Resolving a zoological mystery: the kouprey is a real species

Alexandre Hassanin¹,²,* and Anne Ropiquet¹,²,³

¹UMR 5202, Origine, Structure et Evolution de la Biodiversité, Département Systématique et Evolution, Muséum National d’Histoire Naturelle, Case postale No 51, 55, rue Buffon, 75005 Paris, France
²Service de Systématique Moléculaire, Muséum National d’Histoire Naturelle, 43, rue Cuvier, 75005 Paris, France
³Department of Botany and Zoology, Evolutionary Genomics Group, University of Stellenbosch, Private Bag X1, Matieland 7602, Republic of South Africa

The kouprey was described as a new species, Bos sauveli on the basis of a calf captured in Preah Vihear province of Cambodia and kept alive at the Vincennes Zoo near Paris during the last five decades due to multiple possible factors including uncontrolled hunting, deforestation and competition with domestic livestock linked to human demographic growth, and wars. No sightings of kouprey have been reported by scientists since the 1980s, suggesting that the species is now extinct (MacKinnon & Stuart 1988). The animal is now the national emblem of Cambodia, and an icon of wildlife conservation in southeast Asia.

1. INTRODUCTION

The kouprey was described as a new species, Bos sauveli on the basis of a calf captured in Preah Vihear province of Cambodia and kept alive at the Vincennes Zoo near Paris until 1940 (Urbain 1937). In the middle of the twentieth century, its range was already limited to northern provinces of Cambodia, and slightly beyond the borders with Thailand, Laos and Vietnam (Sauvel 1949). Populations declined dramatically during the last five decades due to multiple possible factors including uncontrolled hunting, deforestation and competition with domestic livestock linked to human demographic growth, and wars. No sightings of kouprey have been reported by scientists since the 1980s, suggesting that the species is now extinct (MacKinnon & Stuart 1988). The animal is now the national emblem of Cambodia, and an icon of wildlife conservation in southeast Asia.

The kouprey is a mysterious animal with striking characters, including spectacular curving horns and a pronounced dewlap, a pendulous skin at the base of the neck that can nearly touch the ground in some older males. Numerous morphological hypotheses have been proposed for the origin of the kouprey: that it is a species closely related to banteng (Bos javanicus) and gaur (Bos frontalis), two other wild oxen of southeast Asia (Urbain 1937; Bohlken 1961; Pfeffer & Kim-San 1967); a morphologically divergent species placed in a separate genus, named Novibus (Goolidge 1940); a wild species linked to aurochs and domestic cattle (Pfeffer & Kim-San 1967; Groves 1981); a vicariant population of banteng (Corbet & Hill 1992); a feral cattle (Wharton 1957; Bohlken 1963); or a hybrid of banteng with either zebu cattle, gaur or water buffalo (Cheminaud 1939; Edmond-Blanc 1947; Bohlken 1958).

In 2004, the holotype of the kouprey (no. 1940–51, MNHN) was included in a molecular phylogeny of the tribe Bovini, and the results suggested close affinities with banteng and gaur (Hassanin & Ropiquet 2004). Seven nucleotide signatures were detected in the mitochondrial cytochrome b gene (Cytb) of the holotype. Surprisingly, four of these signatures were rediscovered in the sequences of Cambodian banteng, and Galbreath et al. (2006) concluded that the kouprey was not a valid species, but a feral hybrid resulting from a crossing between domestic zebu and wild banteng. Unfortunately, this conclusion gained a lot of media coverage (e.g. Bakalar 2006; Casey 2006; Derr 2006). Interviewed for CBS News, Galbreath said 'It is surely desirable not to waste time and money trying to locate or conserve a domestic breed gone wild. The limited funds available for conservation should be used to protect wild species' (Casey 2006).

We suggested, however, that Galbreath et al. (2006) misinterpreted the DNA data (Hassanin & Ropiquet 2007; see also Grigson 2007; Hedges et al. 2007). In the mitochondrial tree, here constructed with three different markers (Cytb, CO2 and D-loop; figure 1), Cambodian banteng are indeed found to be closely related to the kouprey (mean distance: 1.4%) and more distant to gaur (5.0%), but they are unexpectedly found to be highly divergent from Javan banteng (5.4%). Particularly

* Author for correspondence (hassanin@mnhn.fr).
The presence of a large insertion in the mitochondrial D-loop of Javan banteng (176 nt), which is not found in Cambodian banteng. Two conflicting hypotheses can therefore be proposed to interpret the mitochondrial data (figure 2). The first hypothesis assumes that the Cambodian and Javan banteng belong to two distinct species, and that the kouprey diverged morphologically from the former owing to hybridization with another species needing to be identified. This hypothesis is compatible with the conclusions of Galbreath et al. (2006) if we accept that the hybridization of banteng occurred with zebu. The second hypothesis recognizes the kouprey as a valid species, and implies the existence of a mitochondrial introgression event, in which the mitochondrial genome of kouprey was transferred into the ancestor of Cambodian banteng by natural hybridization. Both hypotheses are supported by the fact that viable and potentially fertile hybrids have been produced in captivity between various species of the genus Bos (Van Gelder 1977).

The aim of the present study was to reject one of the two hypotheses in order to conclusively define the taxonomic status of the kouprey. The fact that the mitochondrial genome is maternally inherited limits its application to the evolutionary study of maternal lineage. For this reason, we also sequenced five non-coding nuclear fragments for the holotype of the kouprey and all living species of oxen, bison and yak: two fragments of the Y-chromosome were used to trace the evolutionary history of paternal lineage, and three independent autosomal genes were analysed to evidence possible cases of inter-specific hybridization.

2. MATERIAL AND METHODS
(a) Taxonomic sample
All seven species of the subtribe Bovina (Hassanin & Ropiquet 2004; Wilson & Reeder 2005) are represented in this study (table 1): (i) Bos sauveli, with the holotype of the kouprey (No. 1940–51, MNHN), (ii) Bos javanicus, with four Cambodian banteng and two Javan banteng, (iii) Bos taurus, with two humpless domestic cattle (subspecies B. t. taurus) and two southeast Asian zebu (subspecies B. t. indicus), (iv) Bos frontalis, with two different populations of gaur, (v) Bos grunniens (yak), (vi) Bison bison (American bison), and (vii) Bison bonasus (European bison).
(b) Molecular markers

The nuclear genes were chosen, firstly, because our preliminary analyses revealed that banteng and zebu differ at several nucleotide sites, and, secondly, because they are unlinked markers, with different locations in the genome of *B. taurus*: chromosome 11 for intron 1 of the beta-spectrin non-erythrocytic 1 gene (*SPTBN1*); chromosome 14 for intronic and exonic regions of the thyglobulin gene (*TG*); chromosome 17 for intron 7 of the beta-fibrinogen gene (*FGB*); and the Y-chromosome for two non-coding fragments of the sex-determining region Y (*SRY*). BLAST searches performed on the assembled genome of *B. taurus* indicate that all four nuclear genes are present in single copy, therefore, avoiding PCR amplification of paralogous sequences and facilitating the phylogenetic interpretations.

Three mitochondrial regions were also analysed: the 5′ part of the control region (also named *D-loop*), and two protein-coding genes, i.e., the complete cytochrome *b* (*Cytb*) and subunit II of the cytochrome *c* oxidase (*CO2*).

(c) DNA extraction, amplification and sequencing

Total DNA was extracted from fresh tissues (blood, muscle or hair) or from bones of specimens conserved in the MNHN collections as detailed in Hassanin & Ropiquet (2004). The standard PCR conditions were as follows: 3 min at 94°C; 30–40 cycles of denaturation/annealing/extension with 45 s at 94°C for denaturation, 45 s at 50–60°C for annealing and 1 min at 72°C for extension; and 7 min at 72°C. For DNA extracted from museum specimens, several sets of primers were designed for amplifying and sequencing overlapping PCR products. For DNA extracted from fresh tissues, PCR amplifications were done using external primers only. The protein-coding mitochondrial genes (*Cytb* and *CO2*) were amplified using published primers (Hassanin & Ropiquet 2004). The 5′ part of the *D-loop* region was obtained using the following two primer pairs: (i) 5′-ACT-ATG-ACC-AGC-GGG-CAC-3′ (F: forward) and 5′-GAG-TAC-AAA-GTC-TGT-GTG-3′ (R: reverse) and (ii) 5′-TAG-ACC-AAA-CGC-GAG-3′ (F) and 5′-GCT-GTG-TGG-TTC-AAG-CGG-CAC-3′ (R).

Intron 1 of *SPTBN1* was amplified using the three following primer pairs: (i) 5′-AGT-GCA-GCC-TTG-AAA-GGT-AC-3′ (F) and 5′-CAA-AGT-TCA-CTG-CCC-AGA-AGC-3′ (R), (ii) 5′-CCC-TTC-ATG-ACC-CAA-GTG-CTAC-3′ (F) and 5′-CAA-AGT-TAG-AAA-TAT-CTG-CAC-3′ (R), and (iii) 5′-GCT-CTC-TTG-GCT-TTC-ACT-GGT-GTG-3′ (F) and 5′-ACA-CCC-CTG-TTC-GTC-CTA-G3′ (R). The *TG* fragment was amplified using the following three primer pairs: (i) 5′-GAG-CCC-AAG-AAA-TGT-GAG-TG-3′ (F) and 5′-AGG-CCTG-CGG-CTC-ACT-AAA-TG-3′ (R), (ii) 5′-GAC-AGC-AGC-TGG-TGT-GTC-AGG-3′ (F) and 5′-GAC-CAA-GAT-GCA-TAT-GTG-GTA-AGG-3′ (R), and (iii) 5′-CCC-CCT-GAG-TCC-ATG-GAG-3′ (F) and 5′-GTG-GTG-GGT-GGA-AGC-CTG-GTC-3′ (R). Intron 7 of *FGB* was amplified with the following four primer pairs: (i) 5′-CCA-CAA-CRG-CAT-GTG-CCT-3′ (F) and 5′-AGA-GCT-TAG-ATG-GCT-GCC-3′ (R), (ii) 5′-CAT-GAC-AGC-ATA-GAA-GAT-GTG-CAG-3′ (F) and 5′-GCT-CTA-TTA-TTT-GAG-CAG-ATC-3′ (R), (iii) 5′-GCA-TTG-GAG-CAG-CAC-3′ (F) and 5′-AGG-TGG-AGG-TGG-AGG-3′ (R), and (iv) 5′-GAA-TAT-TTG-TTG-GTA-ATT-TAG-CAC-ATG-3′ (F) and 5′-CAA-GTT-ATGT-TTC-AGG-CAC-3′ (R). The first non-coding fragment of *SRY*, located upstream of the 5′ end of the coding sequence (*SRY*-5′), was amplified with 5′-CCT-GTT-AGG-TAG-TCT-CTG-TTG-AGG-3′ (F) and 5′-CA-CAG-GTC-AGT-AGC-ATC-G3′ (R). The second non-coding fragment of *SRY*, located downstream of the 3′ end of the coding sequence *SRY*-3′, was amplified with 5′-CCT-GTT-AGG-TAG-TCT-CTG-TTG-AGG-3′ (F) and 5′-CA-CAG-GTC-AGT-AGC-ATC-G3′ (R).
Table 1. Origin of the DNA sequences.

<table>
<thead>
<tr>
<th>taxa</th>
<th>N°/origin</th>
<th>sex</th>
<th>Cytb 1140 nt</th>
<th>CO2 582 nt</th>
<th>D-loop 710 nt</th>
<th>FGB 554 nt</th>
<th>SPTBN1 442 nt</th>
<th>TG 627 nt</th>
<th>SRY-5' 224 nt</th>
<th>SRY-3' 303 nt</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bos sauveli</em> kouprey</td>
<td>1940–51 (holotype) MNHN</td>
<td>M</td>
<td>AY689189a</td>
<td>EF693812b</td>
<td>EF693828b</td>
<td>EF693843b</td>
<td>EF693858b</td>
<td>EF693871b</td>
<td>EF693888b</td>
<td></td>
</tr>
<tr>
<td><em>Bos javanicus</em> Cambodian banteng</td>
<td>CKM4 Teuk Chhou zoo</td>
<td>M</td>
<td>EF685912c</td>
<td>EF693806b</td>
<td>EF693822b</td>
<td>EF693837b</td>
<td>EF693852b</td>
<td>EF693867b</td>
<td>EF693878b</td>
<td></td>
</tr>
<tr>
<td><em>Bos javanicus</em> Javan banteng</td>
<td>CKM24 Phnom Tamao zoo</td>
<td>M</td>
<td>EF685914c</td>
<td>EF693808b</td>
<td>EF693824b</td>
<td>EF693838b</td>
<td>EF693853b</td>
<td>EF693868b</td>
<td>EF693879b</td>
<td></td>
</tr>
<tr>
<td><em>Bos javanicus</em> Javan banteng</td>
<td>CKM23 Phnom Tamao Zoo</td>
<td>M</td>
<td>EF693796b</td>
<td>EF693804b</td>
<td>EF693825b</td>
<td>EF693840b</td>
<td>EF693855b</td>
<td>EF693867b</td>
<td>EF693878b</td>
<td></td>
</tr>
<tr>
<td><em>Bos javanicus</em> Javan banteng</td>
<td>CERZA Hermival-les-Vaux</td>
<td>M</td>
<td>EF689188a</td>
<td>EF693809b</td>
<td>EF693827b</td>
<td>EF693842b</td>
<td>EF693852b</td>
<td>EF693870b</td>
<td>EF693881b</td>
<td></td>
</tr>
<tr>
<td><em>Bos taurus taurus</em> domestic cattle</td>
<td>MNHN NCBI database</td>
<td>F</td>
<td>EF693797b</td>
<td>EF693801b</td>
<td>EF693826b</td>
<td>EF693841b</td>
<td>EF693856b</td>
<td>AB039748f</td>
<td>AB039748f</td>
<td></td>
</tr>
<tr>
<td><em>Bos taurus indicus</em> zebu</td>
<td>Limousin France BNX Vietnam</td>
<td>M</td>
<td>EF693799b</td>
<td>EF693802b</td>
<td>EF693829b</td>
<td>EF693844b</td>
<td>EF693858b</td>
<td>EF693872b</td>
<td>EF693882b</td>
<td></td>
</tr>
<tr>
<td><em>Bos frontalis</em> gaur</td>
<td>CMK8 Cambodia Cardamom Cambodia</td>
<td>M</td>
<td>EF685907c</td>
<td>EF693815b</td>
<td>EF693831b</td>
<td>EF693846b</td>
<td>EF693861b</td>
<td>EF693874b</td>
<td>EF693884b</td>
<td></td>
</tr>
<tr>
<td><em>Bos grunniens</em> yak</td>
<td>Jardin des Plantes MNHN</td>
<td>M</td>
<td>NC_006380g</td>
<td>NC_006380g</td>
<td>EF693834b</td>
<td>EF693849b</td>
<td>EF693866b</td>
<td>EF693876b</td>
<td>EF693886b</td>
<td></td>
</tr>
<tr>
<td><em>Bison bison</em> American bison</td>
<td>Vincennes Zoo MNHN</td>
<td>M</td>
<td>AF036273b</td>
<td>AY689197a</td>
<td>EF693810b</td>
<td>EF693835b</td>
<td>EF693850b</td>
<td>EF693863b</td>
<td>EF693877b</td>
<td>EF693887b</td>
</tr>
<tr>
<td><em>Bison bonasus</em> European bison</td>
<td>OSEB MNHN</td>
<td>F</td>
<td>AY689186a</td>
<td>AY689198a</td>
<td>EF693811b</td>
<td>EF693836b</td>
<td>EF693851b</td>
<td>EF693864b</td>
<td>AOY79142j</td>
<td>AOY79142j</td>
</tr>
</tbody>
</table>

sequence (SRY-3'), was amplified with the following two sets of primers: (i) 5'‐CAT‐GTA‐AAG‐AAT‐TCA‐GAC‐TTC‐CC‐3' (F) and 5'‐CCA‐TCT‐AAC‐TG‐GCA‐ATC‐TGC‐3' (R) and (ii) 5'‐CTG‐CTG‐TGA‐TTC‐TCC‐AAG‐GAT‐CAT‐3' (F) and 5'‐AGG‐GAG‐CTT‐TCC‐ATC‐CAA‐GTA‐C‐3' (R). Both strands of all PCR products were sequenced by Genoscreen (Lille, France) with the BigDye Terminator v. 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Sequences generated for this study are available from the GenBank/EMBL/DDBJ databases under accession numbers (Hassanin & Ropiquet 2004).

The expected number a priori of time units between tip and root (rttm) was set at 14 Myr ago, with a standard deviation of 7 Myr ago. The Markov chains were sampled 10 000 times every 100 generations and the 'burn-in' period was set at 100 000 generations. Three calibration points were used for the analyses: the first two correspond to independent domestications of humpless cattle and zebu cattle between 8000 and 10 000 years BP (Loftus et al. 1994), and the third refers to the diversification of the subtribe Bovina (Bos and Bison), estimated between 3.89 and 5.53 Myr ago (Hassanin & Ropiquet 2004).

3. RESULTS AND DISCUSSION

(a) Analyses of nuclear sequences

Five non-coding nuclear fragments, including introns of three independent autosomal genes (FGB, SPTBN1 and TG) and two regions of the SRY gene in the Y-chromosome were sequenced for the holotype of the kouprey and all living species of oxen, bison and yak. What were the expected results with nuclear data (figure 2)? If the holotype of the kouprey was an F1 hybrid between Cambodian banteng (Bc) and zebu (Z), its autosomal genes would be identical to that of zebu, while its Y-chromosome would be identical to that of zebu, while its autosomal genes would be heterozygous Z/Bc, and therefore characterized by several dimorphic nucleotide sites. If the holotype of the kouprey were a member of an outbreeding population, its autosomal loci would be either heterozygous Z/Bc or homozygous (Bc/Bc, or ZZ), and its Y-chromosome would be identical to that of zebu or banteng. As explained previously, the hypothesis of a hybrid origin for the kouprey argues that the Cambodian and Javan banteng belong to two different species. As a consequence, banteng from Cambodia and Java are expected to have different nuclear sequences (Bc and BJ), while the kouprey would differ, at least at some loci, from all other species of the genus Bos.

Sixty-one nucleotide sites were found to be variable between nuclear sequences of Bos and Bison species (figure 3). The results provide conclusive proof for the existence of B. sauveli. Firstly, Cambodian and Javan

### Figure 3. List of the 61 variable sites between nuclear sequences of Bos and Bison species.

The five non-coding nuclear fragments include two regions of the SRY gene in the Y-chromosome (SRY-5' and SRY-3') and introns of three independent autosomal genes (FGB, SPTBN1 and TG). Heterozygous nucleotide sites are highlighted in yellow. Cambodian B. javanicus sequences are labelled in blue and Javan sequences in red.
banteng share identical nuclear alleles, indicating that both populations belong to the same species, *B. javanicus*. Secondly, the sequences of the kouprey holotype do not contain heterozygous sites, which would be expected in the case of hybrid origin, and they differ from those found in other species of *Bos* (with the exception of the *FGβ* gene, for which kouprey, banteng and gaur share the same allele). Thirdly, three nuclear sites are diagnostic for the holotype of the kouprey (figure 3): A in position 92 of *SRY*-3; T in position 51 of *SRY*-2; and G in position 262 of *TG*. The analysis of nuclear data demonstrates therefore that *B. sauveli* and *B. javanicus* are two distinct and valid species, and that the Cambodian banteng acquired a mitochondrial genome from the kouprey by introgressive hybridization.

(b) *Introgression of the mitochondrial genome of kouprey into the common ancestor of Cambodian banteng*

One fundamental question for the conservation of wild populations was to determine whether hybridization between kouprey and banteng occurred as a consequence of human activities. Using a relaxed molecular clock, we estimated that the hybridization occurred during the Pleistocene epoch, at 1.34 ± 0.45 Myr ago. As this estimate largely predates the origin of agriculture and domestication of plants and animals, it can be concluded that it was not a consequence of human intervention. The mitochondrial introgression supposes that at least one kouprey female, which was probably young in order to overcome inter-specific ethological barriers (Kendrick et al. 1998), was adopted into a herd of banteng. The event may have happened in open, dry, deciduous forests of Northern Cambodia, where several field biologists have reported the existence of temporary mixed herds between banteng and kouprey individuals (Edmond-Blanc 1947; Wharton 1957; Pfeiffer 1969). The preservation of this unique habitat is crucial for keeping hope of conserving the kouprey and many other threatened species, such as banteng, gaur, wild water buffalo, Eld’s deer, Asian elephant, tiger and leopard.

We are grateful to the people who provided tissue samples: Jean-Luc Berthier, Jean-Marie Carenton, Norin Chai, Philippe Chardonnat, Denis Debray, Gerard Dousseau, Bernard Dutrillaux, Nick Marx, Jean-Francois Marjorie, Sibyle Moulin, Bui Xuan Nguyen, Nhek Ratanapich, Claire Rejaud, Jacques Rigoulet, Roland Simon, Nhim Thy and SE Nhím Vanda. We also thank Brent Huffman for the photo of the Cambodian wild ox or kouproh. The mammals of the Indomalayan region: a systematic review. Oxford, UK: Natural History Museum publications, Oxford University Press.

REFERENCES


Derr, M. 2006 A celebrity among ungulates may soon be dismissed as a poseur. The New York Times (September 12).


Hassanin, A. & Ropiquet, A. 2007 What is the taxonomic status of the Cambodian banteng (*Bos javanicus*) and does it have close genetic links with the kouprey (*Bos sauveli*)? *J. Zool.*, 271, 246–252.


---

*The kouprey is a real species*  A. Hassanin & A. Ropiquet 2855