Molecular Ecology (2006) **15**, 2083–2093 doi: 10.1111/j.1365-294X.2006.02922.x

Geographic distribution of an extinct equid (*Equus hydruntinus*: Mammalia, Equidae) revealed by morphological and genetical analyses of fossils

LUDOVIC ORLANDO,* MARJAN MASHKOUR,† ARIANE BURKE,‡ CHRISTOPHE J. DOUADY,§ VÉRA EISENMANN¶ and CATHERINE HÄNNI*

**Paléogénétique et Evolution moléculaire, CNRS UMR 5161, INRA LA 1237, Laboratoire de Biologie Moléculaire de la Cellule, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon cedex 07, France, †UMR 5197 Muséum national d'histoire naturelle/CNRS — Département d'Ecologie et Gestion de la Biodiversité — Case postale 56, Bâtiment d'Anatomie comparée-55, rue Buffon, F-75005 Paris, France, ‡Department of Anthropology, University of Manitoba, Winnipeg, MB, Canada R3T 5V5, §CNRS UMR 5023, Laboratoire d'Ecologie des Hydrosystèmes Fluviaux, Université Claude Bernard Lyon 1, 6 rue R. Dubois, Bat. Darwin-C, 69622 Villeurbanne cedex, France, ¶CNRS UMR 5143, Muséum National d'Histoire Naturelle, Département Histoire de la Terre, CP 38, 75005 Paris, France*

Abstract

***Equus hydruntinus* inhabited Europe and the Middle East for more than 300 000 years. For a long time, palaeontological data failed to place *E. hydruntinus* into the equid phylogenetic tree, confronted with the fact that it shares primitive *Equus* characters with both zebras and asses, and derived characters with asses and hemiones. However, the study of a recently discovered skull points to a relationship with hemiones. Extraction of DNA from ancient samples from Crimea (*E. hydruntinus*) and Iran (*E.* cf. *hydruntinus*) yielded 134–288 bp of the mtDNA control region and 143 bp of the cytochrome *b* gene. This DNA analysis supports the proximity of *E. hydruntinus* and *Equus hemionus* suggested by skull and limb bone analyses, and rejects proximity to either *Equus burchelli* or the asses suggested by tooth morphology. Dental morphology may thus be of poor taxonomical value if used alone for establishing equid phylogenetic relationships. Furthermore, the small genetic distance between *E.* cf. *hydruntinus* of Iran and the classical *E. hydruntinus* of Crimea suggests that both samples belong to the same species. Accordingly, the geographic range of *E. hydruntinus* — until now believed to be restricted to Europe, Israel, and Turkey — can be extended towards East as far as Iran.**

Keywords: ancient DNA, equids, *Equus hydruntinus*, mtDNA, past biodiversity, taxonomy

Received 7 December 2005; revision accepted 22 January 2006

Introduction

Since the equid family has been a famous model to describe evolutionary processes for more than a century (Huxley 1870; Gould 1994), it is hard to imagine that the phylogenetic relationships between some equids are still debated. It should be remembered, however, that only skulls, teeth, and bones fossilize, not coat colour, stripes, and behavioural traits, which are used as tools for stating the taxonomy status of extant individuals. Among fossilized remains, skulls are the best taxonomic indicators as they enable researchers to discriminate between the six most commonly recognized extant subgenera (Eisenmann 1979, 1998; Groves & Willoughby 1981). Although less reliable for taxonomy, teeth or limb bones provide useful insights into ecology and feeding. In this context, the taxonomy of the extinct *Equus hydruntinus*, almost exclusively represented by teeth and limb bones until very recently, has always been problematic.

The oldest *E. hydruntinus* specimens excavated in Lunel-Viel (South of France) are about 350 000 years old (Bonifay 1991). Numerous fossils have also been found in Italy and Crimea. The extinction of *E. hydruntinus* is relatively recent: the species still appears in some European Iron Age sites (Wilms 1989) and may be mentioned in Portuguese manuscripts dating to the Middle Ages (Antunes 2006). *E. hydruntinus* is an enigmatic species, combining primitive

Correspondence: Ludovic Orlando, Fax: 33 4 72 72 80 00; E-mail: ludovic.orlando@ens-lyon.fr

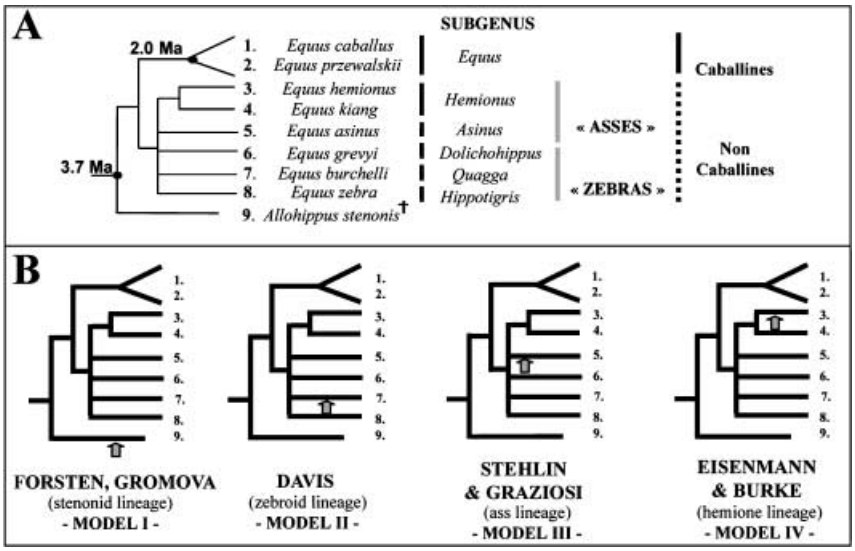

Fig. 1 Taxonomy of *Equus* and relationships of *Equus hydruntinus*. (A) Taxonomy of extant *Equus* and Pliocene *Allohippus*. The problematic nodes are represented with polytomy. (B) Four competing palaeontological models. The hypothetic relationship between *E. hydruntinus* and extant equid species is symbolized by a grey arrow. *E. hydruntinus* has been related to *Allohippus stenonis* (model I; Gromova 1949; Forsten & Ziegler 1995), *Equus burchelli* (model II; Davis 1980), and the subgenus *Asinus* (model III; Stehlin & Graziosi 1935). Stehlin & Graziosi (1935) however, showed that the slender and cursorial proportions of the limbs fit best with hemiones. This relation (model IV) is now confirmed by skull morphology (Burke *et al*. 2003).

characters observable 2 million years ago (Ma) in the Eurasian *Allohippus stenonis* and in African zebras today with derived features typical of modern hemiones and asses. Figure 1 summarizes the lack of consensus between the palaeontological models describing the phylogenetic relationships between *E. hydruntinus* and other equids. Because of its slenderness and cursorial limb proportions, *E. hydruntinus* is most commonly referred as the 'European ass'. But when the teeth, rather than the limbs, are considered, *E. hydruntinus* appears closely related to modern plains zebra (Davis 1980) or to Pliocene stenonines (Gromova 1949; Forsten 1986; Forsten & Ziegler 1995). Formally, *E. hydruntinus* has been placed in the subgenus *Asinus* (Stehlin & Graziosi 1935). Other authors believe that it should belong in a subgenus of its own, *Hydruntinus* (Radulesco & Samson 1962).

As classic studies of teeth and limb proportions lead to opposite conclusions, new data and approaches are needed to address the phylogenetic origin of *E. hydruntinus*. New data provided by the excavation of a nearly complete skeleton in Crimea (Burke *et al*. 2003) and new material from Iran (Eisenmann & Mashkour 1999) have enabled an extensive DNA analysis. Indeed, given that mtDNA has already been used to address the phylogenetic relationships of extant (Ishida *et al*. 1995; Oakenfull *et al*. 2000) and extinct (Higuchi *et al*. 1984, 1987; Orlando *et al*. 2003; Weinstock *et al*. 2005) equids, the ancient DNA technology appears adequate to investigate the relationships between extant equids and *E. hydruntinus*.

Materials and methods

Samples

Five teeth were obtained for ancient DNA analysis. One sample (MM) was obtained from an *Equus hydruntinus* skull excavated in Crimea (Ukraine). Although not perfectly preserved, the Crimean skull is currently the best-known sample for this species (Burke *et al*. 2003). The locality of origin, Kabazi II, is an open-air Middle Palaeolithic site in western Crimea, where other faunal indicators point to the existence of open, dry grasslands (Burke *et al*. 1999). Other samples used in this study were excavated from two sites in the Qazvin plain (Iran) in cold steppic contexts (Mashkour *et al*. 1999; Mashkour 2001): Sagzabad (samples CH22, CH25, and CH26; 1264–863 BC) and Zagheh (sample CH31; 5212–4561 BC). Samples CH22 and CH25 were referred to *E.* cf. *hydruntinus*, since their protocone lengths are similar to classic *E. hydruntinus* fossils from Lunel-Viel and Crimea (Eisenmann & Mashkour 1999). Samples CH26 and CH31, assigned to *Equus caballus* and *Equus hemionus*, respectively (Eisenmann & Mashkour 1999), were analysed (i) to trace any possible contamination (see below), and (ii) to compare *E. hydruntinus* sequences to other equid sequences from the same period of time. The ages of the Sagzabad and Zagheh deposits have been determined through radiocarbon dating of eight bone remains (Mashkour *et al*. 1999). Whole amount of nitrogen per sample, the proportion of nitrogen, and C/N ratios in collagen extracts from 15 and 12 bone or dental remains from Zagheh and Sagzabad, respectively, show values within the range of modern or well-preserved samples, suggesting that fossilization conditions in these deposits are consistent with ancient DNA preservation (Bocherens *et al*. 2000; Gilbert *et al*. 2005).

DNA extraction, amplification, cloning and sequencing

Extraction and amplification were conducted in separate rooms devoted to ancient DNA work as described in Hänni *et al*. (1994) and Orlando *et al*. (2002). The outer layers of the teeth were removed by scraping with a sterile razor blade.

The samples CH25 (1.2 g) and CH26 (1.4 g) were first co-extracted; the samples CH22 (1.3 g) and CH31 (2.4 g) were co-extracted in a second extraction session. Two months later, the extraction of sample MM (1.0 g) was concomitant with extraction of ancient tissue from a wolf. Mock extractions were carried out during each session to monitor contamination.

To target the mtDNA control region (mtDNA CR), we took advantage of three pairs of primers described in Vila *et al.* (2001; L1-H1, L2-H2 and L3-H3 primers). We designed two other pairs of equid-specific primers to cover 296 bp in the cyt *b* gene (cytb1L 5′-CTAATTAAAATCATCAATC and cytb1H 5′-ATAATTCATCCGTAGTTA; cytb2L 5′-AACTGCCTTCTCATCCGTCA and cytb2H 5′-AAAAG-TAGGATGATTCCAAT; Orlando *et al.* 2003). Since they target short overlapping fragments (160–210 bp long), all these primers were adapted to the fragmented nature of ancient DNA. Polymerase chain reactions (PCR) were carried out in a total volume of 50 or 100 μL in a PCR Mastergradient apparatus (Eppendorf®). PCR tubes contained 10 U of *Taq* Gold polymerase (PerkinElmer®), 2–3 mM $MgCl_2$, 1 mg/mL BSA, 250 μM of each dNTP and 300 ng of each primer. Depending on the sample, 0.5–3 μL of ancient extract was added to avoid inhibition of the *Taq* polymerase activity. A 10-min activation step at 92 °C was followed by 50 cycles of denaturation (92 °C, 60 s), annealing (48–50 °C, 60 s), extension (72 °C, 45 s) and a last extension step (5 min). Three independent blanks were carried out during each PCR experiment, as reported in Loreille *et al.* (2001).

PCR products were cloned using the TOPO TA cloning kit (Invitrogen®) following the manufacturer's instructions. Plasmids were purified using the QIAprep spin miniprep kit (QIAGEN®). The sequences of both strands were obtained on a Megabace^{1000} automatic capillary sequencer (Amersham®) after PCR amplification in a volume of 10 μL, with 400 ng of DNA, 4 μL of the DYEnamic ET Mix (Amersham®) and 100 ng of a M13 universal primer (50 cycles, denaturation at 94 °C, 20 s, annealing and extension at 60 °C, 1 min 15 s).

Sequence analyses

Sampling, saturation and data construction. Taxonomic sampling is a key step in phylogenetic inferences. For each gene, available sequences from equid species were extracted from GenBank. Our mtDNA CR data set accounts for 16 extant (Hg refers to the haplogroup labelled in Vila *et al.* 2001; HgA: AF326661, AF326647; HgB: AF072989: Lister *et al.* 1998, AF014411: Kim *et al.* 1999; HgC: AF326662, AF168694 and AF168696: Kavar *et al.* 1999, AF072988: Lister *et al.* 1998; HgD: D23665: Ishida *et al.* 1994, AF326659, AF326660, AF326664; HgE: D14991: Ishida *et al.* 1994, AF072986: Lister *et al.* 1998; HgF; AF326637, AF056071: Kim *et al.* 1999 and 14 ancient horses (AF326668–AF326686: Vila *et al.* 2001; AY049720: Jung *et al.* 2002), 2 *Equus przewalskii* (AF055878: Oakenfull & Ryder 1998; AF072995: Lister *et al.* 1998), 2 *Equus asinus* (X97337: Xu *et al.* 1996; AF220938: Oakenfull *et al.* 2000), 4 *E. hemionus* (AF220934–AF220937: Oakenfull *et al.* 2000), 2 *Equus kiang* (AF220932, AF220933: Oakenfull *et al.* 2000), 3 *Equus grevyi* (AF220928–AF220930: Oakenfull *et al.* 2000), 9 *Equus burchelli* (AF220916–AF220924: Oakenfull *et al.* 2000), 4 *Equus zebra* (AF220925–AF220927 and AF220931: Oakenfull *et al.* 2000) and 3 *Hippidion saldiasi* (AY15261–AY152863: Orlando *et al.* 2003). Our cyt *b* data set includes 3 extant horses (NC_001640: Xu & Arnason 1994; D32190: Chikuni *et al.* 1995; D82932: Ishida *et al.* 1996), 7 *E. asinus* (X97337: Xu *et al.* 1996; AF380130–AF380135: Aranguren-Mendez *et al.* 2004), 1 *E. burchelli* (AY534349, Kimwele *et al.*, unpublished), 1 *E. grevyi* (X56282: Irwin et al. 1991) and 2 *Hippidion saldiasi* (AY15259–AY15260: Orlando *et al.* 2003). Both gene fragments were aligned manually using the SEAVIEW software (Galtier *et al.* 1996). To avoid random error effect (that mostly occurs when data contain too little phylogenetic information), we analysed each locus separately and also both loci simultaneously. Concatenated data were constructed by merging taxa for which both loci were known from the same individual. The only exceptions were *E. grevyi* and *E. burchelli* for which we had to merge independent CR and cyt *b* sequences. Both *E. grevyi* and *E. burchelli* being known by a single cyt *b* sequences but three and nine CR sequences respectively, we constructed a data (later called 'merged consensus', eight taxa) where the CR sequence was represented by a consensus of the two most divergent sequences. Another merged data (later called 'merged chimer', 10 taxa) was constructed by merging the cyt *b* sequence of these two taxa to each of the two most divergent CR sequences.

Data sets must include relevant species but cannot include highly saturated signal. Thus and since available outgroups (rhino or tapir) are phylogenetically distant from our ingroup, saturation level in each gene was investigated using a methods first developed by Philippe *et al.* 1994). The inferred number of substitutions was estimated from patristic distance on a maximum-likelihood (ML) tree (using a GTR + G + I model in PHYML; Guindon & Gascuel 2003), and compared to the observed number of transitions and transversions. Tested outgroups were two rhinoceroses (*Ceratotherium simum* Y07726, and *Rhinoceros unicornis* X97336) for the CR and five Rhinocerotidae (*Ceratotherium simum* NC_001808, *Diceros bicornis* X56283, *Rhinoceros unicornis* NC_001779, *Rhinoceros sondaicus* AJ245725, and *Dicerorhinus sumatrensis* AJ245723; Irwin *et al.* 1991; Xu *et al.* 1996; Xu & Arnason 1997; Tougard *et al.* 2001) plus two Tapiroidae (*Tapirus indicus* AF145734 and *Tapirus terrestris* AF056030) for the cyt *b*.

Phylogenetic reconstruction. Phylogenetic reconstructions were performed under ML and Bayesian approaches. ML trees were performed with PAUP* 4b10 using the

tree-bisection–reconnection branch swapping algorithm and a starting tree obtained via stepwise addition under the best model of evolution according to the AIC criterion of MODELTEST (Posada & Crandall 1998). Bayesian inferences were performed with MRBAYES version 3.0B4 (Huelsenbeck & Ronquist 2001), using a GTR + G + I model of evolution. The tree-space was explored by four chains over 1 000 000 generations sampled every 100. Burn-in value was fixed at 100 000 generations after empirical determination of the convergence.

Robustness assessment. The strength of the phylogenetic signal was first assessed via nonparametric bootstrapping (Felsenstein 1985) applied to an ML framework. Bootstrapping analyses included 1000 pseudoreplicates and were performed using the same setting as for the optimal reconstruction (above). Bayesian posterior probabilities were also recorded even if their significance, in term of robustness, remains an open question (e.g. Douady *et al.* 2003). Subsequently, statistical supports for *a priori* selected hypothesis (Fig. 1) were assessed via the Approximately Unbiased test (AU, Shimodaira 2002) using CONSEL (Shimodaira & Hasegawa 2001) and unilateral Kishino–Hasegawa (KH) test (Kishino & Hasegawa 1989) using PAUP* 4b10.

Molecular divergences. Finally and to get a first grasp on taxonomic implication of our results, CR molecular divergence within and between equids species were evaluated. Based on the most probable topology with the most likely branch lengths, intraspecies (*Hippidion*, $n = 3$; *E. hemionus*, $n = 5$; *E. caballus*, $n = 33$; *E. zebra*, $n = 4$; *E. burchelli*, $n = 9$; *E. grevyi*, $n = 3$; *E. asinus*, $n = 2$; *E. kiang*, $n = 2$) intra-*E. hydruntinus* ($n = 2$), interspecies and inter-*E. hydruntinus*/*E. hemionus* patristic distances were extracted using APE library (Paradis *et al.* 2004) of R 2.0.0 (R development Core Team 2005).

Results

Authentification criteria

We analysed five fossils from different equid species: *Equus caballus* (CH26), *Equus hemionus* (CH31), *Equus hydruntinus* (MM) and *Equus* cf. *hydruntinus* (CH22 and CH25). Sample CH22 never yielded any amplification product although *Taq* polymerase activity was not inhibited (data not shown) and 32 PCRs were attempted. All other samples except for CH26 delivered the L1–H1 mtDNA CR fragment and the cytb2L–cytb2H cyt *b* fragment accounting for 287 bp (Fig. 2). The L3–H3 mtDNA CR fragment was also retrieved for sample CH25, resulting (primers excluded) in supplementary 155 bp information. Sample CH26 yielded only the L2–H2 fragment of the mtDNA CR (162 bp). Each fragment has been amplified independently two or three times. To summarize, amplification succeeded for all of the samples in 13 out of 188 attempts for the mtDNA CR, and in 6 out of 45 attempts for the cyt *b* gene.

We are confident that no cross-contamination occurred between our samples for the following reasons. (i) For each amplification, PCR blanks remained negative (three per experiment). (ii) Amplifications were not successful for all the samples co-extracted. For instance, no DNA was recovered from sample CH22 although it was co-extracted with sample CH31 that yielded mtDNA CR and cyt *b* fragments, and the L3–H3 mtDNA CR fragment was retrieved from sample CH25, but not from the other samples. (iii) No equid DNA was amplified on the ancient wolf extract although it was on the co-extracted sample MM and numerous PCR were attempted (data not shown). (iv) Samples CH26 and CH31, identified by palaeontologists as *E. caballus* and *E. hemionus*, respectively, exhibit mtDNA CR sequences typical of *E. caballus* and *E. hemionus* (see Fig. 2).

A minimum of 4 and up to 10 clones per amplification product (accounting for a final number of 91 clones) were used to deduce a final consensus sequence devoid of *Taq* polymerase damage-induced errors for each sample. The pattern of substitutions observed among clones was characteristic of ancient DNA templates (data not shown). Base composition of consensus sequences and substitution patterns are similar to the ones observed on homologous mtDNA sequences of extant equids (ti/tv > 4.8, Fig. 2; Xu *et al.* 1996; Xu & Arnason 1997). Neither gap nor stop codon disrupt the cyt *b*–coding frame, suggesting that we amplified authentic mtDNA sequences rather than numts.

Phylogenetic analyses

All the analyses were first conducted using rhinoceroses as outgroups. Unrooted analyses were then performed (by excluding rhinoceroses from the dataset) to take into account the substitution saturation bias detected among the sequences (data not shown). In the resulting topology, the root was constrained by hand according to the topology found in the previous rooted analysis. We used mtDNA CR and cyt *b* sequences to make two kinds of concatenated data sets, the first one representing every possible combination between the CR and cyt *b* sequences present in the data set for a given species (referred to below as the 'merged chimer' data set) whereas the second one consisting only in the concatenation of their consensus (referred to below as the 'merged consensus' dataset). Both approaches, if congruent, permit to exclude reassortment and sampling biases that may be generated while concatenating data. Note that phylogenetic analyses conducted on CR and cyt *b* data sets separately lead to similar results and are not further described here.

In all analyses, the samples cluster inside the extant genus *Equus*, contrary to expectations generated by palaeontological

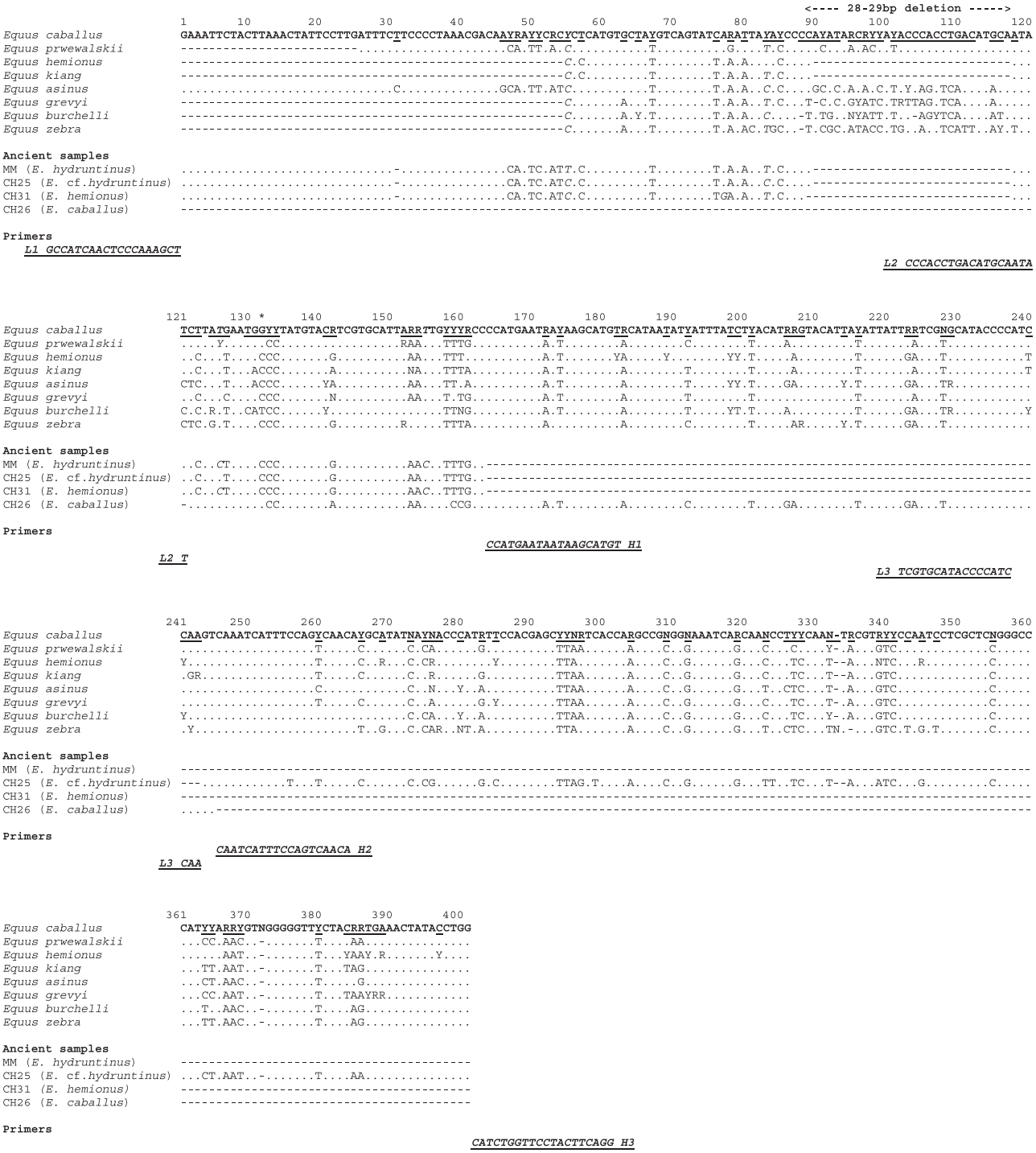

Fig. 2 mtDNA CR alignment between ancient and extant equids. Y = C or T, R = A or G, N = A, C, G, T, or indel. The polymorphic sites between extant equids sequences are underlined. The polymorphic sites between the four ancient sequences are italicized. The asterisk indicates the only transversion observed between the ancient sequences. Dots indicate identity to *Equus caballus*.

model I (Figs 1 and 3). Furthermore, *E.* cf. *hydruntinus* and *E. hydruntinus* samples (CH25 and MM) cluster with *E. hemionus* (Fig. 3). This relationship is in agreement with palaeontological model IV (Fig. 1D) and contra model III (Fig. 1B), which predicted that *E. hydruntinus* would be nested inside *Equus asinus*, and is supported both by strong boostrap values (94–100% of 1000 replicates) and high posterior probabilities (0.98 and 1). Moreover, contrary to palaeontological model II expectations, no relationship between *E. hydruntinus*, *E.* cf. *hydruntinus* and *Equus burchelli*

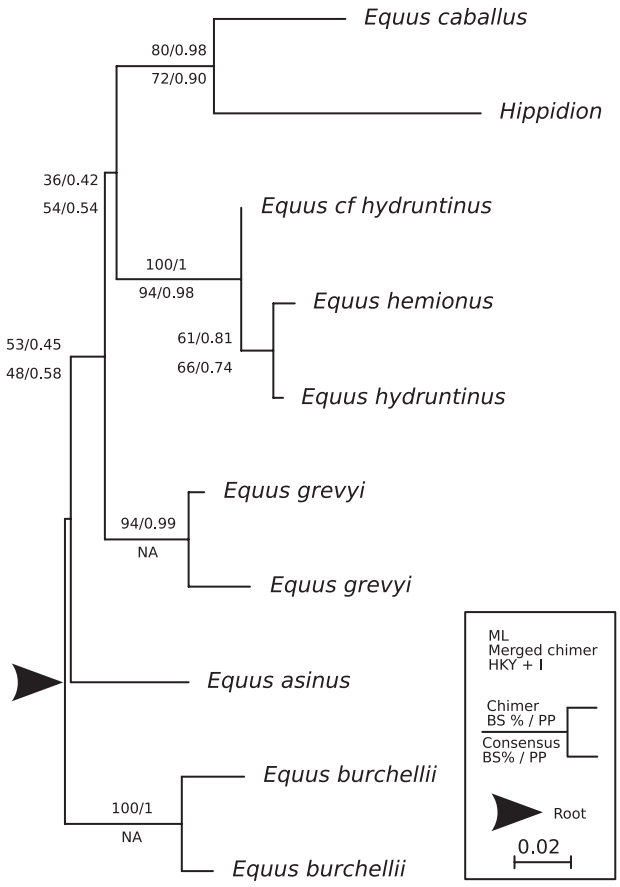

Fig. 3 Most likely tree inferred from our 'merged chimer' data (see Material and methods section). Inference was done under an HKY + I model of substitution. The number above and below branches correspond to bootstrap values and Bayesian posterior probabilities for the 'merged chimer' and the 'merged consensus' data, respectively.

is detected since neither Approximately Unbiased (AU), Kishino–Hasegawa (KH) or Shimodaira–Hasegawa (SH) tests find support for their monophyly (*P* values respectively equal to 0.011-0.025-0.039 and 0.005-0.023-0.032 for the consensus and chimere data sets). To sum up, palaeontological model IV is the only model supported by the molecular phylogenetic analyses.

Interestingly, sequences from samples CH25 and MM share a 28- to 29-bp deletion between positions 89 and 117 with the sequence from sample CH31 (*E. hemionus*) and with sequences from extant *E. hemionus* and *Equus kiang* individuals (Fig. 2). The presence of the deletion on samples CH25, CH31 and MM but not on sample CH26 is also demonstrated by the fact that the L2-H2 mtDNA CR fragment, targeted with the L2 primer which partly overlaps the deletion, was only amplified on the latter sample despite 51 attempts (Fig. 2). Since no other equid species exhibit this deletion, it could have appeared on the ancestral branch of the *E. kiang*/*E. hemionus*/*E. hydruntinus* lineage and thus could be considered as a synapomophy of this clade. This observation reinforces the phylogenetic proximity of *E. hemionus*, *E.* cf. *hydruntinus* and *E. hydruntinus*.

Morphological analyses

In a parallel study, the *E. hydruntinus* skull from Kabazi was compared with many extant and fossil skulls (see description of the samples in Eisenmann & Baylac 2000). Skull measurements were submitted to principal components analysis and mixture discriminant analysis (Burke *et al*. 2003). Results of the study show that (i) the Kabazi skull has a longer cranial length (relative to the length of the palate and muzzle) than *Allohippus*; (ii) the Kabazi skull presents an uncommonly short narial opening, which remains very far from the pattern typical of *A. stenonis*, even if the shortness is possibly exaggerated by postdepositional damage; (iii) the width and shape of the incisor arcade of the Kabazi skull is definitely more hemione-like (very wide snout) than ass- or zebra-like (respectively narrow or long and narrow snouts). Thus, the skull morphology of *E. hydruntinus* supports the hemione connection suggested by the limb bones (model IV).

In *E. hydruntinus*, the teeth are small, the protocone (Fig. 4A) of the upper cheek teeth is short and the ectoflexid of most lower molars is deep (Fig. 4B). In hemiones, the teeth tend to be larger, the protocone is usually longer and the ectoflexid, in general, is shallow. The dimensions of the upper teeth overlap, however, and the ectoflexid pattern is not constant in either *E. hydruntinus* nor in hemiones. Figure 4 (C,D) show the protocone length relative to the mean of occlusal length and width, a measure of occlusal dimension which compensates for the effects of tooth wear (in equids, unworn upper cheek teeth tend to be long and narrow, while worn teeth tend to be short and broad). Figure 4C shows that the third and fourth upper premolars and first and second upper molars from Kabazi (including DNA sample MM, represented by a large square) plot well together with other *E. hydruntinus* teeth collected in Crimea (from the site of Chokurcha). The third and fourth upper premolars (including DNA sample CH25, represented by a large circle) and first and second upper molars from Iran are somewhat larger but have the distinctive small protocone of *E. hydruntinus*. Figure 4D shows how upper tooth dimensions (premolars and molars combined) of *E. hydruntinus*, *E. burchelli*, *E. africanus* and *E. asinus*, and *Allohippus stenonis*, overlap.

Discussion

Phylogenetic position of Equus hydruntinus

Because dental and postcranial characters give contradictory indications, several conflicting models for the phylogenetic

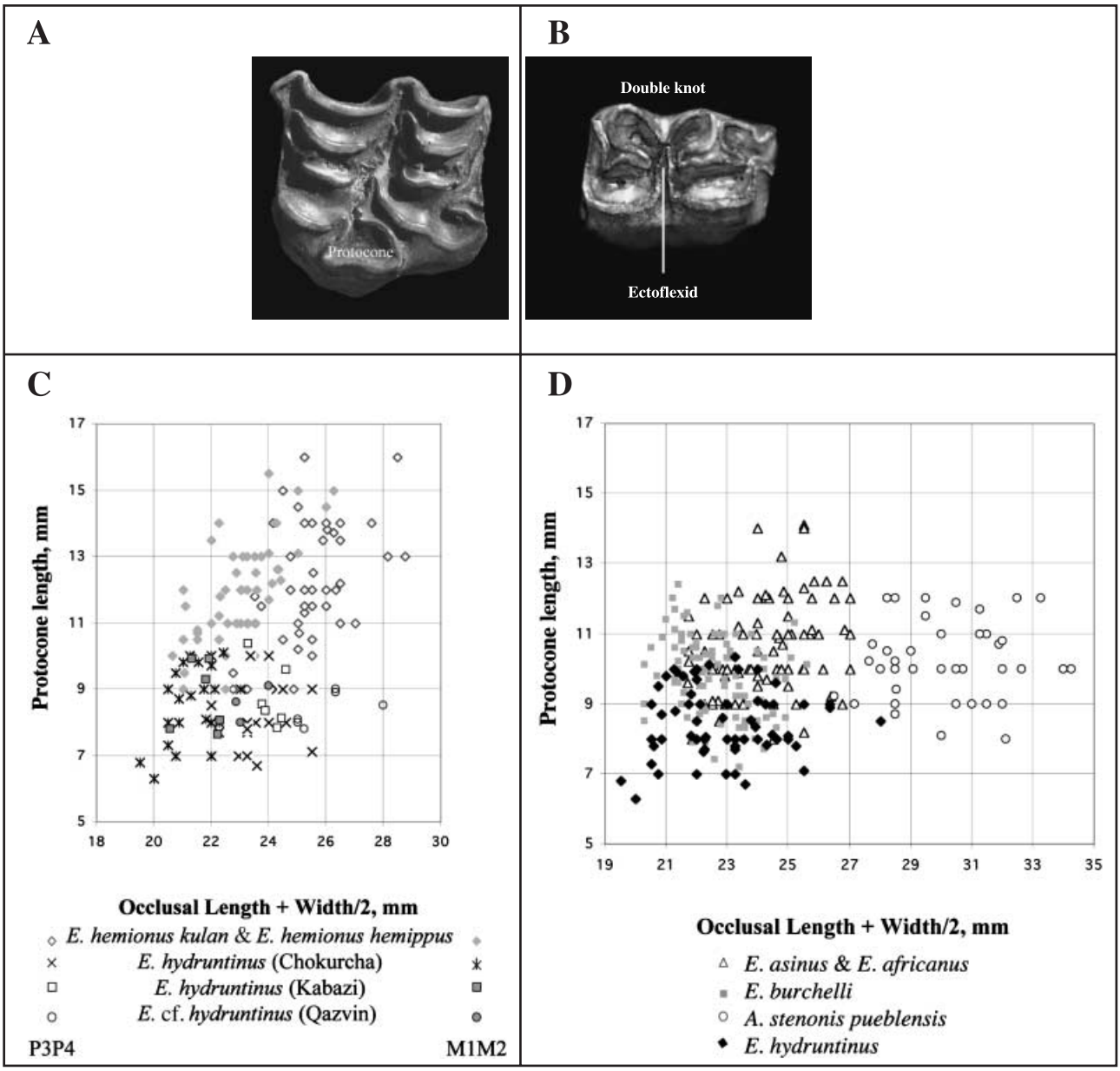

Fig. 4 Analysis of teeth morphology. (A) *Equus* upper left premolar, occlusal view showing a short protocone; (B) *Equus* lower left molar, with a rounded double knot and a deep ectoflexid; (C and D) scatter diagrams of dental measurements. P3, P4, M1, and M2 stand for upper third and fourth premolars and upper first and second molars. *n* is the number of specimens. Measurements of Kabazi and Qazvin fossils were provided by A.B. and M.M. Measurements of extant species, *Allohippus stenonis pueblensis*, and *E. hydruntinus* of Chokurcha were provided by V.E. In (C), the premolars are plotted apart from the molars. In (D), premolars and molars are lumped.

position of *Equus hydruntinus* have been proposed (Fig. 1). *E. hydruntinus* appeared to be a juxtaposition of plesiomorphies (mostly in dental characters) and of adaptations to dry and open landscapes (mostly in limb bones). A recent analysis of cranial material recovered in Crimea (Burke *et al*. 2003) suggested that *E. hydruntinus* is closely related to *Equus hemionus* in agreement with model IV. The results of the genetic analysis presented above leave no ambiguity as to the validity of model IV (Fig. 1). The *Equus* cf. *hydruntinus* and *E. hydruntinus* samples examined here are indeed closely related, exhibiting identical cyt *b* sequences and minimum genetic distance with *E. hemionus* (Figs 3 and 6); they share a deletion in the mtDNA CR which is found only in *E. hemionus* and *Equus kiang*, suggesting a synapomorphy of the *E. kiang*/*E. hemionus*/*E. hydruntinus* lineage. Finally, phylogenetic analyses cluster *E.* cf. *hydruntinus* and *E. hydruntinus* with *E. hemionus* with high bootstrap support and posterior probabilities.

Morphological evolution in E. hydruntinus

Given the agreement between the genetic data, skull and limb bone characters, we believe that dental morphology alone is of poor taxonomic value, partly because plesiomorphic. Short protocones, are considered primitive because they are present in Pliocene *Plesippus* and *Allohippus*. They are also present in *E. hydruntinus* and in some extant plains zebras and donkeys and are supposed to be poor adaptations to an abrasive food (Fig. 4A,C,D). The deep ectoflexids of the lower molars (Fig. 4B), also believed to be 'primitive' since they are found in Pliocene *Plesippus* and *Allohippus*, are present in many extant zebras and again, are considered less adaptive. These supposedly poor adaptations may, however, be compensated by exceedingly high crowns. The rounded double knot of lower cheek teeth (Fig. 4B) is also 'primitive', being the first morphology developed by equids but retained by almost all extant equids (except Caballines). In short, the choice of *Allohippus stenonis*, or *Equus burchelli*, or asses (*Equus asinus* and *Equus africanus*) as the closest relatives to *E. hydruntinus* on the basis of dental characters, although understandable, is not sound because based on plesiomorphies (Eisenmann 1992).

Geographic distribution of E. hydruntinus

Whatever its supposed relationship to other equids, *E. hydruntinus* has always been considered a 'European ass', in opposition to the African (true) asses and the Asian (half) asses ('hemione' means 'half-ass'). But what is presently known about its geographic distribution? In Europe, where equid fossils are unambiguously assigned to *E. hydruntinus* because no other modern species has similar teeth or limb bones, the presence of *E. hydruntinus* seems well founded — from the Iberian Peninsula and Sicily to Germany, where the northernmost fossils originate near the North Sea (Hooijer 1985). To the East, the limit of distribution appears to be the Volga. According to one of us (VE) fossils from Uzbekistan (Batyrov & Kuzmina 1991) and the Kuznetsk Basin (Foronova 1990) do not belong to *E. hydruntinus*. Since limb bone proportions of *E. hydruntinus* are nearly identical with those of hemiones, the assignment of fossils to *E. hydruntinus* in areas inhabited by hemiones is much more debatable as it is almost exclusively based on tooth morphology.

Until recently (Uerpmann 1987), the Middle Eastern distribution of *E. hydruntinus* was limited to Jordan, Syria, Israel, and Turkey. Since then, its distribution has been expanded to include Azerbaidjan (Eisenman & Mashkour 1999). In Africa, various zebra and ass fossils are difficult to characterize and could be confused with *E. hydruntinus*, but *E. hydruntinus* may be present in the Upper Pleistocene of Lybia (Blanc 1956) and possibly Tunisia.

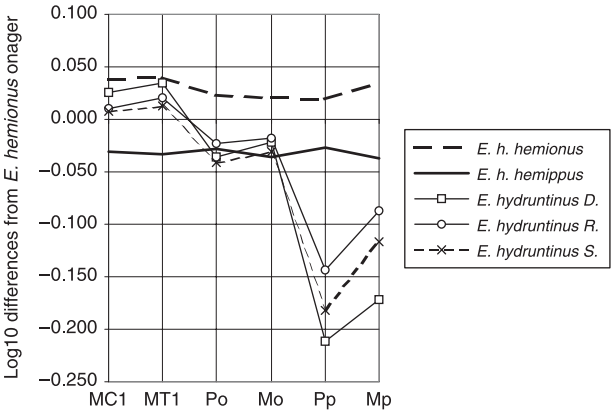

Fig. 5 Simpson's logarithmic ratio diagram comparing limb and tooth dimensions in *Equus hemionus* and *Equus hydruntinus*. MC1: length of third metacarpal; MT1: length of third metatarsal; Po: average occlusal dimension [(length + width)/2] of third and fourth upper premolars; Mo: average occlusal dimension of first and second upper molars. Pp: protocone length of third and fourth upper premolars; Mp: protocone length of first and second upper molars: *E. hydruntinus D., R.* and *S.*: *E. hydruntinus specimens from Dorog (Hungary), Roterberg (Germany) and Staroselie (Crimea)*. The reference (zero line) corresponds to the log10 of the measurements in the Iran subspecies *E. hemionus* onager. The differences of other taxa log10 measurements from the zero line are plotted above if larger, below if smaller. Since the diagram is logarithmic, isometry is represented by parallel segments between measurements. For example, the *Equus hemionus hemippus* graph is parallel, overall, to that of other subspecies of *E. hemionus* (indicating similar proportions) but below them (indicating smaller size).

The teeth belonging to the Crimean skull analysed here, and isolated teeth excavated in Iran both have typical *E. hydruntinus* features (Fig. 4C). But while the Crimea sample falls inside the classical range of *E. hydruntinus* dental dimensions and the classical range of geographic distribution, the Iranian sample does not, which is why the teeth are referred to *E.* cf. *hydruntinus* (Eisenman & Mashkour 1999). It is therefore especially interesting to find similar genetic signatures for the Crimean and Iranian samples. Our results therefore greatly extend the known distribution of *E. hydruntinus*, since we can now confirm *E. hydruntinus* from Iran, and expand its temporal distribution to the Iron Age, as suggested in Mashkour (2001).

The taxonomic status of E. hydruntinus

During the same period, other equid species inhabited the plains of Iran, giving rise to the question of how so many closely related species could coexist in the same ecological niches. The reference to modern models when dealing with past biodiversity may be the cause of our surprise. It took some time before the extraordinary faunal richness and diversity of Beringia (Guthrie 1984) was commonly accepted. The Middle East, where *E. hydruntinus* was apparently

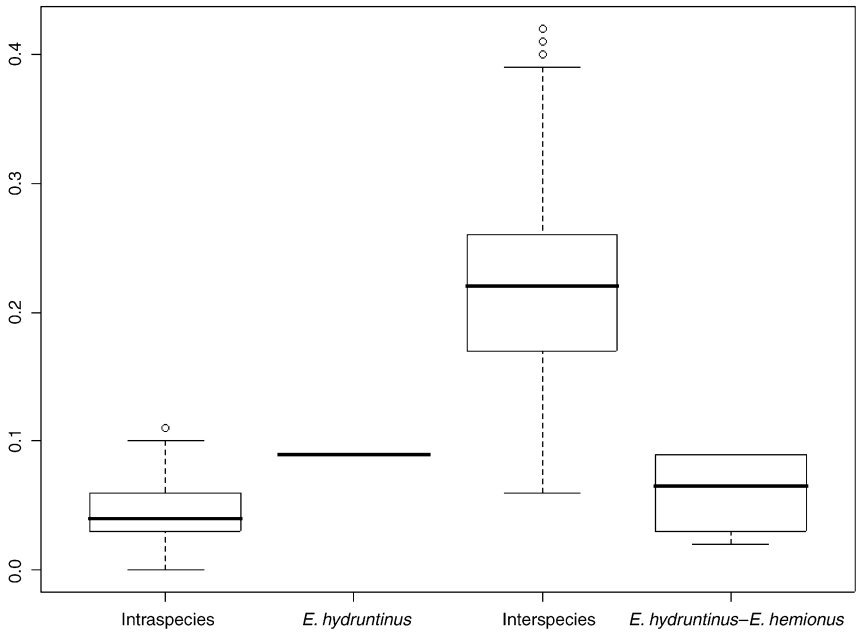

Fig. 6 Intra- and interspecific pairwise distances among equids. Pairwise distances were computed using both the mtDNA control region data set. The estimation for intraspecific distances were estimated for the species which were the most represented in our data sets, i.e. *Equus caballus* and *Equus burchelli*.

sympatric with *E. hemionus* and *Equus caballus* and a further, very large *Equus* was just recently discovered (Eisenmann *et al*. 2002), could possibly have also supported more species than currently believed.

In equids, osteological characters diagnostic of a species are extremely rare and overlap in individual dimensions almost always occurs (Fig. 4C,D). It is the combination of several osteological and dental traits, which is the best way to define fossil species (Eisenmann 1979, 1998; Groves & Willoughby 1981). In the case of *E. hydruntinus* and *E. hemionus*, the relative average dimensions of the metapodials and upper cheek teeth parameters of *E. hydruntinus* differs from that of any other hemione (the metapodials are as long as the Kulan, the teeth are as small as the Hemippes and the protocones even smaller). Accordingly, pending new insights, we consider them as distinct species. The corresponding Simpson's 'ratio diagram' (Fig. 5; Burke *et al*. 2003) illustrates these differences.

Alternatively, it still remains the possibility that *E. hydruntinus* was only conspecific with *E. hemionus* and not a true species. Part of our data may support such a statement: (i) the genetic distance between *E. hemionus* (CH31) and *E. hydruntinus* (CH25 or MM) is in the range of the intraspecific variation observed in *E. caballus* or *E. burchelli* (Fig. 6). (ii) In our phylogenetic tree, *E. hydruntinus* appears as paraphyletic (Fig. 3). (iii) The remarkably short snout and narial opening, the slenderness of its limbs, the short ectoflexids typical of the lower molars, and the small protocones typical of the upper teeth observed in *E. hydruntinus* are not enough to define it as a species since they may at times be observed in *E. hemionus*. Indeed, in equids, osteological characters diagnostic of a species are extremely rare and overlap in individual dimensions almost always occurs (Fig. 4C,D).

Therefore, we consider the question of the real taxonomic status of *E. hydruntinus* still open. More sequence data are needed to conclusively assess *E. hydruntinus* monophyly, to survey its genetic diversity and to estimate more accurately the genetic distance that separate it from *E. hemionus*. Recently, taxonomic and phylogenetic problems among other extinct equids (amerindian *Hippidion* sp.) have been revealed by ancient DNA (Orlando *et al*. 2003; discussed in Alberdi *et al*. 2005; Weinstock *et al*. 2005). Added to the results reported here, these studies suggest the taxonomy of recently extinct equids should be scrupulously revised.

Acknowledgements

We thank A. Thénot for technical help, L. Granger and C. Lerondel in sequencing at the DTAMB (UCB Lyon 1) and the following organizations for supporting this work: CNRS, MENRT (ACI) and UCB-Lyon1. We are especially indebted with Sandrine Hughes, Vincent Laudet, Jean-Denis Vigne, two referees for critical reading of the manuscript and Tristan Lefébure who kindly provided unpublished Perl and R scripts. We would also like to thank Dr Victor Chabai (Ukrainian Academy of Sciences) with whose permission the Kabazi skull could be included in this study.

References

Alberdi MT, Prado JL, Prieto A (2005) Considerations on the paper 'Morphological convergence in *Hippidion* and *Equus* (*Amerhippus*) South American equids elucidated by ancient DNA analysis', by Ludovic Orlando, Véra Eisenmann, Frédéric Reynier, Paul Sondaar, Catherine Hänni. *Journal of Molecular Evolution*, **61**, 145–147.

Antunes MT (2006) The *Zebro* (Equidae) and its extinction in Portugal, with an Appendix on the noun *zebro* and the modern zebra. In: *Equids in Time and Space* (ed. Mashkour M). Oxbow Books, Oxford.

Aranguren-Mendez J, Beja-Pereira A, Avellanet R, Dzama K, Jordana J (2004) Mitochondrial DNA variation and genetic relationships in spanish donkey breeds (*Equus asinus*). *Journal of Animal Breeding and Genetics*, **121**, 319–330.

Aranguren-Mendez J, Jordana J, Gomez M (2001) Genetic diversity in Spanish donkey breeds using microsatellite DNA markers. *Genetics, Selection, Evolution*, **33**, 433–442.

Batyrov B, Kuzmina IE (1991) Pleistotsenovii osel *Equus hydruntinus* Regalia v Evrasii. *Trudy Zoologicheskogo Instituta A.N. SSSR*, **238**, 121–138.

Blanc GA (1956) Sulla esistenza di '*Equus (Asinus) hydruntinus*' Regalia nel Pleistocene del Nord Africa. *Bolletino della Societa Geologica Italiana*, **LXXV**, 3–14.

Bocherens H, Mashkour M, Billiou D (2000) Palaeoenvironmental and archaeological implications of isotopic analyses (^{13}C, ^{15}N) from neolithic to present in Qazvin plain (Iran). *Environmental Archeology*, **5**, 1–19.

Bonifay MF (1991) *Equus hydruntinus Regalia* minor n.ssp. from the caves of Lunel-Viel (Hérault, France). *Beihefte zum Tübinger Atlas des Vorderen Orients, Reihe A*, **19**, 178–216.

Burke A, Markova AK, Mikhailesku C, Patou-Mathis M (1999) The animal environment of Western Crimea. In: *The Middle Palaeolithic of Western Crimea*, Vol. 2 (eds Chabai V, Monigal K), pp. 143–151. ERAUL, Liège.

Burke A, Eisenmann V, Ambler G (2003) The systematic position of *Equus hydruntinus*, an extinct species of Pleistocene equid. *Quaternary Research*, **59**, 459–469.

Chikuni K, Mori Y, Tabata T, Saito M, Monma M, Kosygiyama M (1995) Molecular phylogeny based on the kappa-casein and cytochrome *b* sequences in the mammalian suborder Ruminantia. *Journal of Molecular Evolution*, **41**, 859–866.

Davis SJ (1980) Late Pleistocene and Holocene equid remains from Israel. *Zoological Journal of the Linnean Society*, **70**, 289–312.

Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJ (2003) Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Molecular Biology and Evolution*, **20**, 248–254.

Eisenmann V (1979) Caractères évolutifs et phylogénie du genre *Equus* (Mammalia; Perissodactyla). *Comptes Rendus de l'Académie des Sciences D*, **288**, 497–500.

Eisenmann V (1992) Systematic and biostratigraphical interpretation of the equids from Qafzeh, Tabun, Shkul, and Kebara (Acheulo-Yabrudian) to Upper Palaeolithic of Israel. *Archaeozoologia*, **V**, 43–62.

Eisenmann V (1998) Folivores et tondeurs d'herbe: forme de la symphyse mandibulaire des Equidés et des Tapiridés (Perissodactyla, Mammalia). *Geobios*, **31**, 113–123.

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

Eisenmann V, Mashkour M (1999) The small equids of Binagady (Azerbaidjan) and Qazvin (Iran): *E. hemionus binagadensis* nov. subsp. and *E. hydruntinus*. *Geobios*, **32**, 105–122.

Eisenmann V, Helmer D, Sadia Segui M (2002) The big *Equus* from the Geometric Kebaran of Umm El Tlel, Syria: *Equus valeriani*, *Equus capensis, or Equus caballus*? In: *Proceedings of the Fifth International Symposium on the Archaeozoology of Southwestern Asia and Adjacent Areas* (eds Buitenhuis H, Choyke AM, Mashkour M, Al-Shyiab AH), Archaeozoology of the New East, V. pp. 62–73. Groningen.

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

Foronova IV (1990) *Iskopaemae loshadi Kuznetsovskoï Kotloviny*. AN SSSR, Sibirskoie Otdelenie, Institut geologii i geofiziki, Novossibirsk.

Forsten A (1986) A review of the Süssenborn horses and the origin of *Equus hydruntinus* Regalia. *Quartärpaläontologie*, **6**, 43–52.

Forsten A, Ziegler R (1995) The horses (Mammalia, Equidae) from the early Wuermian of Villa Seckendorff, Stuttgart-Bad Cannstadt, Germany. *Stuttgarter Beiträge zur Naturkunde, Series B (Geologie und Paläontologie)*, **224**, 1–22.

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Computer Applications in the Biosciences*, **12**, 543–548.

Gilbert MTP, Rudbeck L, Willerslev E *et al.* (2005) Biochemical and physical correlates of DNA contamination in archaeological human bones and teeth excavated at Matera, Italy. *Journal of Archaeological Science*, **32**, 785–793.

Gould SJ (1994) *Hen's Teeth and Horse's Toe. Further Reflections in Natural History*. Norton, New York.

Gromova VI (1949) Istorija loshadej (roda *Equus*) v Starom Svete. Chast' 2. Evoljutsija i klassifikatsija roda. *Trudy Palaeontologicheskogo Instituta Akademii Nauk SSSR*, **17**, 162.

Groves CP, Willoughby DP (1981) Studies on the taxonomy and phylogeny of the genus *Equus*. *Mammalia*, **45**, 321–354.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Guthrie D (1984) Mosaics, allelochemics, and nutrients: an ecological theory of Late Pleistocene. In: *Quaternary Extinctions* (eds Martin PS, Klein RG), pp. 259–298. University of Arizona Press, Tucson, Arizona.

Hänni C, Laudet V, Stéhélin D, Taberlet P (1994) Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. *Proceedings of the National Academy of Sciences, USA*, **91**, 12336–12340.

Higuchi RG, Bowman B, Freiberger M, Ryder OA, Wilson AC (1984) DNA sequences from the quagga, an extinct member of the horse family. *Nature*, **312**, 282–284.

Higuchi RG, Wrischnik LA, Oakes E, George M, Tong B, Wilson AC (1987) Mitochondrial DNA of the extinct quagga: relatedness and extent of postmortem change. *Journal of Molecular Evolution*, **25**, 283–287.

Hooijer DA (1985) A further note on the fossil Anglo-Dutch Ass from the North Sea. *Lutra*, **28**, 26–30.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Huxley TH (1870) Address delivered at the anniversary meeting of the Geological Society of London, on the 18th February. 16–38. Reported in: Gould SJ (1987) Life's little joke. *Natural History*, **96**, 16–25.

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution*, **32**, 128–144.

Ishida N, Hasegawa T, Takeda K *et al.* (1994) Polymorphic sequence in the D-loop region of equine mitochondrial DNA. *Animal Genetics*, **25**, 215–221.

Ishida N, Oyunsuren T, Mashima S, Mukoyama H, Saitou N (1995) Mitochondrial DNA sequences of various species of the genus *Equus* with special reference to the phylogenetic relationship

between Przewalskii's wild horse and domestic horse. *Journal of Molecular Evolution*, **41**, 180–188.

Ishida N, Hasegawa T, Oyunsuren T, Mukoyama H (1996) PCR-RFLP analysis of the cytochrome b gene in horse mitochondrial DNA. *Animal Genetics*, **27**, 359–363.

Jung YH, Han SH, Shin T, Oh MY (2002) Genetic characterization of horse bone excavated from the Kwakji archaeological site, Jeju, Korea. *Molecular Cell*, **14**, 224–230.

Kavar T, Habe F, Brem G, Dovc P (1999) Mitochondrial D-loop sequence variation among the 16 maternal lines of the Lipizzan horse breed. *Animal Genetics*, **30**, 423–430.

Kim KI, Yang YH, Lee SS *et al.* (1999) Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. *Animal Genetics*, **30**, 102–108.

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *Journal of Molecular Evolution*, **29**, 170–179.

Lister AM, Kadwell M, Kaagan LM, Jordan WC, Richards MB, Stanley HF (1998) Ancient and modern DNA in a study of horse domestication. *Ancient Biomolecules*, **2**, 267–280.

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. *Current Biology*, **11**, 200–203.

Mashkour M, Fontugne M, Hatte C (1999) Investigations on the evolution of subsistence economy in the Qazvin plain (Iran) from the Neolithic to the Iron Age. *Antiquity*, **73**, 65–76.

Mashkour M (2001) Chasse et élevage du Néolithique à l'Âge du Fer dans la plaine de Qazvin (Iran). Étude archéozoologique des sites de Zagheh, Qabrestan et Sagzabad, 738pp. Ph Thesis: Université Paris I, Pantheon-Sorbonne, Paris.

Oakenfull EA, Ryder OA (1998) Mitochondrial control region and 12S rRNA variation in Przewalski's horse (*Equus przewalskii*). *Animal Genetics*, **29**, 456–459.

Oakenfull EA, Lim HN, Ryder OA (2000) A survey of equid mitochondrial DNA: implications for the evolution, genetic diversity and conservation of *Equus*. *Conservation Genetics*, **1**, 341–355.

Orlando L, Bonjean D, Bocherens H *et al.* (2002) Ancient DNA and the population genetics of cave bears (*Ursus spelaeus*) through space and time. *Molecular Biology and Evolution*, **19**, 1920–1933.

Orlando L, Eisenmann V, Reynier F, Sondaar P, Hänni C (2003) Morphological convergence in *Hippidion* and *Equus* (*Amerhippus*) southern-american equids elucidated by ancient DNA analysis. *Journal of Molecular Evolution*, **57** (Suppl. 1), S29–S40.

Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**, 289–290.

Philippe H, Sorhannus U, Baroin A, Perasso R, Gasse F, Adoutte A (1994) Comparison of molecular and paleontological data in diatoms suggests a major gap in the fossil record. *Journal of Evolutionary Biology*, **7**, 247–265.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Radulesco C, Samson P (1962) Remarques sur quelques Equidés quaternaires de Roumanie. *Zoologischer Anzeiger*, **168**, 171–179.

R Development Core Team (2005) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Shimodaira H (2002) An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, **51**, 492–508.

Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, **17**, 1246–1247.

Stehlin HG, Graziosi P (1935) Ricerche sugli Asinidi fossili d'Europa. *Memoires de la Societe Paléontologique Suisse*, **56**, 1–73.

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. *Molecular Phylogenetics Evolution*, **19**, 34–44.

Uerpmann HP (1987) The ancient distribution of ungulate mammals in the Middle East. Fauna and archaeological Sites in Southwest Asia and Northeast Africa. *Beihefte Zum Tübinger Atlas Des Vorderen Orients, Reihe A*, **27**, 172pp.

Vila C, Leonard JA, Gotherstrom A *et al.* (2001) Widespread origins of domestic horse lineages. *Science*, **291**, 474–477.

Weinstock J, Willerslev E, Sher A *et al.* (2005) Evolution, systematics, and phylogeography of pleistocene horses in the New World: a molecular perspective. *PLoS Biology*, **3**, e241.

Wilms C (1989) Zum Aussterben des europäischen Wildesels. *Germania*, **67**, 143–148.

Xu X, Arnason U (1994) The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. *Gene*, **148**, 357–362.

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*. *Molecular Phylogenetics Evolution*, **7**, 189–194.

Xu X, Gullberg A, Arnason U (1996) The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA comparisons among four closely related mammalian species-pairs. *Journal of Molecular Evolution*, **43**, 438–446.

Drs. Orlando and Hänni are mostly interested in using ancient DNA to reconstruct the phylogenetic origin and the biogeography evolution of extinct mammals. Dr. Eisenmann is a specialist of extant and fossil equids. Dr. Mashkour is an archaeozoologist and takes part to several ongoing archeaological excavations in Iran to study the early events of domestication. Dr. Burke is also a zooarchaeologist and professor of anthropology at the Université de Montreal, currently interested in paleoethology and landscape studies. Dr Douady is interested in phylogenetic and tokogenetic reconstructions from both methodological and biological point of views.
