit., Neuchâtel.

Hänni, C., Laudet, V., Stéhélin, D., Taberlet, P., 1994. Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. Proc. Natl. Acad. Sci. USA 91, 12336–12340.

Hofreiter, M., Jaenicke, V., Serre, D., von Haeseler, A., Pääbo, S., 2001. DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. Nucleic Acids Res. 29, 4793–4799.

Höss, M., Jaruga, P., Zastaway, T.H., Dizdaroglu, M., Pääbo, S., 1996. DNA damage and DNA sequence retrieval from ancient tissues. Proc. Natl. Acad. Sci. USA 24, 1304–1307.

Irwin, D.M., Kocher, T.B., Wilson, A.C., 1991. Evolution of the cytochrome *b* gene of mammals. J. Mol. Evol. 32, 128–144.

Kurten, B., 1968. Pleistocene Mammals of Europe. Weindenfeld and Nicholson, London.

Loreille, O., Orlando, L., Patou-Mathis, M., Philippe, M., Taberlet, P., Hänni, C., 2001. Ancient DNA analysis reveals divergence of the cave bear, *Ursus spelaeus*, and brown bear, *Ursus arctos*, lineages. Curr. Biol. 11, 200–203.

Montgelard, C., Catzeflis, F.M., Douzery, E., 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.

Noro, M., Masuda, R., Dubrovo, I.A., Yoshida, M.C., Kato, M., 1998. Molecular phylogenetic inference of the woolly mammoth *Mammuthus primigenius*, based on complete sequences of mitochondrial cytochrome *b* and 12S ribosomal RNA genes. J. Mol. Evol. 46, 314–326.

Oakenfull, E.A., Lim, H.N., Ryder, O.A., 2000. A survey of equid mitochondrial DNA: implications for the evolution, genetic diversity and conservation of *Equus*. Conservation Genet. 1, 341–355.

Orlando, L., Bonjean, D., Bocherens, H., Thénot, A., Argant, A., Otte, M., Hänni, C., 2002. Ancient DNA and the population gentics of cave bears (*Ursus spelaeus*) through space and time. Mol. Biol. Evol. 19, 1920–1933.

Otte, M., Patou-Mathis, M., Bonjean, D., 1998. La grotte Scladina, vol. 2: L'archéologie. Eraul Liège, Belgium.

Ozawa, T., Hayashi, S., Mikhelson, V.M., 1997. Phylogenetic position of mammoth and Steller's sea cow within Tethytheria demonstrated by mitochondrial DNA sequences. J. Mol. Evol. 44, 406–413.

Prothero, D.-R., Guérin, C., Manning, E., 1989. The history of the Rhinocerotoidea. In: The Evolution of Perissodactyls. Oxford Univ. Press, New York.

Sorenson, M.D., Cooper, A., Paxinos, E.E., Quinn, T.W., James, H.F., Olson, S.L., Fleischer, R.C., 1999. Relationships of the extinct moa-nalos, flightless Hawaiian waterfowl, based on ancient DNA. Proc. R. Soc. Lond. B Biol. Sci. 266, 2187–2193.

Thew, N., Chaix, L., Guérin, C., 2000. Dernier cycle glaciaire et occupations paléolithiques à Alle, Noir Bois (Jura, Suisse). La faune, 93–98. D. Aubry, M. Guélat, J. Detrey, and B. Othenin-Girard. Cahier d'archéologie jurassienne, Porrentruy.

Tougard, C., Delefosse, T., Hänni, C., Montgelard, C., 2001. Phylogenetic relationships of the five extant Rhinoceros species (Rhinocerotidae, Perissodactyla) based on mitochondrial cytochrome *b* and *12S rRNA* genes. Mol. Phylogenet. Evol. 19, 34–44.

Vila, C., Leonard, J.A., Gotherstrom, A., Marklund, S., Sandberg, K., Liden, K., Wayne, R.K., Ellegren, H., 2001. Widespread origins of domestic horse lineages. Science 291, 474–477.

Xu, X., Janke, A., Arnason, U., 1996. The complete mitochondrial DNA sequence of the greater Indian rhinoceros, *Rhinoceros unicornis*, and the phylogenetic relationship among Carnivora, Perissodactyla, and Artiodactyla (+Cetacea). Mol. Biol. Evol. 13, 1167–1173.

Xu, X., Arnason, U., 1997. The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum*, and comparison with the mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis*. Mol. Phylogenet. Evol. 7, 189–194.