TÜBİTAK

Turkish Journal of Zoology

http://journals.tubitak.gov.tr/zoology/

Research Article

Turk J Zool
(2016) 40:
© TÜBİTAK
doi:10.3906/zoo-1501-57

Phylogenetic relationships of the vulnerable wild cattle, Malayan gaur (*Bos gaurus hubbacki*), and its hybrid, the selembu, based on maternal markers

Muhammad Khairul Anwar ROSLI[1], Syed Mohamad Fahmi SYED-SHABTHAR[1], Jeffrine Japning ROVIE-RYAN[2], Ramlah ZAINUDDIN[3], Md Nor SHUKOR[1], Salmah YAAKOP[1], Badrul Munir MD-ZAIN[1,*]

[1]School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

[2]Department of Wildlife and National Parks (PERHILITAN), Peninsular Malaysia, Kuala Lumpur, Malaysia

[3]Molecular Ecology Laboratory, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

Received: 29.01.2015 • **Accepted/Published Online:** 30.09.2015 • **Final Version:** 00.00.2016

Abstract: The gaur (*Bos gaurus*) is one of the two extant wild cattle species that can be found in several Asian countries. This species is threatened by extinction due to declining wild populations. Selembu is the name of the Malayan gaur × domestic cattle hybrid. We planned this study to determine the position of the Malayan gaur and its hybrid, the selembu, in the phylogenetics of the genus *Bos* (*Bos gaurus*, *Bos javanicus*, *Bos indicus*, and *Bos taurus*). The mitochondrial 12S rRNA gene and the control region (D-loop) were sequenced in 43 *Bos* samples. Sequences from one water buffalo (*Bubalus*) were used as an outgroup. Phylogenetic trees were reconstructed using neighbor-joining and maximum parsimony in PAUP 4.0b10 and Bayesian inference in MrBayes 3.1. All tree topologies indicated that the Malayan gaur belongs to its own monophyletic clade that is distinct from other species of the genus *Bos.* Selembu samples were grouped in zebu and/or taurine cattle clades. The results also indicated that there are significant embranchment differences in the tree topologies between wild (Malayan gaur and banteng/Bali cattle) and domestic (taurine cattle, zebu cattle, and selembu) cattle. The results showed the complete maternal inheritance situation among the studied samples of all cattle species.

Key words: *Bos gaurus hubbacki*, gaur, seladang, selembu, wild cattle, hybrid

1. Introduction

Cattle species in the genus *Bos* can be subdivided into wild and domestic types. Wild cattle comprise *Bos gaurus* (gaur) and *Bos javanicus* (banteng), while domestic cattle comprise *Bos indicus* (zebu cattle) and *Bos taurus* (taurine cattle) (Syed-Shabthar et al., 2013). In Malaysia, all four species can be found throughout the Malay Peninsula and Borneo. Wild Malayan gaurs (*Bos gaurus hubbacki*) can only be found in Peninsular Malaysia and wild banteng (*Bos javanicus lowi*) can only be found in Sabah, a part of Borneo (Medway, 1983; Wilson and Reeder, 2005). Both species of domestic cattle can be found in the Malay Peninsula and Borneo. The selembu is a hybrid cattle bred in Malaysia by crossbreeding wild Malayan gaur with domestic cattle (zebu and/or taurine) (Mamat-Hamidi et al., 2009). In other countries, gaur has been domesticated and produces fertile offspring called gayal (*Bos frontalis*). The Bali cattle is a domestic form of wild banteng originally bred on Bali Island, Indonesia (Mohamad et al., 2012).

Three subspecies of gaurs are generally recognized: *Bos gaurus gaurus*, found in India, southern Nepal, and Bhutan; *Bos gaurus laosiensis*, distributed in Myanmar, Laos, Vietnam, Thailand, and Cambodia; and *Bos gaurus hubbacki*, which exists only in Peninsular Malaysia and southern Thailand (Duckworth et al., 2008). The Malayan gaur, locally known as 'seladang', can be found in several states, including Pahang, Kelantan, Kedah, Perak, and Terengganu (Yusof, 1981). According to Sahir (2001), there are about 500 remaining individuals in Malaysia. Wild populations of gaur have declined significantly in this country (Conry, 1989). It has been declared as a "Totally Protected" animal under Wildlife Protection Act 76/72, Schedule I (Wild Animals) by the Malaysian government. Since 1982, in situ conservation efforts for the Malayan gaur by the Department of Wildlife and National Parks have included steps to prevent its extinction in response to the growing concern that survival in the wild may be threatened by severe habitat reduction (Sahir, 2001).

* Correspondence: abgbadd@ukm.edu.my

Malayan gaurs share many morphological characteristics with the other two subspecies. According to Groves (2003), there are multivariate overlaps in the analysis of skull and horn measurements between Indian and Southeast Asian subspecies. Southeast Asian subspecies are much bigger, with relatively shorter nasal bones, a narrower occiput, and a narrower horn span. Compared to the Indian gaur, the ascending branch of the premaxilla of the Southeast Asian subspecies generally does not reach the nasal area, as it does in the Indian subspecies (Duckworth et al., 2008). Some of the shared characteristics of gaurs include being sexually dimorphic, a muscular-like bump on the male shoulder with less muscles formed on females, white stockings, gray-brown hair on the forehead and between the horns, an average body size of about 2.5 to 4 m, and a weight of around 700 to 1500 kg (Medway, 1983).

According to Mamat-Hamidi et al., (2009), selembu history began in early April 1983 when Malaysia first saw the birth of a calf from a dairy cow of the Heifer breed (Sahiwan-Friesian) and a wild Malayan gaur bull that had gotten lost on a dairy farm owned by the Veterinary Services of Malaysia in Kluang, Johor. The selembu F1 hybrid shows morphological characteristics as follows: it has a large body size and strong muscular legs, and a higher growth rate in every stage of growth than the mother cows. However, it is hard to find the selembu since it is the least popular cross-breed and the least popular of breeders' choices due to fertility issues in Malaysia. Because of the hybrid's high body mass production and vulnerability, a selembu breeding and production project is currently ongoing at the Agro-Biotechnology Institute Malaysia, MOSTI and Jukimas farm, Selangor.

Phylogenetic trees from earlier reports were constructed based on morphological characteristics of fossils and extant taxa (Groves, 1981) and further reclassified using molecular data from several different techniques. To date, studies on the relationship of species in the genus *Bos* have been conducted by researchers around the world using several types of molecular data and techniques: a) mitochondrial DNA (mtDNA) sequence data (Matthee and Davis, 2001; Cai et al., 2007; Gu et al., 2007; Ginja et al., 2010), b) nuclear DNA sequence data and microsatellites (Kikkawa et al., 2003; MacEachern et al., 2009), and c) DNA fingerprinting and amplified fragment length polymorphism (AFLP) techniques (Buntjer et al., 2002; Vasil'ev et al., 2002). Phylogenetically, the gaur belongs to the tribe Bovini and is further assigned to the subtribe Bovina, which includes some other extant species such as *B. javanicus* (banteng), *B. taurus* (taurine cattle), *B. indicus* (zebu cattle), *B. grunniens* (yak), *Bison bison* (American bison), *Bison bonasus* (European bison), and other *Bubalus* species (Hassanin et al., 2004, 2006). On the other hand, the gayal was previously grouped in the domestic cattle clade (Li et al., 2008), the gaur clade (Ma et al., 2007), or a monophyletic clade of its own in the wild cattle clade (Verkaar et al., 2004). The introduction of molecular systematic approaches and genetic analysis algorithms reduced the anomalies and uncertainties of relationships between taxa in the tribe Bovini (Hassanin et al., 2004). According to Buntjer et al., (2002) the only consistent outcome of comparisons of morphological or molecular characters is the early branching of buffalo-like species (*Bubalus* and *Syncerus*). These two genera were reported to produce fertile hybrid offspring (Wall et al., 1992; Lenstra and Bradley, 1999) and consistently resolve themselves as dichotomous groups (MacEachern et al., 2009). The phylogenetic relationships of the other species of the tribe Bovini are still being debated.

The mitochondrial genome has been extensively used to amplify many genes of interest for phylogenetic studies (Lim et al., 2010; Md-Zain et al., 2010; Ang et al., 2011; Vun et al., 2011). Sequence divergence accumulates more rapidly in mtDNA than nuclear DNA due to a faster mutation rate and the low efficiency of the repair system in mtDNA (compared to nuclear DNA), often leading to high levels of informative variation (Ang et al., 2011). However, different rates of evolution exist within mtDNA regions and these differences suggest different viewpoints for phylogenetic analysis (Simon et al., 1996). Within the ribosomal sequence, the 12S rRNA gene is considered a highly conserved region in mtDNA (Hillis et al., 1996). Conserved sequences are usually applied to illustrate the phylogeny of higher categorical levels, rather than at species level. The D-loop is a noncoding control region that is highly variable and often used for phylogenetic studies including cattle biogeographical analysis (Kim et al., 2003; Cai et al., 2011). To date, there is no phylogenetic study that includes all wild and domestic cattle species (gaur, banteng, taurine cattle, zebu cattle, Bali cattle, and selembu). This study was conducted to portray the phylogenetic relationships of the gaur and its hybrid, the selembu, along with other cattle species based on two mitochondrial DNA regions, the 12S rRNA gene and the control region (D-loop).

2. Materials and methods

2.1. Samples and DNA extraction

For this research, 43 cattle genetic samples were assessed for DNA isolation (Table 1). Several authorized institutions, including the Department of Wildlife and National Parks (PERHILITAN), the National Veterinary Biotechnology Institute (IBVK), the Agro-Biotechnology Institute (ABI) MOSTI, Taiping Zoo, and Zoo Melaka, provided the genetic samples used in this study. Samples received were in the form of blood, tissue, and fecal matter. All samples were extracted and stored in a –20 °C freezer. Total

Table 1. List of samples used in the study.

No.	Species	Domestic name	Code	Phylogeny code
1	*Bos gaurus*	Seladang	Seroja	Gaur C
2	*Bos gaurus*	Seladang	Sarum	Gaur I
3	*Bos gaurus*	Seladang	Damak	Gaur K
4	*Bos gaurus*	Seladang	Kakak	Gaur M
5	*Bos gaurus*	Seladang	BGH 8	Gaur W
6	*Bos gaurus*	Seladang	BGH 17	Gaur Y
7	*Bos javanicus*	Banteng	BJM 1	Gaur 1
8	*Bos javanicus*	Banteng	BJM 2	Banteng 2
9	*Bos javanicus*	Banteng	BJM 4	Banteng 4
10	*Bos javanicus*	Banteng	BJM 5	Banteng 5
11	*Bos javanicus*	Bali cattle	24474	Bali A
12	*Bos javanicus*	Bali cattle	111208	Bali B
13	*Bos javanicus*	Bali cattle	16639	Bali C
14	*Bos javanicus*	Bali cattle	2874	Bali D
15	*Bos javanicus*	Bali cattle	10018	Bali E
16		Selembu	P3	Selembu A
17		Selembu	P4	Selembu B
18		Selembu	P5	Selembu D
19		Selembu	P7	Selembu E
20	*Bos taurus*	Droughtmaster	Droughtmaster	Dom Tau J2
21	*Bos taurus*	Drankenberger	Drankenberger J4	Dom Tau J4
22	*Bos taurus*	Limousin	Limousin	Dom Tau J11
23	*Bos taurus*	Jersi	Jersey	Dom Tau J13
24	*Bos indicus*	Kedah-Kelantan	Mafriwal	Dom Zeb J6
25	*Bos indicus*	Fresian	Freisian	Dom Zeb J9
26	*Bos indicus*	Santa Getrudis	Santa Getrudis	Dom Zeb J10
27	*Bos indicus*	Kedah-Kelantan	KK 3	Dom Zeb KK3
28	*Bos indicus*	Kedah-Kelantan	KK 4	Dom Zeb KK4
29	*Bos indicus*	Kedah-Kelantan	KK 46	Dom Zeb KK 46
30	*Bubalus bubalis*	Water buffalo	-	Water buffalo

genomic DNA was extracted using a standard extraction kit and protocol provided by the QIAGEN DNeasy Blood and Tissue Kit and QIAGEN QIAamp Stool Mini Kit.

2.2. DNA amplification and sequencing

A polymerase chain reaction (PCR) was performed using 25 µL of reaction mixture containing 10–15 ng of genomic DNA, 2.5 µL 10X PCR buffer, 2.0 mM $MgCl_2$, 0.2 mM dNTP mix, 0.6 µM of each primer, and 4 U of Taq DNA polymerase (Vivantis Sdn. Bhd.) in a PTC-100 Thermal Cycler (MJ Research Inc.). The partial 12S rRNA gene fragment of approximately 450 bp was amplified using mammal universal forward primer L1091

(5'-CTGGGATTAGATACCCCACTAT-3') and reverse primer H1478 (5'-GAGGGTGACGGGCGGTGTGT-3') (Kocher et al., 1989). The partial D-loop fragment of approximately 859 bp was amplified using specifically designed primers for cattle, namely Walid_F (5'-TCACCGTCAACTCCCAAAGCTGA–3') and Walid_R (5'-AGGGGGAAGTTTTATGGAAGGGGG-3'). PCR conditions for both genes were as follows: 4 min of denaturation at 94 °C, followed by 30 cycles of 30 s at 94 °C, 30 s at 58 °C, and 1 min at 72 °C, with a final extension of 7 min at 72 °C before cooling to 4 °C for 10 min. DNA from PCR products was purified using the Vivantis G-F1 PCR Clean-up Kit and was sent immediately to the sequencing service company, First Base Sdn. Bhd. (Malaysia), for sequencing.

2.3. Phylogenetic analysis

Sequencing results were exported as FASTA sequence files. The 12S rRNA gene and D-loop region sequences of studied samples were aligned using the ClustalW multiple alignment algorithm of BioEdit, together with a water buffalo (*Bubalus bubalis*) DNA sequence obtained from GenBank as an outgroup. All sequences were analyzed using MEGA 4.0, PAUP 4.0b10, and MrBayes 3.1 for phylogeny reconstruction. Genetic distance and sequence polymorphisms were analyzed using MEGA 4.0. Two methods of analysis in PAUP included neighbor-joining (NJ) with Kimura's 2-parameter method (Pevsner, 2009), which takes into account the unequal rates of evolution of transition and transversion but assumes an equal distribution of nucleotide composition, and maximum parsimony (MP) with stepwise addition and 1000 replicates in heuristic search (Swofford, 2002) and 50% majority rule consensus. In MP, gaps were treated as missing data and equal weighting was given for transitions and transversions, with heuristic search with TBR branch-swapping algorithm. All of the trees were subjected to bootstrap resampling with 1000 replicates to get bootstrap value support. Modeltest 3.7 (Posada and Crandall, 1998) was used to choose the substitution model that best fit the data using the AIC criterion. The best suggested model was subsequently used for Bayesian analysis in MrBayes 3.1.

3. Results

All 43 samples were extracted, amplified, and sequenced. Each sequence was tested for statistically significant DNA matching with sequences in an online database using the Basic Local Alignment Search Tool (BLAST) prior to sequence analysis. All the samples matched the genus *Bos* DNA sequences in the database. Sequences were aligned together with the outgroup sequence according to the respective gene for further analysis.

3.1. Sequence polymorphism

Approximately 450 bp sequences of 12S rRNA were aligned and a total of 379 bp sequences were used for further analysis. Sequences were analyzed for 342 constant and 37 variable characters. From these 37 variable characters, 15 were parsimony-uninformative and 22 were parsimony-informative. The parsimony-informative characters represent 5.8% of the sequences and these sequences contain 10% variable characters with a transition/transversion ratio (Ti/Tv) of 14.23.

For the D-loop region, approximately 859 bp were sequenced and aligned. Final 694 bp sequences were used for further analysis. From these characters, 531 were constant and 163 were variable. The polymorphic characters for both loci are shown in Tables 2 and 3. A total of 120 variable characters were parsimony-informative, representing 17.29% of the total sequence length. The Ti/Tv ratio was recorded to be 16.0. These sequence polymorphisms for both loci are summarized in Table 4, including additional information about the sequences.

Table 2. The polymorphic characters of 12S rRNA.

	1111	1112222223	33
	13791289	9993467890	04
	1282766601	2697175890	21
Gaur_M	ATTCTGAGGA	TCAAGTCCTT	AT
Gaur_I	....	....	..
Gaur_C	....	....	..
Gaur_K	....	....	..
Gaur_W	....	....	..
Gaur_Y	....	....	..
Banteng_1	G..T......	C..GA.....	GC
Banteng_2	G..T......	C..GA.....	GC
Banteng_4	G..T......	C..GA.....	GC
Banteng_5	G..T......	C..GA.....	GC
Bali_A	G..TC.....	C..GA.....	G.
Bali_B	G..T......	C..GA.....	GC
Bali_C	G..T......	C..GA.....	GC
Bali_D	G..T......	C..GA.....	GC
Bali_E	G..TC.....	C..GA.....	G.
Selembu_A	.CCT.A.AAG	.TG..CTTCC	..
Selembu_B	.CCT.A.AAG	.TG..CTTCC	..
Selembu_D	.CCT.A.AAG	.TG...TTCC	..
Selembu_E	.CCT.A.AAG	.TG..CTTCC	..
Dom_Tau_J2	.CCT.A.AAG	.TG..CTTCC	..
Dom_Tau_J11	.CCT.A.AAG	.TG..CTTCC	..
Dom_Tau_J13	.CCT.AGAAG	.TG..CTTCC	..
Dom_Tau_J4	.CCT.A.AAG	.TG..CTTCC	..
Dom_Zeb_J6	.CCT.A.AAG	.TG...TTCC	..
Dom_Zeb_J9	.CCT.A.AAG	.TG...TTCC	..
DDom_Zeb_J10	.CCT.A.AAG	.TG...TTCC	..
Dom_Zeb_KK4	.CCT.A.AAG	.TG...TTCC	..
Dom_Zeb_KK3	.CCT.A.AAG	.TG...TTCC	..
Dom_Zeb_46	.CCT.A.AAG	.TG...TTCC	..
Water_buffalo	..C.CAGACG	C....C.T..	..

Table 3. The polymorphic characters of the D-loop.

		111	1111111111	1111111111	1111111112	2222222222	2222222222	2233333333	3333333333	3333334444	4444444444	4555555566
	1223355566	6788999011	1122333334	4444455555	5556667891	1234445555	6777777888	9900011111	1122222334	5589991222	3344558999	9333567734
	7124508903	5725128523	7909036790	3467801234	5790132622	7053780156	5124568023	2907801456	7903489279	4560599023	8916152123	9147387858
Bali_A	TCGGCCTTCT	GATCCACAGC	TACCCACCTC	AGCATATATT	ATTAAATGAC	CAGTACCCAC	GTATCCCATC	CTACTTCTGA	CACGCTCTCA	CCAACCGATT	TTATTACATT	CTAACCCCTC
Bali_B	C	....	....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Bali_C	C	....	....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Bali_D	....	....	....	....	....	....	....	.....C....	....	T.........	....	..G.......
Bali_E	C	....	.....G....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Banteng_1	C	....	....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Banteng_2	C	....	....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Banteng_4	C	....	....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Banteng_5	C	....	....	....	.......A..	....	....	TC...C....	....	....	....	..G.......
Gaur_C	C..A.T..A.	AC.TTG.GAA	C..T...A.A	.A..CCCC.C	CACGGTCA.T	TGA...T..T	.G.CTTT...	..GTC....G	C	TTGGTT..CC	...C..T..C	TC..T.ATCA
Gaur_I	C..A.T..A.	AC.TTG.GAA	C..T...A.A	.A..CCCC.C	CACGGTCA.T	TGA...T..T	.G.CTTT...	..GTC....G	C	TTGGTT..CC	...C..T..C	TC..T.ATCA
Gaur_K	C..A.T..A.	AC.TTG.GAA	C..T...A.A	.A..CCCC.C	CACGGTCA.T	TGA...T..T	.G.CTTT...	..GTC....G	C	TTGGTT..CC	...C..T..C	TC..T.ATCA
Gaur_M	C..A.T..A.	AC.TTG.GAA	C..T...A.A	.A..CCCC.C	CACGGTCA.T	TGA...T..T	.G.CTTT...	..GTC....G	C	TTGGTT..CC	...C..T..C	TC..T.ATCA
Gaur_W	C..A.T..A.	AC.TTG.GAA	C..T...A.A	.A..CCCC.C	CACGGTCA.T	TGA...T..T	.G.CTTT...	..GTC....G	C	TTGGTT..CC	...C..T..C	TC..T.ATCA
Gaur_Y	C..A.T..A.	AC.TTG.GAA	C..T...A.A	.A..CCCC.C	CACGGTCA.T	TGA...T..T	.G.CTTT...	..GTC....G	C	TTGGTT..CC	...C..T..C	TC..T.ATCA
Dom_Tau_J13	.TAATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	TG.AC..TGT	AATATTTGCT	T.GTC.T...	TGTAT.T.TC	TTGGTTAGCC	CC..CGT.CG	TC.GT.AT..
Dom_Tau_J11	.TAATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	TG.AC..TG.	AATATTTGCT	T.GTC.T...	TGTAT.T.TC	TTGGTTA.CC	CC..CGT.CG	TC.GT.AT..
Dom_Tau_J4	.TAATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	T..ACT.T..	AATATTTGCT	T.GTC.....	TGTAT.T.TC	TTGGTTA.CC	CC.CCGT.CG	TC.GT.AT..
Dom_Tau_J2	.TAATTCCG.	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	TG.AC..TGT	AATATTTGCT	T.GTC.T...	TG.AT.T.TC	TTGGTTA.CC	CC..CGT.CG	TC.GT.AT..
Dom_Zeb_KK4	.T.ATTCCG.	A.C...T...	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	T..AT..T.T	AACATTT..T	T...C.TCA.	...ATC...C	TT.GTTAGCC	....CGTG.G	TC..T.ATC.
Dom_Zeb_KK3	.T.ATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	T..ATT.T.T	AATATTT..T	T...C..CA.	...ATC..TC	TT.GTTAGCC	....CGTG.G	TC..T.ATC.
Dom_Zeb_46	.T.ATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	T..AT..T.T	AACATTT..T	T...C.TCA.	...ATC...C	TT.GTTAGCC	....CGTG.G	TC..T.ATC.
Dom_Zeb_J10	.T.ATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	T..AT..T.T	AACATTT..T	T...C..CA.	...ATC...C	TT.GTTAGCC	C...CGTG.G	TC..T.ATC.
Dom_Zeb_J6	.T.ATTCC..	A.C...T.A.	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	TG.AT..T.T	AACATTT..T	T...C..CA.	...ATC...C	TT.GTTAGCC	....CGTG.G	TC..T.ATC.
Dom_Zeb_J9	.T.ATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	T..AT..T.T	AACATTT.C.	T...C..CA.	...ATC.C.C	TT.GTTAGC.	....CGTG.G	TC..TTATC.
Selembu_A	.TAATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	TG.AC..TG.	AACATTTGCT	T.GTC.T...	TGTAT.T.TC	TTGGTTA.CC	CCG.CGT.CG	TC.GT.AT..
Selembu_B	.TAATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	TG.AC..TG.	AACATTTGCT	T.GTC.T...	TGTAT.T.TC	TTGGTTA.CC	CCG.CGT.CG	TC.GT.AT..
Selembu_E	.TAATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	TCC..TCA.A	TG.AC...G.	AATATTTGCT	T.GTC.T...	TGTAT.T.TC	TTGGTTA.CC	CC..CGT.CG	TC.GT.AT..
Selembu_D	.T.ATTCC..	A.C...T...	CTAAT.GACA	GATGAT..CC	CGC..TCAGA	T..AT..T.T	AACATTT...	T...C..CA.	...ATC.C.C	TT.GTTAGC.	....CGTG.G	TC..TTATC.
Water_buffalo	CT.CT..G.C	ACCT.TT.A.	C....GTAAA	TAGG.T..CC	..CG.CC...	TG.A..G..T	.ACATTTG..	.AG.C.TCAT	........TC	T.GGTT.TA.	...CA.TG.C	TC.TT.ATCA

Table 4. Information of sequence characteristics of 12S rRNA and D-loop.

Description	Region/value	
	12S	D-loop
Total characters	379	694
Conserved characters	342	531
Variable characters	37	163
Singleton characters	15	43
Parsimony-informative characters	22	120
Percentage of parsimony-informative characters	5.80%	17.29%
Bias of Ti/Tv (R)	25.43	4.74
Frequency of base T	23.10	28.00
Frequency of base C	20.50	24.40
Frequency of base A	38.40	32.70
Frequency of base G	18.00	14.80

3.2. Phylogenetic reconstruction

3.2.1. 12S rRNA gene

The NJ tree was portrayed using the Kimura 2-parameter algorithm with 1000 bootstrap replications. Cattle samples were grouped together in a major cattle clade distinct from the outgroup water buffalo (Figure 1). The Malayan gaur, banteng, and Bali cattle were grouped in Clade A while zebu, taurine cattle, and selembu were in Clade B. In Clade A, the Malayan gaur formed its own monophyletic clade distinct from the other mixed clade of banteng and Bali cattle. Two sister clades formed in Clade B, namely the zebu clade and taurine cattle clade. The selembu samples were observed mixed in zebu and taurine cattle clades. The domestic zebu cattle clade comprised Kedah-Kelantan, Friesian, Santa Getrudis, Local Indian Diary, Mafriwal, Charolais, and selembu S4. The domestic taurine cattle clade consisted of the Jersey, Limousin, Drakenberger, Droughtmaster, selembu S1, selembu S2, and selembu S5 samples. Bootstrap support values can be found on each branch of the tree (Figure 1).

The MP tree was analyzed in PAUP 4.0 and the best tree was chosen based on several criteria, as follows: CI = 0.7959, HI = 0.2041, RI = 0.9685, RC = 0.7708, and tree length of 49. The topology observed was fairly similar to the NJ tree, with a slight difference in Clade B, where polytomy occurred between samples of zebu cattle and the monophyletic clade of taurine cattle. However, Clade B was supported by a high bootstrap value (98%).

Bayesian inference was run with two metropolis-coupled Monte Carlo Markov chains simultaneously. A total of 1,000,000 generations were generated and ended

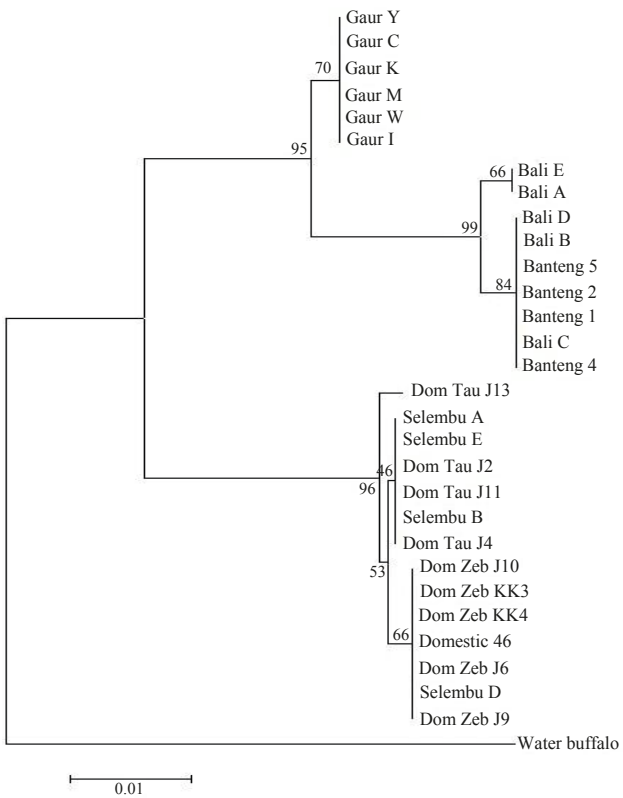

Figure 1. Neighbor-joining tree of the 12S rRNA gene. The numbers at the branches stand for bootstrap values (%) of 1000 replications.

with a standard deviation of split frequencies value of 0.005232. A tree was generated for every 100 generations and the consensus tree topology for 7501 was obtained by eliminating 2500 earliest trees from 10,001 trees (25% Bayesian tree burn-in decided based on our Bayesian history plots). The final consensus tree was similar to the NJ tree topology, with the only differences being probability support values at each branch.

3.2.2. D-loop region

For the D-loop region sequences, the NJ, MP, and Bayesian phylogenetic trees showed the same topologies as 12S rRNA. Tree topologies for both MP and Bayesian trees were not shown as their topologies were the same as that of the NJ tree in both loci. The best MP tree was chosen based on several criteria: CI = 0.7106, HI = 0.2894, RI = 0.9671, RC = 0.6872, and tree length of 311. The best model for ModelTest is GTR+I+G. As with the 12S rRNA, a total of 1,000,000 generations were generated and ended with a standard deviation of split frequencies value of 0.0027. A tree was generated for every 100 generations and the consensus tree topology for 7501 was obtained by eliminating 2500 earliest trees from 10,001 trees (25% Bayesian tree burn-in decided based on our Bayesian history plots). Each monophyletic clade for each tree was supported by high bootstrap values and probability support (above 90% or 0.9).

Cattle samples were grouped in a major clade distinct from the outgroup of water buffalo (Figure 2). In all trees, the gaur, banteng, and Bali cattle were grouped together in one clade while the zebu, taurine, and selembu cattle were grouped together in another clade. The gaur formed its own monophyletic clade distinct from the banteng/Bali cattle clade. Selembu samples were mixed in the domestic cattle clade in both zebu and taurine clades.

3.3. Genetic distance

A pairwise analysis was performed on each region in PAUP 4.0 to get genetic distance values (Table 5). For both regions, results showed that the Malayan gaur is closely related to banteng and Bali cattle. For banteng and Bali cattle, there are less than 1% genetic distance values, indicating that the two species are closely related to each other. The domestic cattle, zebu, and taurine cattle are closer to each other compared to other species. Genetic distance values showed that the selembu is closely related to zebu and taurine cattle rather than to other species.

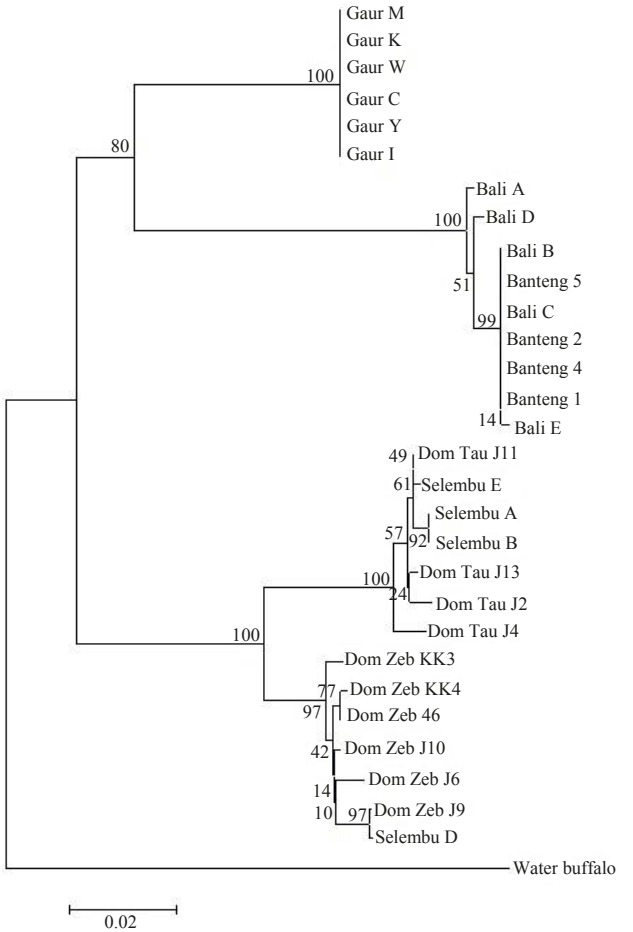

Figure 2. Neighbor-joining tree of the D-loop gene. The numbers at the branches stand for bootstrap values (%) of 1000 replications.

4. Discussion

This study demonstrates that the application of mitochondrial markers is useful in understanding the relationship among cattle species in the genus *Bos*. mtDNA is a haploid molecule and maternally inherited. Therefore, it has one-fourth the effective population size (Ne) of nuclear genes. This makes mtDNA more sensitive than nuclear genes to demographic processes such as population fragmentation and bottlenecks (Dadi et al., 2009). The two mtDNA regions of interest, 12S rRNA and the D-loop, showed two different evolution rates, with the D-loop having a higher mutation rate compared to 12S rRNA. Investigation of variable sites with parsimony-informative characters in this study showed that the 12S rRNA gene is conserved enough and suitable as a tool to identify the relationship of *Bos* species from genus up to species level. The D-loop region was also significant for this study. The D-loop showed that the clade formed is up to population level with more parsimony-informative characters. The D-loop region has been extensively used by other researchers for population genetic studies (Lim et al., 2010; Ang et al., 2012). The bootstrap values and probability for the Bayesian analysis obtained from both regions are strong enough to support each clade in all phylogenetic trees.

Data from the 12S rRNA gene and the D-loop region show a concurrent pattern of genetic distances, which

Table 5. Genetic distance of 12S rRNA and D-loop (italic) based on Kimura 2-parameter.

	Gaur	Banteng	Bali cattle	Selembu	Taurine cattle	Zebu cattle
Banteng	0.0188 / 0.1076					
Bali cattle	0.0188 / 0.1055	0.0021 / 0.0035				
Selembu	0.0377 / 0.1091	0.0521 / 0.1372	0.0521 / 0.1365			
Taurine cattle	0.0391 / 0.1104	0.0536 / 0.1432	0.0536 / 0.1425	0.0013 / 0.0179		
Zebu cattle	0.0355 / 0.1010	0.0499 / 0.1262	0.0499 / 0.1255	0.0020 / 0.0371	0.0033 / 0.0453	
Water buffalo	0.0699 / 0.1498	0.0851 / 0.1913	0.0826 / 0.1890	0.0736 / 0.1672	0.0721 / 0.1719	0.0759 / 0.1552

indicate the genetic divergence between *Bos gaurus hubbacki* and other cattle species. From the genetic distance data, the Malayan gaur was closer to banteng and Bali cattle compared to other species. Tree topologies from different phylogenetic analyses clearly indicated that the Malayan gaur forms its own distinct monophyletic clade. The Malayan gaur has been grouped together with banteng. These topologies were supported by values of genetic distances (average value of 0.1%), with banteng and Bali cattle as its closest descendants. This is true for classifications assigned for both gaur and banteng species as wild cattle with earlier divergence than other cattle species based on tree topologies. In earlier reports, based on molecular data, *Bos gaurus* was grouped with *Bos javanicus* (banteng) in a wild cattle clade (Schreiber et al., 1999; Lai et al., 2006; Stock et al., 2009). This arrangement concurs with research using autosomal genes (MacEachern et al., 2009). Our molecular data further corroborated this taxonomic distinction.

Tree topologies showed that selembu in Malaysia have a mixture of zebu and taurine cattle maternally. There is no significantly close relationship of the selembu with its ancestor, the Malayan gaur. A close relationship between Malayan gaur and selembu cannot be determined only based on maternal bias of the mtDNA. However, earlier reports stated that there are three systematic deposition possibilities of this hybrid form of gaur: 1) gayal will be deposited in an independent clade that has a close relation with gaur (Verkaar et al., 2004), 2) gayal will be deposited in a clade together with gaur (Ma et al., 2007), and/or 3) gayal will be deposited in a clade of either zebu cattle or taurine cattle (Li et al., 2008). Maternal inheritance is concurrent with the history of selembu in Malaysia; a wild Malayan gaur bull got lost on a dairy farm owned by the Veterinary Services of Malaysia in Kluang, Johor, and mated with a domestic cow (Mamat-Hamidi et al., 2009). With the presence of paternal and biparental markers in future studies, systematic relationships between Malayan gaur and its hybrid, the selembu, can be clearly viewed and resolved.

For domestic cattle, *B. indicus* and *B. taurus* were together in a group with distinct separation from each other. Genetic distance and high values of bootstrapping (>97%) and Bayesian posterior probabilities (>0.9) supported the distinction of these two domestic species. These findings are identical to a few earlier reports based on sequences of mitochondrial DNA of cattle species (Bradley et al., 1996; Hassanin et al., 2004, 2006; Verkaar et al., 2004; Ma et al., 2007) concluded that the association of taurine and zebu cattle reflects the fertility of females as well as male hybrid offspring with a divergence time of 100,000 to 200,000 years. Through a domestication event that occurred between 8000 and 10,000 years ago, both species are believed to have originated from the aurochs, *B. primigenius* (Epstein, 1971; MacHugh et al., 1997). However, there appears to be an insertion of domestic Bali cattle in the banteng subclade within the wild cattle group instead of in the domestic cattle group. According to Mohamad et al. (2009), Bali cattle are a domestic type of banteng in which domestication took place around 3500 BC. It is kept on several Indonesian islands, as well as in other countries. This thus explains the matrilineal genetic inheritance of mtDNA of banteng in Bali cattle in our study. Theoretically, the mtDNA of banteng was transferred into its Bali cattle ancestor through a mating between a zebu/taurine male and a banteng female during historic times. Therefore, the grouping of Bali cattle and banteng seems to make sense (Mohamad et al., 2009). Similarly, the molecular phylogeny obtained here is concordant with the morphological phylogeny of domestic cattle, the humpless taurine and the humped zebu, excluding Bali cattle.

The independent distinction of Malayan gaur infers conservation relevance to its decreased population numbers in the wild. Any translocation, reintroduction, or breeding-in-captivity program should consider the best plan ecologically and genetically, in order to help increase these subspecies. Studies of genetic diversity including phylogeography are essential in order to have a better understanding of the relationships among Malayan gaur individuals and between the other two subspecies

of gaurs. These goals can be achieved with the help of different molecular markers (i.e. Y chromosomes and microsatellites) and could provide a more clear view of phylogenetic relationships of species of the genus *Bos*. For a better understanding of the selembu in cattle phylogenetics, several markers such as nuclear DNA (recombinant representative) and Y chromosomes (paternal inheritance) can be used in future studies.

Acknowledgments
We would like to express our sincere appreciation to the Faculty of Science and Technology, UKM. Our special thanks to the Department of Wildlife and National Parks (PERHILITAN) and especially the Director General, Director Ex-situ Conservation Division, Dr Abas Mazni Othman from MARDI, the National Institute of Veterinary Biodiversity (IBVK), the Agro-Biotechnology Institute MOSTI, Zoo Taiping, and Zoo Melaka (banteng fecal samples) for providing genetic samples. This study was funded by research grants STGL-003-2009, Sciencefund 02-01-02-SF0762, INDUSTRI-2011-003, and DLP-2013-006. We also express our gratitude to all who were directly or indirectly involved in this study.

References

Ang A, Srivasthan A, Md-Zain BM, Ismail MRB, Meier R (2012). Low genetic variability in the recovering urban banded leaf monkey population of Singapore. Raffles B Zool 60: 589–594.

Ang KC, Leow JWH, Yeap WK, Hood S, Mahani MC, Md-Zain BM (2011). Phylogenetic relationships of the Orang Asli and Iban of Malaysia based on maternal markers. Genet Mol Res 10: 640–649.

Bradley DG, MacHugh DE, Cunningham P, Loftus RT (1996). Mitochondrial diversity and the origin of African and European cattle. P Natl Acad Sci USA 93: 5131–5135.

Buntjer JB, Otsen M, Nijman IJ, Kuiper MT, Lenstra JA (2002). Phylogeny of bovine species based on AFLP fingerprinting. Heredity 88: 46–51.

Cai X, Chen H, Lei CZ, Wang S, Xue K, Zhang B (2007). MtDNA diversity and genetic lineages of eighteen cattle breeds from *Bos taurus* and *Bos indicus* in China. Genetica 131: 175–183.

Cai X, Gou X, Jiang L (2011). Bioinformatic analysis of phylogenetic evolution of Chinese cattle. In: The 5th International Conference on Bioinformatics and Biomedical Engineering; Wuhan, China. New York, NY, USA: IEEE, pp. 1–3.

Conry PJ (1989). Gaur *Bos gaurus* and development in Malaysia. Biol Conserv 49: 47–65.

Dadi H, Tibbo M, Takahashi Y, Nomura K, Hanada H, Amano A (2009). Variation in mitochondrial DNA and maternal genetic ancestry of Ethiopian cattle populations. Anim Genet 40: 556–559.

Duckworth JW, Steinmetz R, Timmins RJ, Pattanavibool A, Than Z, Do T, Hedges S (2008). *Bos gaurus*. IUCN Red List of Threatened Species. Gland, Switzerland: IUCN.

Epstein H (1971). The Origin of the Domestic Animals of Africa. 1st ed. New York, NY, USA: African Publishing Corporation.

Ginja C, Penedo MC, Melucci L, Quiroz J, Martinez Lopez OR, Revidatti MA, Martinez-Martinez A, Delgado JV, Gama LT (2010). Origins and genetic diversity of New World Creole cattle: inferences from mitochondrial and Y chromosome polymorphisms. Anim Genet 41: 128–141.

Groves C (2003). Taxonomy of ungulates of the Indian Subcontinent. J Bombay Nat Hist Soc 100: 341–362.

Groves CP (1981). Systematic relationships in the Bovini (Artiodactyla, Bovidae). J Zool Syst Evol 19: 264–278.

Gu Z, Zhao X, Li N, Wu C (2007). Complete sequence of the yak (*Bos grunniens*) mitochondrial genome and its evolutionary relationship with other ruminants. Mol Phylogenet Evol 42: 248–255.

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, Cornette R, Tranier M, Pfeffer P, Candegabe P, Lemaire M (2006). Has the kourpey (*Bos sauveli* Ubain, 1937) been domesticated in Cambodia? C R Biol 329: 124–135.

Hillis DM, Moritz C, Mable BK (1996). Molecular Systematics. 2nd ed. Sunderland, MA, USA: Sinauer Associates Inc.

Kikkawa Y, Takada T, Sutopo, Nomura K, Namikawa T, Yonekawa H, Amano T (2003). Phylogenies using mtDNA and SRY provide evidence for male-mediated introgression in Asian domestic cattle. Anim Genet 34: 96–101.

Kim KI, Lee JH, Lee SS, Yang YH (2003). Phylogenetic relationships of Northeast Asian cattle to other cattle populations determined using mitochondrial DNA D-Loop sequence polymorphism. Biochem Genet 41: 91–98.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. P Natl Acad Sci USA 86: 6196–6200.

Lai SJ, Liu YP, Liu Y X, Li XW, Yao YG (2006). Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. Mol Phylogenet Evol 38: 146–154.

Lenstra JA, Bradley DG (1999). Systematics and phylogeny of cattle. In: Fries R, Ruvinsky A, editors. The Genetics of Cattle. Oxford, UK: CABI Publishing, pp. 1–14.

Li SP, Chang H, Ma GL, Cheng HY (2008). Molecular phylogeny of the gayal in Yunnan China inferred from the analysis of cytochrome b gene entire sequences. Asian Austral J Anim 21: 789–793.

Lim LS, Ang KC, Mahani MC, Shahrom AW, Md-Zain BM (2010). Mitochondrial DNA polymorphism and phylogenetic relationships of Proto Malays in Peninsular Malaysia. J Biol Sci 10: 71–83.

Ma G, Chang H, Li S, Chen H, Ji D, Geng R, Chang C, Li Y (2007). Phylogenetic relationships and status quo of colonies for gayal based on analysis of cytochrome *b* gene partial sequences. J Genet Genomics 34: 413–419.

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.

MacHugh DE, Shriver MD, Loftus RT, Cunningham P, Bradley DG (1997). Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). Genetics 146: 1071–1086.

Mamat-Hamidi, Idris I, Hilmi M (2009). Karyotype of Malayan Gaur (*Bos gaurus hubbacki*), Sahiwal-Friesian cattle and Gaur x cattle hybrid backcrosses. Pak J Biol Sci 12: 896–901.

Matthee CA, Davis SK (2001) Molecular insights into the evolution of the family Bovidae: a nuclear DNA perspective. Mol Biol Evol 18: 1220–1230.

Md-Zain BM, Lee SJ, Lakim M, Ampeng A, Mahani MC (2010). Phylogenetic position of *Tarsius bancanus* based on partial cytochrome *b* DNA sequences. J Biol Sci 10: 348–354.

Medway L (1983). Wild mammals of Malaya (Peninsular Malaysia) and Singapore. Kuala Lumpur, Malaysia: Oxford University Press.

Mohamad K, Olsson M, Andersson G, Purwantara B, Van Tol HTA, Rodriguez-Martinez H, Colenbrander B, Lenstra JA (2012). The origin of Indonesian cattle and conservation genetics of the Bali cattle breed. Reprod Dom Anim 47: 18–20.

Mohamad K, Olsson M, Van-Tol HTA, Rodriguez-Martinez H, Colenbrander B, Lenstra JA (2009). On the origin of Indonesian cattle. PLoS One 4: e5490.

Pevsner J (2009). Bioinformatics and Functional Genomics. 2nd ed. Hoboken, NJ, USA: Wiley-Blackwell.

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

Sahir O (2001). Breeding performance of the Malaysian captive gaur. J Wildl Parks 19: 25–32.

Schreiber A, Seibold I, Notzold G, Wink M (1999). Cytochrome *b* gene haplotypes characterize chromosomal lineages of anoa, the Sulawesi dwarf buffalo (Bovidae: *Bubalus* sp.). J Hered 90: 165–176.

Simon C, Nigro L, Sullivan J, Holsinger K, Martin, A, Grapputo A, Franke A, McIntosh C (1996). Large differences in substitutional pattern and evolutionary rate of 12S ribosomal RNA genes. Mol Biol Evol 13: 923–932.

Stock F, Edwards CJ, Bollongino R, Finlay EK, Burger J (2009). Cytochrome *b* sequences of ancient cattle and wild ox support phylogenetic complexity in the ancient and modern bovine population. Anim Genet 40: 694–700.

Swofford DL (2002). Phylogenetic Analysis Using Parsimony (and Other Methods). Version 4.0 Beta. Sunderland, MA, USA: Sinauer Associates.

Syed-Shabthar SMF, Rosli MKA, Mohd-Zin NAA, Romaino SMN, Fazly-Ann ZA, Mahani MC, Abas-Mazni O, Zainuddin R, Yaakop S, Md-Zain BM (2013). The molecular phylogenetic signature of Bali cattle revealed by maternal and paternal markers. Mol Biol Rep 40: 5165–5176.

Vasil'ev VA, Steklenev EP, Morozova EV, Semenova SK (2002). DNA fingerprinting of individual species and intergeneric and interspecific hybrids of genera *Bos* and *Bison*, subfamily Bovinae. Genetika 38: 515–520.

Verkaar ELC, Nijman IJ, Beeke M, Hanekamp E, Lenstra JA (2004). Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? Mol Biol Evol 21: 1165–1170.

Vun VF, Mahani MC, Lakim M, Ampeng A, Md-Zain BM (2011). Phylogenetic relationships of leaf monkeys (*Presbytis*; Colobinae) based on cytochrome *b* and 12S rRNA genes. Genet Mol Res 10: 368–381.

Wall DA, Davis SK, Read BM (1992). Phylogenetic relationships in the subfamily Bovinae (Mammalia: artiodactyla) based on ribosomal DNA. J Mammal 73: 262–275.

Wilson DE, Reeder DAM (2005). Mammal Species of the World. A Taxonomic and Geographic Reference. 3rd ed. Baltimore, MD, USA: Johns Hopkins University Press.

Yusof E (1981). Seladang (*Bos gaurus hubbacki*) in Malaysia. Kuala Lumpur, Malaysia: Department of Wildlife and National Parks.