Molecular Phylogenetics and Evolution 66 (2013) 766–775

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Combining multiple autosomal introns for studying shallow phylogeny and taxonomy of Laurasiatherian mammals: Application to the tribe Bovini (Cetartiodactyla, Bovidae)

Alexandre Hassanin [a,b,*], Junghwa An [a,b], Anne Ropiquet [c], Trung Thanh Nguyen [a], Arnaud Couloux [d]

[a] Muséum national d'Histoire naturelle (MNHN), Département Systématique et Evolution, UMR 7205 – Origine, Structure et Evolution de la Biodiversité, 75005 Paris, France
[b] MNHN, UMS 2700, Service de Systématique Moléculaire, 75005 Paris, France
[c] Department of Conservation Ecology and Entomology, Stellenbosch University, Private Bag X1, Matieland 7602, Western Cape, South Africa
[d] Genoscope, Centre National de Séquençage, 91057 Evry, France

ARTICLE INFO

Article history:
Received 15 May 2012
Revised 27 September 2012
Accepted 1 November 2012
Available online 15 November 2012

Keywords:
Laurasiatheria
Bovinae
Nuclear introns
Phylogeny
Taxonomy
Mitochondrial introgression
Indels

ABSTRACT

Mitochondrial sequences are widely used for species identification and for studying phylogenetic relationships among closely related species or populations of the same species. However, many studies of mammals have shown that the maternal history of the mitochondrial genome can be discordant with the true evolutionary history of the taxa. In such cases, the analyses of multiple nuclear genes can be more powerful for deciphering interspecific relationships.

Here, we designed primers for amplifying 13 new exon-primed intron-crossing (EPIC) autosomal loci for studying shallow phylogeny and taxonomy of Laurasiatherian mammals. Three criteria were used for the selection of the markers: gene orthology, a PCR product length between 600 and 1200 nucleotides, and different chromosomal locations in the bovine genome. Positive PCRs were obtained from different species representing the orders Carnivora, Cetartiodactyla, Chiroptera, Perissodactyla and Pholidota.

The newly developed markers were analyzed in a phylogenetic study of the tribe Bovini (the group containing domestic and wild cattle, bison, yak, African buffalo, Asian buffalo, and saola) based on 17 taxa and 18 nuclear genes, representing a total alignment of 13,095 nucleotides. The phylogenetic results were compared to those obtained from analyses of the complete mitochondrial genome and Y chromosomal genes. Our analyses support a basal divergence of the saola (*Pseudoryx*) and a sister-group relationship between yak and bison. These results contrast with recent molecular studies but are in better agreement with morphology. The comparison of pairwise nucleotide distances shows that our nuDNA dataset provides a good signal for identifying taxonomic levels, such as species, genera, subtribes, tribes and subfamilies, whereas the mtDNA genome fails because of mtDNA introgression and higher levels of homoplasy. Accordingly, we conclude that the genus *Bison* should be regarded as a synonym of *Bos*, with the European bison relegated to a subspecies rank within *Bos bison*. We compared our molecular dating estimates to the fossil record in order to propose a biogeographic scenario for the evolution of Bovini during the Neogene.

1. Introduction

In animals, mitochondrial DNA (mtDNA) sequences are commonly used for species identification (DNA barcoding) or for studying phylogenetic relationships among closely related species (molecular systematics) or among populations of the same species (phylogeography and population genetics). Its utility is primarily explained by three major characteristics of the mitochondrial genome. First, its evolutionary rates are substantially faster than the nuclear genome (Brown et al., 1979; Saccone et al., 2006), which ensures obtaining sufficient nucleotide variation for comparing closely related organisms. Second, its simple and conserved structure, with two rRNA genes and 13 protein-coding genes, which lack introns and are separated by well-conserved tRNA genes (Boore, 1999), has facilitated the development of primers for PCR amplification of popular genes, such as the cytochrome *b* (*Cytb*) and the first subunit of the cytochrome *c* oxidase (CO1). Third, there are many mtDNA molecules per cell, which provides DNA extracts of sufficient quantity and quality from all kinds of tissues, including those collected from late Pleistocene remains (e.g., Reich et al., 2010).

* Corresponding author at: Muséum national d'Histoire naturelle (MNHN), Département Systématique et Evolution, UMR 7205 – Origine, Structure et Evolution de la Biodiversité, 75005 Paris, France. Fax: +33 1 40 79 30 63.
E-mail address: hassanin@mnhn.fr (A. Hassanin).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.11.003

A growing number of studies have, however, shown that the maternal history of the mitochondrial genome can be discordant with the evolutionary history of the species. In other words, the mtDNA tree and the species tree can show different relationships because of female introgression or incomplete lineage sorting of ancestral polymorphism (e.g., Verkaar et al., 2004; Ropiquet and Hassanin, 2006; Hassanin and Ropiquet, 2007; Alves et al., 2008; Nesi et al., 2011). Sequences of the Y chromosome have been used to detect mtDNA introgression in various orders of mammals, such as Carnivora, Cetartiodactyla, and Primates (Verkaar et al., 2004; Tosi et al., 2005; Hassanin and Ropiquet, 2007; Cabria et al., 2011; Roos et al., 2011). However, the use of genetic markers on the Y chromosome is not a panacea. First, it is often difficult to collect tissue from males, in particular for rare species. Second, male introgression has been recorded in a few studies, indicating that the signal of the Y chromosome can also be misleading at the species level (e.g., Roos et al., 2011). Genes located on autosomal chromosomes have been found to provide accurate information for identifying the direction of hybridization and for reconstructing shallow phylogenies (Ropiquet and Hassanin, 2006; Hassanin and Ropiquet, 2007; Fischer et al., 2011). However, multiple alignments of autosomal sequences often produce a small amount of nucleotide variation at the genus, species and subspecies levels. For instance, the nucleotide divergence between chimpanzee and human is 1.44% for the whole genome, 1.78% for the Y chromosome (Kuroki et al., 2006) and 8.79% for the mitochondrial genome (GenBank accession numbers NC_001643 versus NC_012920). Another problem is the incomplete lineage sorting of ancestrally polymorphic alleles, which is expected to be more frequent with autosomal markers because their coalescence times are four times longer, on average, than that of the uniparental markers (mtDNA and Y chromosome). This means that gene trees from different autosomal loci will reveal a significant proportion of paraphyletic or polyphyletic patterns at the species level. For example, the study of Fischer et al. (2011), based on 15 nuclear loci of chimpanzees, showed that the monophyly of *Pan paniscus* was supported by all loci, whereas that of *Pan troglodytes* was supported by only four loci (27%). In addition, six loci (40%) produced various paraphyletic patterns for *Pan troglodytes*. This study illustrates that the phylogenetic signal of a given autosomal marker may be too weak and potentially biased by incomplete lineage sorting, whereas the analysis of multiple autosomal loci is expected to provide a strong and reliable signal for interpreting relationships among closely related taxa.

Here, we designed primers for amplifying 13 autosomal introns that can be used for studying species of Laurasiatheria, one of the four major clades of placental mammals that includes the six orders Carnivora (feliformes and caniformes), Cetartiodactyla (artiodactyls and cetaceans), Chiroptera (bats), Eulipotyphla (hedgehogs, shrews and moles), Perissodactyla (horses, tapirs, and rhinos), and Pholidota (pangolins) (Murphy et al., 2001; Springer et al., 2011). The primers were tested on several species representing several orders of Laurasiatheria. The markers were then used for studying the phylogeny and taxonomy of the tribe Bovini (Cetartiodactyla, Bovidae), the group containing domestic and wild cattle, bison, yak, African buffalo, Asian buffalo, and saola. This taxonomic group is particularly interesting because two cases of ancient mtDNA introgression have been previously identified between species of *Bos* and *Bison* (Hassanin and Ropiquet, 2004, 2007; Verkaar et al., 2004) and because several molecular markers were already available for phylogenetic reconstruction, including the complete mtDNA genome (Hassanin et al., 2012) and three genes of the Y chromosome (Nijman et al., 2008).

2. Materials and methods

2.1. Selecting autosomal markers and primer design

The choice of intron DNA sequences was motivated by the presence of unique insertions and deletions (indels) at different taxonomic levels (e.g., Ropiquet and Hassanin, 2005a; Matthee et al., 2007). The three criteria used for the selection of the introns were gene orthology, a PCR product length between 600 and 1200 nucleotides (nt) in the three model species of Laurasiatheria (*Bos taurus*, *Canis familiaris*, *Equus caballus*), and different chromosomal locations in the genome of *B. taurus*.

The OrthoMaM database (Ranwez et al., 2007; http://www.orthomam.univ-montp2.fr/orthomam/html/) was used to edit a list of orthologous gene candidates using the sequences available for cow, dog and horse. We then chose several genes located on different bovine chromosomes and conducted BLAST runs on NCBI (http://www.ncbi.nlm.nih.gov/) to confirm that there was only one significant hit (score > 200 bits and query coverage > 50%) for each selected marker in the complete human genome (criterion of gene orthology). For each gene candidate the complete genomic sequences of *B. taurus*, *C. familiaris*, *E. caballus*, and *H. sapiens* were extracted from GenBank, and DNA sequences were aligned automatically using MUSCLE (Edgar, 2004). The primer pairs used for PCR amplification and DNA sequencing of nuclear introns were designed by comparing conserved exonic regions, and by selecting only introns with a length between 600 and 1200 nt in the three model species of Laurasiatheria.

A total of 13 primer pairs were drawn for PCR amplification from 13 different nuclear introns (Table 1): intron 6 of CCAR1 (cell division cycle and apoptosis regulator 1), intron 2 of CHPF2 (chondroitin polymerizing factor 2), intron 20 of DHX29 (DEAH (Asp-Glu-Ala-His) box polypeptide 29), introns 8 and 19 of DIS3 (DIS3 mitotic control), intron 4 of EXOSC9 (exosome component 9), intron 6 of HDAC1 (histone deacetylase 1), intron 10 of HDAC2 (histone deacetylase 2), introns 2 and 6 of PABPN1 (poly(A) binding protein, nuclear 1), intron 6 of RIOK3 (RIO kinase 3), intron 9 of TUFM (elongation factor Tu, mitochondrial precursor), and intron 6 of ZFYVE27 (zinc finger, FYVE domain containing 27).

2.2. DNA extraction, amplification and sequencing

Total genomic DNA was extracted from blood, hair, muscle, skin, or cryopreserved cells using QIAGEN DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany).

Amplifications were done in 20 μl using 3 μl of Buffer 10× with $MgCl_2$, 2 μl of dNTP (6.6 mM), 0.12 μl of Taq DNA polymerase (2.5 U, Qiagen, Hilden, Germany) and 0.75 μl of the two primers at 10 μM. The standard PCR conditions were as follows: 4 min at 95 °C; 5 cycles of denaturation/annealing/extension with 45 s at 95 °C, 1 min at 60 °C and 1 min at 72 °C, followed by 30 cycles of 30 s at 95 °C, 45 s at 55 °C, and 1 min at 72 °C, followed by 10 min at 72 °C. PCR products were resolved by electrophoresis on a 1.5% agarose gel stained with ethidium bromide and visualized under UV light.

We analyzed 18 independent autosomal genes for studying phylogenetic relationships within the tribe Bovini (Cetartiodactyla, Bovidae). The markers included 11 introns specifically developed for this study, and seven additional markers that were amplified and sequenced using the primers published in previous papers: exon 4 of CSN3 (Ropiquet and Hassanin, 2005b), intron 8 of FGB (Hassanin and Ropiquet, 2007), intron 2 of LALBA (Hassanin and Douzery, 2003), the promotor of LF (Hassanin and Douzery, 1999a), intron 9 of PRKCI (Ropiquet and Hassanin, 2005a), intron 13 of SPTBN1, and intron 9 of TG (Ropiquet and Hassanin, 2006).

Table 1
The 13 markers developed in this study.

Markers[a]	Primers		Sequence length and chromosome location in			
			Bos taurus	*Canis familiaris*	*Equus caballus*	*Homo sapiens*
CCAR1 I6	CCAR1-EX6U	GATGTATTCTTTCAGCTTAGGT	960/ch 28	898/ch 4	896/ch 1	4618/ch 10
	CCAR1-EX7L	GCTTCAACCAATACTCTGTCAC				
CHPF2 I2	CHPF2-EX2U	GAGTGGCTYGGCCGCTGCCTCAT	872/ch 4	763/ch 16	910/ch 4	974/ch 7
	CHPF2-EX3L	CACTGTASGCCCGCTCCAGCTCCA				
DHX29 I20	DHX29-EX20U	GCCTTTGGAGGAATTATGTCTTCA	683/ch 20	741/ch 2	742/ch 21	784/ch 5
	DHX29-EX21L	TCAGGAGAACCAAGATTGCA				
DIS3I8	DIS3-EX8U	GTTGCTRTTGATGGTTGGC	713/ch 12	726/ch 22	708/ch 17	990/ch 13
	DIS3-EX9L	CAGTRATGCTCCAGGGCAT				
DIS3 I19	DIS3-EX19U	CAAAGTATGGTTTAGAAGGTACAGT	839/ch 12	726/ch 22	709/ch 17	407/ch 13
	DIS3-EX20L	GGACATTCGGATTTTCTGATGCTG				
EXOSC9 I4	EXOSC9-EX4U	CAGATCTCTTGGTGAAGTTGAATCG	984/ch 6	973/ch 19	987/ch 2	1677/ch 4
	EXOSC9-EX5L	GTCCACACGTATTTGCCAAAC				
HDAC1 I6	HDAC1-EX6U	AATAYGGAGAGTACTTCCCAG	1110/ch 2	1065/ch 2	1142/ch 2	1425/ch 1
	HDAC1-EX7L	GTAGTTAACAGCATAATACTTGCC				
HDAC2 I10	HDAC2-EX10U	GACAAACCAGAACACTCCAG	690/ch 9	864/ch 12	694/ch 10	990/ch 6
	HDAC2-EX11L	GGTGCATGAGGYAACATGCGT				
PABPN1 I2	PABPN1-E2U	GAAGCAGATGAATATGAGTCCACC	721/ch 10	743/ch 14	756/ch 1	725/ch 14
	PABPN1-E3L	CTTCTCCTCAATGGACATGAT				
PABPN1 I6	PABPN1-E6U	GCCCGRACCACCAACTACAAC	986/ch 10	960/ch 14	1002/ch 1	1025/ch 14
	PABPN1-E7L	ATACCATGATGTCGCTCTAGC				
RIOK3 I6	RIOK3-EX6U	GCCTACTCAGAAGAACGTCGAAG	701/ch 24	689/ch 7	746/ch 8	1166/ch 18
	RIOK3-EX7L	ACGTGTCTTAGGATCAACTG				
TUFM I9	TUFM-EX9U	CTGACTTGGGACATGGCCTGTCG	838/ch 25	658/ch 6	652/ch 13	663/ch 16
	TUFM-EX10L	ACGCTGGCCTTTYTCTAAGATCAT				
ZFYVE27 I6	ZFYVE27-EX6U	GAATGTGGAGTTCTTCCGAG	693/ch 26	747/ch 28	757/ch 1	766/ch 10
	ZFYVE27-EX7L	GGGTTCATCCGCCGCTGCAGA				

[a] Intron number in *Bos taurus*. Abbreviations: ch = chromosome; nt = nucleotide.

Both strands of PCR products were sequenced using Sanger sequencing on an ABI 3730 automatic sequencer at the *Centre National de Séquençage* (Genoscope) in Evry (France). The sequences were edited and assembled using Sequencher 5.0 (Gene Codes Corporation). Heterozygous positions (double peaks) were scored using the IUPAC ambiguity codes. Sequences generated for this study were deposited in the EMBL/DDBJ/GenBank database (accession numbers JX861896–JX862178; see Appendix A for details).

2.3. Phylogenetic analyses

The phylogeny of the tribe Bovini was investigated using 18 autosomal genes for 17 taxa (Appendix A). The outgroup species were chosen on the basis of previous molecular studies on the family Bovidae (e.g., Hassanin and Douzery, 1999b; Hassanin et al., 2012). They include three species representing three different tribes of the family Bovidae: *Boselaphus tragocamelus* (Bovinae, Boselaphini), *Tragelaphus oryx* (Bovinae, Tragelaphini), and *Capra hircus* (Antilopinae, Caprini).

For each gene, DNA sequences were aligned with MUSCLE using the default settings (Edgar, 2004) and refined manually on Se–Al v2.0a11 (A. Rambaut. Sequence Alignment Editor Version 2.0 alpha 11. 2002; http://evolve.zoo.ox.ac.uk/). The concatenated dataset represents a total alignment of 13,095 nucleotides (Appendix B).

The best-fitting models of sequence evolution were selected under jModelTest (Posada, 2008) for each marker and the final concatenation. Using the Akaike Information Criterion (AIC), we selected the K80 model for LF, the K80+I+G model for TG, the HKY model for PRKCI, the HKY+I model for DHX29, HDAC1, and SPTBN1, the HKY+G model for DIS3, TUFM, and ZFYVE27, the SYM+I+G model for CSN3, the GTR model for HDAC2, PABPN1 and LALBA, the GTR+I model for RIOK3 and FGB, and the GTR+G model for CCAR1, CHPF2, EXOSC9, and the concatenation of the 18 autosomal markers.

Phylogenetic analyses were performed using Bayesian inference (BI) and maximum likelihood (ML) methods, with gaps treated as missing data. Bayesian inferences were conducted using MrBayes v3.1.2 (Huelsenbeck and Ronquist, 2001) on each of the 18 markers and on the concatenated dataset. The posterior probabilities (PP) were calculated using four independent Markov chains run for 10,000,000 Metropolis-coupled MCMC generations, with tree sampling every 1000 generations and a burn-in of 5000 trees. The concatenated dataset was also analyzed for ML reconstruction using PAUP* version 4b10 (Swofford, 2003). The most-parsimonious tree was used as a starting tree to estimate the likelihood parameters of the GTR+G model selected under jModelTest. These parameters were then included in a heuristic search using Tree Bisection Reconnection (TBR) branch-swapping. The resulting tree was saved and the new parameters used in the subsequent ML analysis. The whole procedure was reiterated on the basis of the newly generated ML tree until stabilization of both parameters and ML score was achieved. ML bootstrap values (BPs) were computed on 1000 pseudo-replicates using the parameters of the final analysis (base frequencies: A = 0.267664, C = 0.201144, G = 0.216134, and T = 0.315058; rate matrix: AC = 1.21057, AG = 5.87923, AT = 0.63178, CG = 1.48500, CT = 4.86388, and GT = 1.00000; shape = 0.381620).

The lists of bipartitions obtained from the Bayesian analyses of the 18 independent autosomal markers were transformed into a weighted binary matrix for supertree construction using SuperTRI v57 (Ropiquet et al., 2009; Python script available at http://www.normalesup.org/~bli/Programs/programs.htmlupon). Each binary character corresponded to a node, which was weighted according to its frequency of occurrence in one of the 18 lists of bipartitions. This way, the SuperTRI method accounts for both principal and secondary signals, as all phylogenetic hypotheses found during the analyses are represented in the weighted binary matrix used for supertree construction. The reliability of the nodes was assessed using three measures: supertree bootstrap percentages (SBPs) were obtained from PAUP* after 1000 BP replicates of the matrix of 3959 binary characters generated by SuperTRI v57, and the mean posterior probabilities (MPP) and reproducibility indices (Rep) were directly calculated on SuperTRI v57. The SBP, MPP, and Rep values were reported on the Bayesian tree found with the concatenated dataset of 13,095 nt (Fig. 1). Here, the

SuperTRI analyses were conducted to test for repeated phylogenetic signals in the different autosomal genes. If a node was recovered with high SBP, MPP, and Rep values (SBP > 50; MPP > 0.5; Rep > 0.5), we concluded that the signal was robust and reliable, i.e., found in most nuclear markers. If a node was recovered with low Rep values (⩽0.11), we concluded that the signal was restricted to one or two nuclear markers.

2.4. Molecular dating analyses

Divergence dates were estimated using a Bayesian relaxed molecular clock approach implemented in PhyloBayes 3.1 (Lartillot et al., 2009) using the concatenated nucleotide dataset and the Bayesian topology of Fig. 1. All absolute ages of the geological periods and chronostratigraphic references were taken from the 2012 Geological Time Scale of the International Chronostratigraphic Chart (http://www.stratigraphy.org). For fossil calibrations, we selected two time constraints: the split between American bison and European bison was set at 230 ± 10 ka, as the first entrance of bison into North America was dated between 220 and 240 ka (Scott, 2010); and second, the first occurrence of the Bovidae, which is also well documented in the Early Miocene of Africa and Eurasia, was set at 20 ± 2 Ma (Vrba and Schaller, 2000). As recommended by Lartillot et al. (2009), the analyses were conducted using the CAT-GTR+Gamma 4 model and a log-normal autocorrelated clock relaxation model. In addition, we applied the BD-SOFT model, which used birth–death (BD) priors on divergence times, combined with soft fossil calibrations (Yang and Rannala, 2006). The calculations were conducted by running a MCMC chain for 10,000 cycles sampling posterior rates and dates every 10 cycles. Posterior estimates of divergence dates were then computed from the last 7500 cycles accounting for the initial burnin period.

3. Results

3.1. Characteristics of the newly developed markers

Table 1 details the characteristics of the 13 newly developed autosomal markers, including primer sequences, chromosomal location in the genome of *B. taurus*, *C. familiaris*, *E. caballus*, and *H. sapiens*, and product lengths in these four model species. These 13 markers represent 11 unlinked genes, which are located on different chromosomes in the bovine genome. Note, however, that some of these genes are located on the same chromosome in other model species. In *C. familiaris*, DHX29 and HDAC1 genes are located on chromosome 2. In *E. caballus*, the three genes CCAR1, PABPN1 and ZFYVE27 are located on chromosome 1, whereas EXOSC9 and HDAC1 are located on chromosome 2. In *H. sapiens*, CCAR1 and ZFYVE27 are located on chromosome 10.

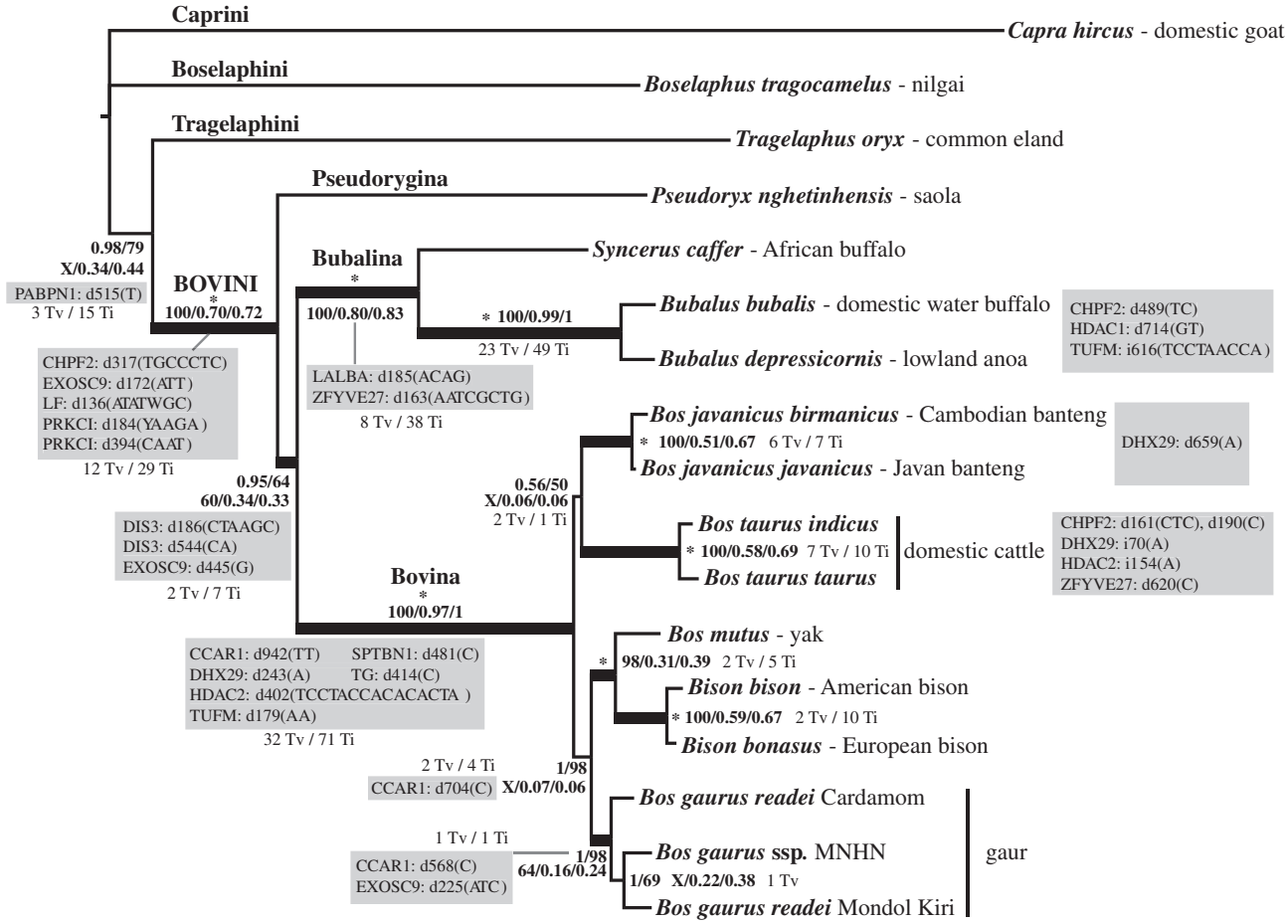

Fig. 1. Phylogeny of the tribe Bovini. The Bayesian tree was reconstructed from the concatenated dataset combining all the 18 independent autosomal genes (17 taxa and 13,095 nucleotide characters) with a GTR+G model. For each internal branch, the two first values correspond to the Bayesian posterior probability (PP, to the left of the slash) and maximum likelihood bootstrap percentage (BP, the right of the slash). The asterisk indicates that the node was supported by maximal values of robustness (PP = 1; BP = 100). The values below are the three measures of reliability obtained from the SuperTRI analyses of the 18 nuclear genes: from left to right, Supertree Bootstrap percentage (SBP), mean posterior probability (MPP) and the Reproducibility index (Rep). Thick branches indicate nodes that were considered reliable (for more details, see Section 3). The letter "X" indicates that the node was not found in the analysis. The position and nature of all diagnostic indels (i: insertion; d: deletion) shared by at least two taxa in the alignments of nuclear genes are highlighted in grey.

All the 13 primer pairs designed for this study (Table 1) were tested for PCR amplification in a few species of Laurasiatheria representing the orders Cetartiodactyla (17 bovid species), Carnivora (*Genetta servalina*), Chiroptera (*Hipposideros armiger*), Perissodactyla (*Equus quagga*), and Pholidota (*Phataginus tricuspis*). A single band was amplified for all samples using standard PCR conditions, suggesting that these markers can be easily used for studying Laurasiatherian species.

3.2. Phylogeny of the tribe Bovini

Among the 13 newly generated nuclear markers, 11 were sequenced and aligned for studying the phylogeny of the tribe Bovini using 17 taxa (Appendix A): CCAR1 I6 (1025 nt), CHPF2 I2 (887 nt), DHX29 I20 (685 nt), DIS3 I19 (848 nt), EXOSC9 I4 (985 nt), HDAC1 I6 (1106 nt), HDAC2 I4 (722 nt), PABPN1 I2 (730 nt), RIOK3 I6 (719 nt), TUFM I9 (855 nt), ZFYVE27 I6 (694 nt). In addition, seven markers used in previous studies were also sequenced and aligned: CSN3 E4 (418 nt), FGB I8 (693 nt), LALBA I2 (485 nt), LF (337 nt), PRKCI I9 (525 nt), SPTBN1 I13 (595 nt), and TG I9 (786 nt). The 18 markers belong to unlinked genes, which have different chromosomal locations in the bovine genome: chromosome 1 for PRKCI, chromosome 2 for HDAC1, chromosome 4 for CHPF2, chromosome 5 for LALBA, distant regions of the chromosome 6 for EXOSC9 and CSN3 (corresponding to positions 3429840–3430823 and 87390280–87390679, respectively; http://genome.ucsc.edu/), chromosome 9 for HDAC2, chromosome 10 for PABPN1, chromosome 11 for SPTBN1, chromosome 12 for DIS3, chromosome 14 for TG, chromosome 17 for FGB, chromosome 20 for DHX29, chromosome 22 for LF, chromosome 24 for RIOK3, chromosome 25 for TUFM, chromosome 26 for ZFYVE27, and chromosome 28 for CCAR1.

The final concatenated dataset combining the 18 autosomal markers contained 13,095 aligned positions. The Bayesian tree reconstructed from this dataset is depicted in Fig. 1. The tribe Bovini and the subtribes Bovina and Bubalina were found to be monophyletic with maximal support values (PP = 1; BP = 100), and these three groups were highly supported by the separate analyses of the 18 independent markers (SBP = 100; 0.70 < MPP < 0.97; 0.72 < Rep < 1; Appendix C). They are also characterized by a large number of non-homoplastic synapomorphies, including substitutions and deletions (Fig. 1; Appendix D).

Regarding interrelationships between the three subtribes of Bovini, our dataset supports a clade uniting Bovina and Bubalina (PP = 0.95; BP = 64; SBP = 62; MPP = 0.34) to the exclusion of Pseudorygina. This node was recovered in six separate Bayesian analyses of the 18 markers (Rep = 0.33), but with PP > 0.5 for only four markers: CCAR1 (PP = 1), DIS3 (PP = 0.72), EXOSC9 (PP = 0.97), FGB (PP = 0.77). The two other possible hypotheses, Pseudorygina as sister to either Bovina or Bubalina, were less supported (BP = 0/34; SBP = 0/31; MPP = 0.21/0.28). The association between *Pseudoryx* and Bovina was found in three separate analyses of the 18 markers (Rep = 0.17), but with PP > 0.5 for only two markers: CHPF2 (PP = 0.59) and LALBA (PP = 0.63). The association between *Pseudoryx* and Bubalina was found in seven separate analyses of the 18 markers (Rep = 0.39), but with PP > 0.5 for only two markers: DHX29 (PP = 0.96) and SPTBN1 (PP = 0.98). All other markers provided a weak signal for relationships between the three subtribes. The analysis of molecular signatures also confirmed a basal position for Pseudorygina (Fig. 1 and Appendix D). There were 12 diagnostic mutations linking Bovina and Bubalina, including three deletions (DIS3: CTAAGC at pos. 186 and CA at pos. 544; EXOSC9: G at pos. 445), two transversions (CCAR1: T → G at pos. 74 and A → C at pos. 165) and seven transitions (DIS3: C → T at pos. 471; EXOSC9: A → G at pos. 737, T → C at pos. 833; HDAC1: C → T at pos. 632; TUFM: T → C at pos. 322; FGB: T → C at pos. 28 and 122). In comparison, neither of the two other hypotheses is supported by any indel; seven non-homoplastic substitutions support the clade uniting *Pseudoryx* and Bubalina, including two transversions (LALBA: A → T at pos. 401; SPTBN1: T → A at pos. 75) and five transitions (DHX29: G → A at pos. 297 and A → G at pos. 441; HDAC1: T → C at pos. 393; TUFM: T → C at pos. 277; SPTBN1: T → C at pos. 269), whereas only two non-homoplastic transitions support the clade uniting *Pseudoryx* with Bovina (CHPF2: G → A at pos. 299; EXOSC9: A → G at pos. 695).

All species of Bovina for which at least two members were represented were recovered as monophyletic with high support values (PP = 1; BP > 98): *Bos gaurus*, *B. javanicus* and *B. taurus*. Each of these three species can be diagnosed by one or two specific indels and by several substitutions (Fig. 1). The monophyly of the genus *Bison* was also strongly supported by the supermatrix analyses (PP = 1; BP = 100), and by PP > 0.5 in the separate analyses of nine markers (CCAR1: 0.98; CHPF2: 1; DHX29: 0.97; DIS3: 1; HDAC1: 0.82; HDAC2: 1; PABPN1: 0.82; TUFM: 1; FGB: 0.94). In contrast, a paraphyletic pattern was obtained for the genus *Bos*, due to the sister-group relationship between *Bos mutus* and *Bison*. The association of yak and bison was highly supported in the supermatrix analyses (PP = 1; BP = 100) and also found to be the most highly supported hypothesis in the SuperTRI analyses (SBP = 100; MPP = 0.31; Rep = 0.39). It was also strongly supported in the Bayesain analyses of five markers (0.92 < PP < 1 for DHX29, EXOSC9, LALBA, SPTBN1, and ZFYVE27; Appendix C). Moreover, we found that they share seven exclusive substitutions, which are located on each of these five different genes (Appendix D). The clade uniting *B. mutus*, *B. bison*, *B. bison*, and *B. gaurus* was found to be robust in the supermatrix analyses (PP = 1; BP = 98), but the SuperTRI analyses indicate that this node is not reliable (not found in the SBP consensus tree/MPP = 0.07/Rep = 0.06). This grouping was supported by only one marker, CCAR1 (PP = 1; Appendix C). Similarly, the association of *B. javanicus* and *B. taurus* was only found with CCAR1 (PP = 1; Appendix C). In this case, however, the node was not robust in the supermatrix analyses (PP = 0.56; BP = 50).

For comparison, we reanalyzed the three Y chromosomal genes studied by Nijman et al. (2008; SRY, ZFY, and DBY). Our alignment included two additional taxa, for which DNA sequences were downloaded from GenBank (Appendix E): *Bubalus bubalis carabanensis* (SRY: GQ259331; ZFY: GQ259329; DBY: GQ259327) and *Capra hircus* (SRY: D82963; ZFY: AY082500 and FJ349608; DBY: not available). The Bayesian tree obtained from the concatenated dataset (4921 characters for 12 taxa) is presented in Appendix F. Several taxa were recovered as monophyletic, including the subtribes Bovina and Bubalina, the genus *Bison*, and the species *Bos gaurus*, *Bos taurus*, and *Bubalus bubalis*. The genus *Bison* was not found to be associated with *Bos mutus*, the latter being the first species to diverge within Bovina. However, the basal position of the yak was found with weak support values (PP = 0.75; BP = 59), and it was diagnosed by only one non-homoplastic substitution, corresponding to a transition A → G at position 3186 of our Y chromosomal alignment (Appendix E). Similarly, all other interspecific relationships within Bovina are not supported by more than one molecular signature, indicating that the Y chromosomal dataset contains a very weak phylogenetic signal. There are only two exceptions: the genus *Bison* and the species *Bos gaurus*, which are both characterized by four exclusive substitutions (Appendix G).

We also reanalyzed the alignment used in MacEachern et al. (2009), but without the ancestral sequence inferred for Bovini, named "Ancient" in their study. The topologies of our Bayesian and ML trees were identical (Appendix H). The species *Bos javanicus* was associated to the clade uniting *Bos taurus*, *Bos mutus*, and *Bison bison* (PP = 0.63; BP = 47), whereas it was a sister-group of *Bos gaurus* in the NJ tree of MacEachern et al. (2009) (BP = 89). Other relationships were identical to those previously published:

Bos mutus was associated wih *Bos taurus* (PP = 1; BP = 80), and the clade was linked to *Bison bison* (PP = 1; BP = 74).

3.3. Nucleotide variation at different taxonomic levels

Fig. 2 compares the pairwise nucleotide distances calculated from our nuclear dataset with those calculated from an alignment of complete mitochondrial genomes (data from Hassanin et al., 2012). Interestingly, nuDNA distances can be ranked according to recognized taxonomic levels: between 5.40% and 4.85% for comparisons between the subfamilies Bovinae and Antilopinae; between 4.06% and 3.62% for comparisons between the tribes Bovini, Boselaphini, and Tragelaphini; between 2.84% and 2.36% for comparisons between the subtribes Bovina, Bubalina and Pseudorygina; between 1.57% and 1.55% for comparisons between genera of Bubalina; between 0.71% and 0.36% for comparisons between species of Bovina; and only 0.26% between *Bubalus bubalis* and *B. depressicornis*. Surprisingly, the nuDNA distance between *Bison bison* and *Bison bonasus* was only 0.08%, which is similar to the nuDNA distances found for intraspecific comparisons: 0.06% for *Bos javanicus*, 0.10% for *Bos taurus*, and 0.18% for *Bos gaurus* (not shown).

The mtDNA distances were always higher than the nuDNA distances, but they cannot be used to distinguish subfamilial, tribal, and subtribal levels, because the ranges of percentages are overlapping between these three taxonomic levels: 13.18–12.60% for subfamilial comparisons, 12.94–11.34% for tribal comparisons, and 12.09–10.42% for subtribal comparisons (Fig. 2). In addition, intraspecific and interspecific distances were highly variable. For instance, intraspecific distances were 3.71% for *Bos javanicus*, but only 1.22% for *Bos taurus*. The distance between species of *Bison* was 5.68%, which is in the range of interspecific comparisons within Bovina (2.71% between *Bos mutus* and *Bison bison* and 6.15% between *Bos taurus taurus* and *Bos javanicus javanicus*).

3.4. Molecular estimates of divergence time

Table 2 presents the estimates of divergence time obtained from our nuclear dataset, as well as a comparison with the recent analyses of Hassanin et al. (2012) based on complete mtDNA genomes. The results of the two studies are very similar, although molecular divergence times were estimated using different datasets (nuDNA versus mtDNA), different taxonomic samples (Bovini versus Cetartiodactyla), and different calibration points (two in the present study versus five in the mtDNA study).

4. Discussion

4.1. Supermatrix versus SuperTRI analyses

In the supermatrix approach, different molecular markers are combined together into a single data matrix, which therefore contains a large number of characters for phylogenetic reconstruction. The main advantage of a supermatrix is that phylogenetic signals from different markers, including weak signals, can be combined and thereby enhanced. For this reason, a tree constructed from a supermatrix is often more resolved and robust than a tree developed from separate analyses of markers (e.g., Murphy et al.,

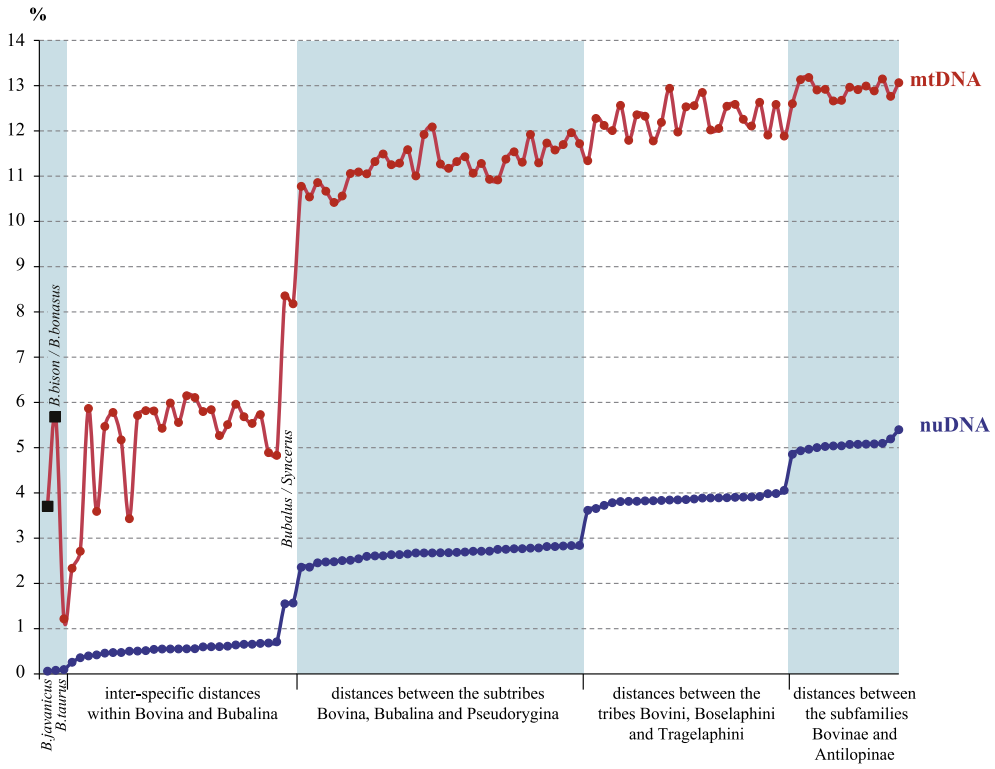

Fig. 2. Comparisons of pairwise nucleotide distances between nuclear and mtDNA datasets. The pairwise nucleotide distances were estimated using the nuclear alignment used in this study (13,095 nt, in blue) and the alignment of mtDNA genomes (14,902 nt, in red) published in Hassanin et al. (2012). All the 17 taxa in Fig. 1 were included in the comparisons, with the exception of two individuals of *Bos gaurus* that were not available for mtDNA sequencing. The nuclear distances were ranked in ascending order, and they were compared to the corresponding pairwise distances obtained from mtDNA sequences. The nuclear distances follow the different taxonomic levels, with the exception of the divergence between *Bison bison* and *Bison bonasus*, which is comparable to intraspecific distances calculated for *Bos javanicus* and *Bos taurus*. The two black squares indicate the two mtDNA distances distorted by an ancient event of mtDNA introgression: first, the mtDNA of the Cambodian banteng, *Bos javanicus birmanicus*, was acquired from the kouprey, *Bos sauveli*; second, the mtDNA of the European bison, *Bison bonasus*, was acquired from a species related to *Bos taurus* (for details, see Hassanin and Ropiquet, 2004, 2007). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Divergence times estimated with our nuDNA dataset, and comparisons with the dates calculated from complete mtDNA genomes (Hassanin et al., 2012).

Node	nuDNA		mtDNA		Geologic epoch
	Mean (Ma)	SD (Ma)	Mean (Ma)	SD (Ma)	
Bovidae[a]	19.67	1.20	19.75	1.86	Early Miocene
Bovinae	17.47	1.93	16.41	2.03	Early/Middle Miocene
Bovini + Tragelaphini	16.80	1.94	15.77	1.96	Early/Middle Miocene
Bovini	13.53	1.86	13.33	1.95	Middle Miocene
Bovina + Bubalina	13.06	1.85	NA	NA	Middle Miocene
Bovina	3.75	1.14	3.31	0.92	Pliocene
Bison + *Bos mutus*	1.98	0.60	NA	NA	Lower Pleistocene
Bison[a]	0.23	0.01	NA	NA	Middle Pleistocene
Bos gaurus	1.85	0.71	NA	NA	Lower Pleistocene
Bos javanicus	0.57	0.42	NA	NA	Middle Pleistocene
Bos taurus	0.35	0.20	0.50	0.19	Middle Pleistocene
Bubalina	8.94	1.70	8.16	1.68	Late Miocene
Bubalus	1.29	0.48	1.29	0.47	Lower Pleistocene

[a] Two calibration points used in this study (see Section 2). Abbreviations: Ma = million years; NA = not available; SD = standard deviation.

2001; Regier et al., 2010; Kocot et al., 2011). However, the supermatrix approach can be misleading if the species-tree signal is not dominant after the combination of the data sets (Ropiquet et al., 2009). Under certain circumstances, one marker can impose its own erroneous signal upon all other markers. For example, shallow phylogenies can be strongly biased by mtDNA introgression if the supermatrix corresponds to a concatenation of mtDNA and nuDNA markers. In such cases, the mtDNA markers impose their erroneous phylogenetic signal because mtDNA alignments generally contain higher levels of nucleotide variation than nuclear alignments. Similarly, the retention of ancestral alleles in one or several nuclear markers (incomplete lineage sorting) can lead to erroneous interpretations of recent relationships. Separate analyses of the markers are therefore necessary to avoid misinterpretations. The SuperTRI method developed by Ropiquet et al. (2009) is a robust way to summarize the results of separate analyses of markers. The conceptual basis of the SuperTRI approach is that a node recovered in several phylogenetic analyses of independent loci is reliable, whereas a node that is highly supported by only one marker, but never found by others, is less reliable.

Our analyses have shown that there is no significant incongruence between supermatrix and SuperTRI results for most nodes of the tree in Fig. 1. This means that the nodes highly supported by the supermatrix analyses are retrieved as being the best hypotheses in the SuperTRI analyses. However, there is one major exception concerning the clade uniting the yak, bison and gaur: it is highly supported by the supermatrix analyses (PP = 1; BP = 98), but not by the SuperTRI indices (not found in the SBP consensus tree/MPP = 0.07/Rep = 0.06). Our separate analyses show that this node is only supported by CCAR1 (PP = 1; Appendix C). Since this node was not retrieved with any other marker, it cannot be considered reliable. In fact, the retention of ancestral polymorphism in the CCAR1 alleles sequenced for both *Bos taurus* and *Bos javanicus* may explain the artificial grouping of yak, bison, and gaur.

4.2. Systematic relationships in the tribe Bovini

Most phylogenetic studies based on DNA sequences have identified three divergent lineages within the tribe Bovini corresponding to the subtribes Bovina, Bubalina, and Pseudorygina, but their interrelationships were unresolved due to the absence of a robust signal (Hassanin and Douzery, 1999a; Gatesy and Arctander, 2000; Hassanin and Ropiquet, 2004; Hassanin et al., 2012). Molecular dating estimates have suggested that the three subtribes diverged from one another over a short period of time during the middle Miocene, between 13.5 and 13.1 ± 1.9 Ma (Table 2). Our analyses of 18 autosomal markers indicate, however, that a sister-group relationship between Bovina and Bubalina is the most strongly supported phylogenetic hypothesis, being supported by 12 molecular signatures detected in six independent different markers (CCAR1, DIS3, EXOSC9, FGB, HDAC1, and TUFM) and including three diagnostic deletions (Fig. 1).

The question of interspecific relationships among species of Bovina remains highly debated among molecular biologists. Although the genus *Bison* was found to be polyphyletic using mtDNA sequences, it was retrieved as monophyletic in the analyses of Y chromosomal genes (Verkaar et al., 2004; Nijman et al., 2008). Here, our analyses have shown that nine of the total 18 autosomal genes strongly support the monophyly of *Bison* ($0.82 < PP < 1$). Although discordant with mtDNA data, the nuclear results are in strong agreement with morphology (Groves, 1981; Geraads, 1992). Verkaar et al. (2004) proposed two hypotheses for explaining the anomalous divergence of the mtDNAs between the two bison species: the first hypothesis is incomplete lineage sorting, which implies that two distinct mitochondrial lineages coexisted until the recent divergence of American and European bison; the second hypothesis is that the European bison has emerged by species hybridization initiated by introgression of bison bulls in another ancestral species. Our data indicate that American and European bison share very similar autosomal sequences, as the nucleotide distance is only 0.08%. This percentage of divergence is within the range of intraspecific variation in *B. gaurus*, *B. javanicus*, and *Bos taurus* (between 0.06% and 0.18%). Therefore, the analyses of autosomal markers strongly suggest that the mitochondrial genome of European bison was acquired by introgression after one or several past events of interspecific hybridization between a male of European bison and a female of an extinct species that was closely related to the aurochs (the wild ancestor of domestic cattle). The mtDNA introgression probably arose somewhere in Europe after the divergence between American and European bison estimated at 230 ± 10 ka (Scott, 2010). From a taxonomic point of view, our molecular results suggest, first, to synonymize the genus *Bison* with *Bos*, and second, to treat the American and European bison as two subspecies of the same species, *Bos bison bison* and *Bos bison bonasus*, respectively. Our molecular results corroborate previous taxonomic propositions based on morphology (e.g., Bohlken, 1958; Groves, 1981; Geraads, 1992), and illuminate the fact that hybrids between American and European bison are fertile in both sexes (Van Gelder, 1977).

Our nuclear analyses also support a sister-group relationship between bison and yak. Although in agreement with the analyses of amplified fragment length polymorphism (AFLP) fingerprinting published by Buntjer et al. (2002), this result contrasts with recent molecular studies, which have found different relationships: a ba-

sal position of *Bos mutus* using Y chromosomal genes (Nijman et al., 2008) or an association of *Bos mutus* with *Bos taurus* using other autosomal genes (MacEachern et al., 2009). However, our reanalyses have shown that the datasets used in Nijman et al. (2008) and MacEachern et al. (2009) do not contain a strong signal for the position of both bison and yak ($59 < BP_{ML} < 80$; Appendices F–H). For example, the early divergence of *Bos mutus* is only supported by the fact that all other species of Bovina share a G nucleotide in position 3186 of the Y chromosomal alignment (Appendix E). Moreover, MacEachern et al. (2009) hypothesized that their result showing the association of yak with domestic cattle was erroneous, because their autosomal sequences were probably generated from a yak–taurine hybrid rather than a pure wild yak. By contrast, our analyses show that the clade uniting bison and yak is robust and reliable (i.e., repeated with strong support in the analyses of five independent nuclear markers). Interestingly, this result is in perfect agreement with most morphological studies (Groves, 1981; Olsen, 1990; Geraads, 1992). Groves (1981) indicated several possible derived characters uniting bison and yak: in their skulls, the intercornual ridge is flattened and craniocaudally expanded in the midline; they have 14 thoracic vertebrae and 5 lumbars, whereas other Bovina have 13 thoracic vertebrae and 6 lumbars; their rhinarium is more overgrown with hair, leaving only a narrow bare strip around the nostrils, as well as the philtrum; and they have long hair and a beard. Geraads (1992) wrote that the association of yak and bison "is supported by a broad skull, supra-orbital foramina which are wide apart, acute parieto-frontal suture, loss of ethmoidal fissure, widely inserted horn-cores, loss of contact between nasal and premaxilla, short premolar row, mesial lobe of P_4 wide open, and more simple molars". In addition, a sister-group relationship between bison and yak agrees with the paleogeographic scenario recently proposed by Deng et al. (2011), who suggests that the Tibetan Plateau was the center of origin of their most recent common ancestor. A further argument in favor of this hypothesis is that the mtDNA genome contains a robust signal in favor of the clade composed of yak and American bison (Hassanin and Ropiquet, 2007; Hassanin et al., 2012).

4.3. Biogeographic evolution of the Bovini

The earliest fossils of Bovini are represented by the genus *Selenoportax*† from the late Miocene of Pakistan (Siwaliks; 10.2 Ma) and Myanmar (Irrawaddy Formation; 10.4–8.8 Ma (Bibi, 2007). The biogeographic analyses of Hassanin and Ropiquet (2004) also support an Asian origin for Bovini. Since all three subtribes are currently more diversified in Southeast Asia, this region is likely to constitute the center of origin of Bovini. A basal position of Pseudorygina, as supported by our nuclear analyses, gives even greater credence to this hypothesis. Molecular dating estimates have suggested that the common ancestor of crown Bovini experienced a rapid diversification during the middle Miocene, i.e. between 13.5 and 13.1 ± 1.9 Ma (Table 2). The evolution of Bovini in Southeast Asia may have been driven by changes in vegetation distribution, as paleoecological data have documented oscillations between tropical woodlands and grasslands in northern Thailand between 13.3 and 13.1 Ma (Sepulchre et al., 2010).

Today, there are two divergent geographic lineages within the subtribe Bubalina: *Bubalus*, which is found across the Indian subcontinent and Southeast Asia, and *Syncerus caffer*, which is endemic to sub-Saharan Africa (IUCN, 2011). Our molecular estimations have indicated that the two groups diverged from each other in the late Miocene, between 8.2 and 8.9 ± 1.7 Ma (Table 2). However, the fossil record indicates that these two groups appeared in the Pliocene. The lineage leading to *Bubalus* is represented by *Proamphibos*† in the Pliocene of the Siwaliks between 3.5 and 2.2 Ma, followed by *Hemibos*† during early Pleistocene times in Eurasia, and by *Bubalus*, which is well diversified in the middle and late Pleistocene of the Indian subcontinent, Europe, China, and Southeast Asia (Khan and Akhtar, 2011; Martinez-Navarro et al., 2011). The lineage leading to *Syncerus* is represented in Africa by *Ugandax*†, between 5.7 and 2 Ma, followed by several Plio-Pleistocene species of *Syncerus* (Bibi, 2009; Gentry, 2010). Therefore, a gap of 3–4 Ma exists in the fossil record of Bubalina. However, a simple biogeographic explanation can be proposed if we accept that several Miocene fossils of *Pachyportax*† can be related to crown Bubalina. Indeed, several fossils of *Pachyportax*† were described in the late Miocene of eastern Africa, Arabia, South Asia, and Southeast Asia (Bibi, 2007, 2009; Khan et al., 2009). The age and distribution of these fossils suggest that a dispersal event occurred during the late Miocene, from South or Southeast Asia into Africa. In Arabia, tropical savannah and grasslands replaced tropical woodlands during the Tortonian, between 11.6 and 7.2 Ma (Pound et al., 2012). Due to these favorable conditions, the ancestor of the *Syncerus* lineage may have migrated from Asia to Africa around 8–9 Ma.

Our analyses suggest that the last common ancestor of extant Bovina underwent a rapid diversification into at least four different lineages during the middle Pliocene, i.e., between 3.3 ± 0.9 Ma and 3.8 ± 1.1 Ma (Table 2). Interestingly, several extinct genera of Bovina first appeared during the Pliocene, such as *Pelorovis*† in Africa and *Leptobos*† in Africa and Europe: the first genus has been linked to *Bos taurus*, the second to bison species (Martinez-Navarro et al., 2007). In addition, fossil remains assigned to *Bison sivalensis*† were found in the Pliocene of the Indian subcontinent (Khan et al., 2010). These data suggest that the genus *Bos* sensu lato (i.e., including *Bison*) was already represented by different morphological lineages during the Pliocene in Africa, Europe, and Asia. At that time, Arabia and a large part of North Africa were covered with temperate and tropical xerophytic shrublands and grasslands (Salzmann et al., 2011). Simulations have also indicated that there was a northward expansion of temperate forests and grasslands in Eurasia, with vast tracts of grassland in Siberia (Haywood et al., 2009). We suggest therefore that these favorable conditions may have triggered the radiation and dispersal of Bovina across a huge area comprising North Africa, southern Europe, West Asia, Central Asia, and Siberia.

Our nuclear data support a close relationship between bison and yak. To date, fossils related to the yak have been found in the late Pleistocene of Siberia, Tibet, and Nepal (Leslie and Schaller, 2009), whereas fossils assigned to *Bison* are much older, with *Bison sivalensis*† from the late Pliocene of the Upper Siwaliks (3.3–2.6 Ma; Khan et al., 2010). *Bison* appeared in Eastern Europe at 1.8 Ma, and was present in the Middle East around 1.6–1.2 Ma, and thereafter in Western Europe around 1.5 Ma (Martinez-Navarro et al., 2007). *Bison* entered North America around 240–220 ka, and then rapidly spread across the continent, where it diversified into several species (Scott, 2010). The paleontological data suggest therefore that the common ancestor of yak and bison emerged somewhere in Asia, and possibly during the Pliocene. Based on the discovery of a Himalayan woolly rhino in the Pliocene (dated to 3.7 Ma), Deng et al. (2011) have proposed that some Ice Age megaherbivores, such as the woolly rhino and yak, first evolved in the Tibetan Plateau before the beginning of the Ice Age. Such a scenario is compatible with our molecular estimates, since the bison + yak lineage separated from other Bovina between 3.7 ± 1.1 Ma and 2 ± 0.6 Ma (Table 2). Subsequently, yak remained in East Asia, while bison spread in the Siwaliks, Middle East and Europe, and crossed into North America through the Bering land bridge during the middle Pleistocene.

5. Conclusions

The markers developed in this study have proven to be very powerful for resolving the systematic relationships within the tribe

Bovini, a difficult and recently diversified group for which contradicting results have been published, due to both mtDNA introgression and absence of robust nuclear signals. We found that the nucleotide distances calculated on the concatenated nuclear dataset can be used to distinguish different taxonomic levels, such as species, genus, subtribe, tribe and subfamily. Such an approach combining multiple independent autosomal introns is anticipated to be highly promising for studying problematic taxonomic groups that diversified during the Plio-Pleistocene epoch.

Acknowledgments

We are very grateful to people who provided tissue samples: Jean-Luc Berthier, Jean-Marie Carenton, Norin Chai, Philippe Chardonnet, Gerard Dousseau, Bernard Dutrillaux, Sylvie Laidebeure, Alexis Lécu, Jean-Francois Marjarie, Sibyle Moulin, Eric Pellé, Claire Rejaud, and Jacques Rigoulet. We thank Nicolas Puillandre for the photo of American bison and William G. Robichaud for the photo of saola. We also acknowledge Andrew T. Knight, Link Olson, and the two anonymous reviewers for their helpful comments on the manuscript. J.A. was supported by the Korea Research Foundation (Grant KRF-2007-357-C00079). This work was supported by the MNHN, CNRS, and "*Consortium National de Recherche en Génomique*". It is part of Agreement No. 2005/67 between the Genoscope and the MNHN on the projects "Macrophylogeny of Life" and "SpeedID".

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.11.003.

References

Alves, P.C., Melo-Ferreira, J., Freitas, H., Boursot, P., 2008. The ubiquitous mountain hare mitochondria: multiple introgressive hybridization in hares, genus *Lepus*. Philos. Trans. R. Soc. B 363, 2831–2839.

Bibi, F., 2007. Origin, paleoecology, and paleobiogeography of early Bovini. Palaeogeogr. Palaeoclimatol. Palaeoecol. 248, 60–72.

Bibi, F., 2009. Evolution, Systematics, and Paleoecology of Bovinae (Mammalia: Artiodactyla) from the Late Miocene to the Recent. PhD Dissertation, Yale University.

Bohlken, H., 1958. Vergleichende Untersuchungen an Wildrindern (Tribus Bovini, Simpson 1945). Zool. Jahrb. Allg. Zool. 68, 113–202.

Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Res 27, 1767–1780.

Brown, W.M., George Jr., M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76, 1967–1971.

Buntjer, J.B., Otsen, M., Nijman, I.J., Kuiper, M.T., Lenstra, J.A., 2002. Phylogeny of bovine species based on AFLP fingerprinting. Heredity 88, 46–51.

Cabria, M.T., Michaux, J.R., Gómez-Moliner, B.J., Skumatov, D., Maran, T., Fournier, P., de Luzuriaga, J.L., Zardoya, R., 2011. Bayesian analysis of hybridization and introgression between the endangered european mink (*Mustela lutreola*) and the polecat (*Mustela putorius*). Mol. Ecol. 20, 1176–1190.

Deng, T., Wang, X., Fortelius, M., Li, Q., Wang, Y., Tseng, Z.J., Takeuchi, G.T., Saylor, J.E., Säilä, L.K., Xie, G., 2011. Out of Tibet: Pliocene woolly rhino suggests high-plateau origin of Ice Age megaherbivores. Science 333, 1285–1288.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Fischer, A., Prüfer, K., Good, J.M., Halbwax, M., Wiebe, V., André, C., Atencia, R., Mugisha, L., Ptak, S.E., Pääbo, S., 2011. Bonobos fall within the genomic variation of chimpanzees. PLoS One 6, e21605.

Gatesy, J., Arctander, P., 2000. Hidden morphological support for the phylogenetic placement of *Pseudoryx nghetinhensis* with bovine bovids: a combined analysis of gross anatomical evidence and DNA sequences from five genes. Syst. Biol. 49, 515–538.

Gentry, A.W., 2010. Bovidae. In: Werdelin, L., Sanders, W.J. (Eds.), Cenozoic Mammals of Africa. University of California Press, Berkeley, pp. 741–796.

Geraads, D., 1992. Phylogenetic analysis of the tribe Bovini (Mammalia, Artiodactyla). Zool. J. Linn. Soc. – Lond. 104, 193–207.

Groves, C.P., 1981. Systematic relationships in the Bovini (Artiodactyla, Bovidae). Z. Zool. Syst. Evol. 19, 264–278.

Hassanin, A., Douzery, E.J., 1999a. Evolutionary affinities of the enigmatic saola (*Pseudoryx nghetinhensis*) in the context of the molecular phylogeny of Bovidae. Proc. R. Soc. B – Biol. Sci. 266, 893–900.

Hassanin, A., Douzery, E.J., 1999b. The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome *b* gene. Mol. Phylogenet. Evol. 13, 227–243.

Hassanin, A., Douzery, E.J., 2003. Molecular and morphological phylogenies of Ruminantia and the alternative position of the Moschidae. Syst. Biol. 52, 206–228.

Hassanin, A., Ropiquet, A., 2004. Molecular phylogeny of the tribe Bovini (Bovidae, Bovinae) and the taxonomic status of the Kouprey, *Bos sauveli* Urbain 1937. Mol. Phylogenet. Evol. 33, 896–907.

Hassanin, A., Ropiquet, A., 2007. Resolving a zoological mystery: the kouprey is a real species. Proc. R. Soc. B – Biol. Sci. 274, 2849–2855.

Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., Jansen van Vuuren, B., Matthee, C., Ruiz-Garcia, M., Catzeflis, F., Areskoug, V., Nguyen, T.T., Couloux, A., 2012. Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. C. R. Biol. 335, 32–50.

Haywood, A.M., Chandler, M.A., Valdes, P.J., Salzmann, U., Lunt, D.J., Dowsett, H.J., 2009. Comparison of mid-Pliocene climate predictions produced by the HadAM3 and GCMAM3 General Circulation Models. Global Planet. Change 66, 208–224.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

IUCN, 2011. IUCN Red List of Threatened Species. Version 2011.2. <http://www.iucnredlist.org> (downloaded 24.04.12).

Khan, M.A., Akhtar, M., 2011. *Proamphibos kashmiricus* (Bovini: Bovidae: Mammalia) lower case from the Pinjor formation of Pakistan. Pak. J. Zool. 43, 615–621.

Khan, M.A., Iqbal, M., Ghaffar, A., Akhtar, M., 2009. *Proamphibos* (Bovini, Bovidae, Mammalia) from the Tatrot formation in the Upper Siwaliks of Pakistan. J. Anim. Plant Sci. 19, 104–107.

Khan, M.A., Kostopoulos, D.S., Akhtar, M., Nazir, M., 2010. Bison remains from the Upper Siwaliks of Pakistan. Neues Jahrb. Geol. P.-A. 258, 121–128.

Kocot, K.M., Cannon, J.T., Todt, C., Citarella, M.R., Kohn, A.B., Meyer, A., Santos, S.R., Schander, C., Moroz, L.L., Lieb, B., Halanych, K.M., 2011. Phylogenomics reveals deep molluscan relationships. Nature 477, 452–456.

Kuroki, Y., Toyoda, A., Noguchi, H., Taylor, T.D., Itoh, T., Kim, D.S., Kim, D.W., Choi, S.H., Kim, I.C., Choi, H.H., Kim, Y.S., Satta, Y., Saitou, N., Yamada, T., Morishita, S., Hattori, M., Sakaki, Y., Park, H.S., Fujiyama, A., 2006. Comparative analysis of chimpanzee and human Y chromosomes unveils complex evolutionary pathway. Nat. Genet. 38, 158–167.

Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288.

Leslie Jr., D.M., Schaller, G.B., 2009. *Bos grunniens* and *Bos mutus* (Artiodactyla: Bovidae). Mamm. Species 836, 1–17.

MacEachern, S., McEwan, J., Goddard, M., 2009. Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (Bovidae, Bovinae). BMC Genomics 10, 177.

Martinez-Navarro, B., Perez-Claros, J.A., Palombo, M.R., Lorenzo, R., Palmqvist, P., 2007. The Olduvai buffalo *Pelorovis* and the origin of *Bos*. Quatern. Res. 68, 220–226.

Martinez-Navarro, B., Ros-Montoya, S., Patrocinio, E.M., Palmqvist, P., 2011. Presence of the Asian origin Bovini, *Hemibos* sp. aff. *Hemibos gracilis* and *Bison* sp., at the early Pleistocene site of Venta Micena (Orce, Spain). Quatern. Int. 243, 54–60.

Matthee, C.A., Eick, G., Willows-Munro, S., Montgelard, C., Pardini, A.T., Robinson, T.J., 2007. Indel evolution of mammalian introns and the utility of non-coding nuclear markers in eutherian phylogenetics. Mol. Phylogenet. Evol. 42, 827–837.

Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., O'Brien, S.J., 2001. Molecular phylogenetics and the origins of placental mammals. Nature 409, 614–618.

Nesi, N., Nakouné, E., Cruaud, C., Hassanin, A., 2011. DNA barcoding of African fruit bats (Mammalia, Pteropodidae). The mitochondrial genome does not provide a reliable discrimination between *Epomophorus gambianus* and *Micropteropus pusillus*. C. R. Biol. 334, 544–554.

Nijman, I.J., van Boxtel, D.C.J., van Cann, L.M., Marnoch, Y., Cuppen, E., Lenstra, J.A., 2008. Phylogeny of Y chromosomes from bovine species. Cladistics 24, 723–726.

Olsen, S.J., 1990. Fossil ancestry of the yak, its cultural significance, and domestication in Tibet. Proc. Acad. Natl. Sci. Phila. 142, 73–100.

Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.

Pound, M.J., Haywood, A.M., Salzmann, U., Riding, J.B., 2012. Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33 Ma). Earth-Sci. Rev. 112, 1–22.

Ranwez, V., Delsuc, F., Ranwez, S., Belkhir, K., Tilak, M.K., Douzery, E.J., 2007. OrthoMaM: a database of orthologous genomic markers for placental mammal phylogenetics. BMC Evol. Biol. 7, 241.

Regier, J.C., Shultz, J.W., Zwick, A., Hussey, A., Ball, B., Wetzer, R., Martin, J.W., Cunningham, C.W., 2010. Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature 463, 1079–1083.

Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L., Maricic, T., Good, J.M., Marques-Bonet, T., Alkan, C., Fu, Q., Mallick, S., Li, H., Meyer, M., Eichler, E.E., Stoneking, M., Richards, M., Talamo, S., Shunkov, M.V., Derevianko, A.P., Hublin, J.J., Kelso, J., Slatkin, M., Pääbo, S., 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468, 1053–1060.

Roos, C., Zinner, D., Kubatko, L.S., Schwarz, C., Yang, M., Meyer, D., Nash, S.D., Xing, J., Batzer, M.A., Brameier, M., Leendertz, F.H., Ziegler, T., Perwitasari-Farajallah, D., Nadler, T., Walter, L., Osterholz, M., 2011. Nuclear versus mitochondrial DNA: evidence for hybridization in colobine monkeys. BMC Evol. Biol. 11, 77.

Ropiquet, A., Hassanin, A., 2005a. Molecular evidence for the polyphyly of the genus *Hemitragus* (Mammalia, Bovidae). Mol. Phylogenet. Evol. 36, 154–168.

Ropiquet, A., Hassanin, A., 2005b. Molecular phylogeny of Caprines (Bovidae, Antilopinae): the question of their origin and diversification during the Miocene. J. Zool. Syst. Evol. Res. 43, 49–60.

Ropiquet, A., Hassanin, A., 2006. Hybrid origin of the Pliocene ancestor of wild goats. Mol. Phylogenet. Evol. 41, 395–404.

Ropiquet, A., Li, B., Hassanin, A., 2009. SuperTRI: a new approach based on branch support analyses of multiple independent data sets for assessing reliability of phylogenetic inferences. C. R. Biol. 332, 832–847.

Saccone, C., Lanave, C., De Grassi, A., 2006. Metazoan OXPHOS gene families: evolutionary forces at the level of mitochondrial and nuclear genomes. Biochim. Biophys. Acta 1757, 1171–1178.

Salzmann, U., Williams, M., Haywood, A.M., Johnson, A.L.A., Kender, S., Zalasiewicz, J., 2011. Climate and environment of a Pliocene warm world. Palaeogeogr. Palaeoclimatol. Palaeoecol. 309, 1–8.

Scott, E., 2010. Extinctions, scenarios, and assumptions: changes in latest Pleistocene large herbivore abundance and distribution in western North America. Quatern. Int. 217, 225–239.

Sepulchre, P., Jolly, D., Ducrocq, S., Chaimanee, Y., Jaeger, J.J., Raillard, A., 2010. Mid-tertiary paleoenvironments in Thailand: pollen evidence. Clim. Past 6, 461–473.

Springer, M.S., Meredith, R.W., Janecka, J.E., Murphy, W.J., 2011. The historical biogeography of Mammalia. Philos. Trans. R. Soc. B 366, 2478–2502.

Swofford, D.L., 2003 PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0b. Sinauer Associates, Sunderland, MA.

Tosi, A.J., Detwiler, K.M., Disotell, T.R., 2005. X-chromosomal window into the evolutionary history of the guenons (Primates: Cercopithecini). Mol. Phylogenet. Evol. 36, 58–66.

Van Gelder, R.G., 1977. Mammalian hybrids and generic limits. Am. Mus. Novit. 2635, 1–25.

Verkaar, E.L., Nijman, I.J., Beeke, M., Hanekamp, E., Lenstra, J.A., 2004. Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? Mol. Biol. Evol. 21, 1165–1170.

Vrba, E., Schaller, G.B., 2000. Antelopes, Deer, and Relatives. Fossil Record, Behavioral Ecology, Systematics and Conservation. Yale University Press, New Haven and London, 341 p.

Yang, Z., Rannala, B., 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23, 212–226.