Genetic diversity and phylogenetic relationship using mitochondrial ND5 partial sequences in three cattle breeds

Aslinda, K.[1,4]*, Wan Somarny, W.M.Z.[1], Mat Tasol, S.[1], Fadzlirahimi, I.[2], Mohd. Hafiz, A.W.[3] and Siti Azizah, M.N.[4,5]

[1]Strategic Livestock Research Centre, MARDI Headquarters, 43000 Serdang, Selangor, [2]Kemaman MARDI Station, P.O.Box 44, 24007 Kemaman, Terengganu,
[3]Muadzam Shah MARDI Station, 26700 Muadzam Shah, Pahang,
[4]School of Biological Sciences, Universiti Sains Malaysia, Penang,
[5]Centre for Marine and Coastal Studies, Universiti Sains Malaysia, 11800 Minden, Penang.
*Corresponding author: linda_3685@yahoo.com

Abstract

Partial sequences of mtDNA ND5 (404 bp) were determined for 107 bovine samples from three cattle breeds: Kedah-Kelantan (KK, n=86), Yellow cattle (YC, n=18) and Banteng (n=3). Seven haplotypes (Hap01-Hap07) were generated with Hap01 having the highest frequency shared by KK and YC populations, whilst Hap03-04, Hap06 and Hap07 were unique to KK cattle from Pahang (KKB), YC and Banteng, respectively. The YC was the most diverse (h= 0.3072, π= 0.003899) than other populations. Among the Kedah-Kelantan cattle populations, KKB (h= 0.1871, π= 0.002438) was the most diverse, while KK cattle from Kelantan (KKA) was the least (h= 0.0690, π= 0.000171). Low genetic distance between KKA and KK cattle from Terengganu (KKC, 0.0002) showed it to be more closely related in maternal lineage than between KKA and KKB cattle (0.0014) and KKA and YC (0.0022). The NJ tree differentiated *B. taurus* from *B. indicus* with 99% bootstrap support indicating distinct evolutionary lineages. *B. indicus* further bifurcated into KKA, KKB and KKC populations and YC populations revealing its maternal relationship with *B. indicus*.

Key words: Kedah-Kelantan cattle, ND5 gene, phylogenetic, genetic diversity

Introduction

Globally, domestic breeds of cattle are mostly originated from two related species: *Bos indicus* (zebu) and *B. taurus* (taurine), both descending from *B. primigenius*. *Bos grunniens* is the origin of the domestic yak, the gayal and wild gaur which were initially classified as *B. frontalis* and *B. gaurus,* respectively, but are now both categorized as *B. frontalis* (domesticated gayal and wild gaur) while the domestic Bali cattle originated from *B. javanicus* (Banteng) (Mohamad *et al*., 2009). Most of these breeds are crossbred cattle which have naturally adapted to the local environment through domestication (Xuan *et al.,* 2010a).

The Kedah-Kelantan (KK) is a native breed of Malaysia through many generations of domestication and is well adapted to the local climate (Devendra, 1975). This domesticated breed is a crossbred of Shorthorn cattle from west Asia and Zebu cattle from India. It is an important breed for beef and hide production (Devendra, 1975). Thus, the KK breed is an invaluable asset to the country and its genetic integrity needs to be conserved to maintain breeding programmes to meet market demand. An important approach to achieve this is through a comprehensive study on the genetic diversity and phylogenetic relationship of the cattle breed as has been done in other regions: in Korea (Kyu-Il *et al.,* 2003), in Portugal (Mateus *et al.,* 2004), in China (Cai *et al.,* 2007) and in Romania (Xuan *et al.,* 2010a, b). The Yellow cattle (YC) and Banteng are indigenous to China and Indonesia and categorized in the *B. indicus* and *B. javanicus* lineages, respectively.

In the past few decades, mitochondrial DNA (mtDNA) has been widely used in phylogenetic studies because of several advantageous characteristics it possesses: maternal inheritance, no recombination and fast rate of evolution (Bailey *et al.*, 1996) which permits investigation of recent history. It is also an efficient genetic marker for population studies in Chinese cattle using mtDNA ND5 gene (Zhang *et al*., 2008), Korean cattle (Kyu-II *et al*., 2003), Asian cattle using mtDNA D-loop sequence (Jia *et al*., 2010) and *B. indicus* and *B. taurus* cattle breeds using the complete mtDNA as marker (Hiendleder *et al*., 2008). The objective of this study was to elucidate the genetic diversity and phylogenetic relationships of three cattle breeds: Kedah-Kelantan (KK), Yellow cattle (YC) and Banteng using mtDNA ND5 partial sequences.

Materials and Methods

Sample collection and extraction

Blood samples of the KK breed were collected from three populations, namely Tanah Merah, Kelantan (KKA, n=29), Tembangau, Pahang (KKB, n=31) and the Kemaman MARDI Station, Terengganu (KKC, n=26). The Yellow cattle samples (n=18) were collected from the herd in Muadzam MARDI Station, Pahang while Banteng (n=3) from animals kept at the National Zoo in Kuala Lumpur. Genomic DNA was obtained using Genomic DNA Purification Kit (Promega, USA). The DNA concentrations were quantified using a Nanodrop Spectrophotometer, ND-1000 (USA) or run under 0.7% agarose gel.

DNA amplification and sequencing

The DNA extract was then amplified using published primers of cattle mtDNA ND5 gene fragment; F-5'CCC AAC GAG GAA AAT ATA CC 3' and R-5'GGA AGA GGT TGT TTG CGG TT 3' (Rogerio *et al.,* 2007). A total of 25 µl Polymerase Chain Reaction (PCR) reaction of 1 X Buffer, 0.6 ρmol primers, 200 µM dNTPs, 2.0 mM $MgCl_2$, 50 ng genomic DNA and 1.5 U Taq DNA polymerase (Promega, USA) was prepared. Amplification was conducted in a Bio-Rad C1000TM Thermal Cycler with initial denaturation at 94 °C, 5 minutes followed by 40 cycles of denaturation at 94 °C, 30 second, annealing at 62 °C, 30 second and extension at 72 °C, 35 second, and a final extension of 72 °C for 10 minutes. The PCR products were purified using the PCR Clean-Up System Kit (Promega, USA) and sent for sequencing to a service provider (Medigene Sdn Bhd).

Statistical Analysis

All sequences were manually confirmed and aligned using Clustal W 1.6 implemented in MEGA 4.0 software (Tamura *et al.*, 2007). The software was also used to calculate the average genetic distance within and between populations. The haplotypes were identified using Collapse 1.2 software (Posada, 2004). The Neighbour-Joining (NJ) phylogenetic tree (Saitou and Nei, 1987) was constructed using Kimura two-parameter model in MEGA with 1000 replicates bootstrap (Felsenstein, 1988). Arlequin 3.11 software (Excoffier *et al*., 2005) was used to calculate the number of haplotypes, number of polymorphic sites, haplotypes and nucleotide diversities, transition and transversions.

Results and Discussion

Final alignment after editing using MEGA 4.0 software (Tamura *et al.*, 2007) was 404 bp of ND5 partial sequence with 134 amino acids revealing seven haplotypes (Hap01-Hap07) by Collapse 1.2 software (Posada, 2004) as presented in Table 1. Hap01 had the highest frequency in all populations except Banteng: KKA (0.966), KKB (0.903), KKC (0.962) and YC (0.833). Hap03-04 and Hap06 occurred at very low frequencies but were unique to KKB (0.0323) and YC (0.0556), respectively. Banteng was monomorphic (1.000) and unique for Hap07. Hap02 also occurred at low frequencies but was shared by KKA (0.0345) and KKC (0.0385) while Hap05 was shared by KKB (0.0323) and YC (0.111).

The generated haplotypes were compared with a reference *B. indicus* mtDNA sequence from GenBank: GI: 33321647. A total of 44 variable sites (10.9% of all sites) with seven parsimony informative sites and 37 singletons were observed. From the 44 variable sites the transition and transversion site ratio was 21:1, similar to that observed in Xuan *et al*. (2010b) but higher transition bias compared to Irwin *et al*. (1991, ratio 10:1). The transition rate of T↔C (pyrimidines) was higher than of A↔G (purines), similar to the finding of Xuan *et al*. (2010a) and Tamura and Nei (1993). There was no insertion or deletion in the aligned sequence. The average nucleotide frequencies of C, T, A and G were 27.3, 31.8, 35.0 and 5.9%, respectively, indicating that the average number of G nucleotide was lower than the rest. The nucleotide composition was richer in A/T nucleotide (66.8%) than C/G (33.2%).

Table 1. The frequencies, number of haplotypes and polymorphic sites, nucleotide and haplotype diversity, transition and transversion in five cattle populations

	Populations				
Parameter	KKA	KKB	KKC	YC	Banteng
Haplotype					
01	0.966	0.903	0.962	0.833	0
02	0.0345	0	0.0385	0	0
03	0	0.0323	0	0	0
04	0	0.0323	0	0	0
05	0	0.0323	0	0.111	0
06	0	0	0	0.0556	0
07	0	0	0	0	1.0
n[1]	29	31	26	18	3
NH[1]	2	4	2	3	1
NPS[1]	1	7	1	8	0
π[1]	0.000171	0.002438	0.000190	0.003899	0.0
h[1]	0.0690	0.1871	0.0769	0.3072	0.0
ts:tv[1]	1:0	7:0	1:1	8:0	0:0

[1]*n*=No. of samples, NH=No. of haplotypes, NPS=No. of polymorphic sites, π=Nucleotide diversity, *h*=Haplotype diversity, ts:tv=Transition:Transversion

Among all five populations, the Yellow cattle was the most diverse (*h*= 0.3072, π= 0.003899) but the value was lower than the overall haplotype diversity (0.848) and nucleotide diversity (0.00923) of Chinese cattle breed in China (Cai *et al*., 2007). On the other hand, among the Kedah-Kelantan breeds, KKB (*h*=0.1871, π= 0.002438) was the most diverse, while KKA was the least (*h*= 0.0690, π= 0.000171).

Table 2. Kimura's two-parameter pairwise distance within and between populations with standard errors[1]

Population	KKA	KKB	KKC	YC	Banteng
KKA	**0.0002**	0.0005	0.0002	0.0007	0.0165
KKB	0.0014	**0.0025**	0.0005	0.0011	0.0164
KKC	0.0002	0.0014	**0.0002**	0.0007	0.0165
YC	0.0022	0.0031	0.0022	**0.0040**	0.0164
Banteng	0.1038	0.1038	0.1038	0.1042	**0.0000**

[1]Intrapopulation distance along the diagonal in bold and Interpopulation distance below the diagonal and standard error of interpopulation distance above the diagonal

The average pairwise distances within and between populations were analysed using the Kimura's two-parameter pairwise distance with standard error by 1000 bootstrap replications (Table 2). This study showed that average pairwise distance of YC (0.0040) was in the same range in the Chinese cattle breeds studies (0.00048 to 0.01591) (Cai *et al.*, 2007). The Kedah-Kelantan population from Kelantan (KKA) and Terengganu (KKC) had lower average pairwise distance (0.0002) compared to European, African and Indian cattle breeds (0.0011 to 0.0092) (Loftus *et al.*, 1994). These findings also showed that genetic distance between KKA and KKC (0.0002) was closely related in maternal lineage compared to KKA and KKB (0.0014) as well as KKA and YC (0.0022) populations. Thus, the average genetic distance intrapopulation of KK and YC (0.0002-0.0025) indicated that these cattle populations were closely related in their mitochondrial DNA content (Xuan *et al.*, 2010b).

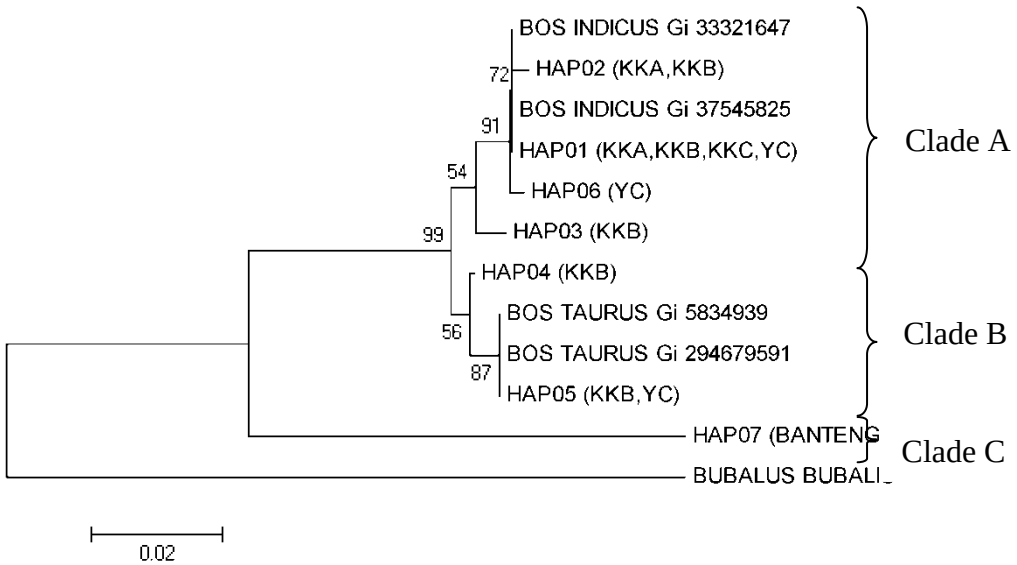

Figure 1. Neighbour Joining tree of ND5 partial sequence of seven haplotypes and five reference sequences with the bootstrap value.

The NJ phylogenetic tree (Figure 1) was constructed with the seven generated haplotypes and additional five sequences cited in GenBank based on the Kimura-2 parameter model with bootstrap confidence of 1000. *Bubalus bubalis* (Gi: 51173180) was used as the outgroup. Three distinct lineages were observed depicting the three maternal origins namely Clade A (*B. indicus*), Clade B (*B. taurus*) and Clade C (*B. javanicus*) which interpreted as three maternal origins. The separation of *B. taurus* and *B. indicus* lineages with 99 % support indicated that these two groups are distinct evolutionary lineages.

Most of the Kedah-Kelantan haplotypes clustered with *B. indicus* reference GenBank haplotypes confirming ancestry of this breed from a crossbred of Shorthorn cattle from west Asia and *B. indicus* (zebu) cattle from India (Devendra, 1975). Interestingly, two haplotypes (Hap04 and Hap05) of Kedah-Kelantan from Pahang (KKB) were clustered into the *B. taurus* lineage. This suggested the occurrence of genetic introgression (Kiesow *et al.*, 2011) between the Kedah-Kelantan cattle from FELDA Farm, Pahang with a female bovine carrying the *B. taurus* maternal

information as reflected in the higher genetic variation within KKB population (0.0025).

On the other hand, YC individuals in Malaysia had more mixed ancestry based on their clustering in both the *B. indicus* clade (Hap01-03 and Hap06) as well as the *B. taurus* clade (Hap04-05). Studied by Cai *et al*. (2007) showed a similar finding in their investigation on the Chinese cattle breeds. The tree also clearly illustrated that Banteng (Hap07) was separated from clade A and B, forming its own distinct lineage support by the higher average pairwise distance between Banteng than other populations (0.1038-0.1042). Thus, further studies with increased sample size and a bigger population of Banteng need to be conducted in order to verify the findings of the present study.

Conclusion

This study showed that all Kedah-Kelantan cattle from Kelantan clustered into *B. indicus* group which meant the population inherited *B. indicus* maternal information from female zebu bovine and showed that the KK was closely related to its ancestral breed. The Kedah-Kelantan cattle population from Kelantan and Terengganu also showed low genetic distance within population and low gene diversity and nucleotides diversity. Realising that the Kedah-Kelantan breed is a valuable genetic resource found in this country and attempts must be made to maintain its genomic through serious conservation and gene preservation initiatives.

References

Bailey, J.F., Healy, B., Jianlin, H., Sherchand, L., Pradhan, S.L., Tsendsuren, T., Foggin, J.M., Gaillard, C., Steane, D., Zakharov, I. and Bradley, D.G. 1996. Genetic variation of mitochondrial DNA within domestic yak population. TAAAS/ IYIC/ Yak Foundation/ FAO-ROAP/ ICIMOD/ ILRIworkshop 181-189.

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

Devendra, C. 1975. Productivity of Kedah/Kelantan cattle in Malaysia. The Proceeding of the 4th continuing education program in animal health and production beef cattle. 18-22 August 1975, UPM Serdang Selangor.

Excoffier, L., Laval, G. and Schneider, S. 2005. Arlequin ver3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatic Online 1: 47-50.

Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. Ann. Rev. Genet. 22: 521-565.

Hiendleder, S., Lewalski, H. and Janke, A. 2008. Complete mitochondrial genomes of *Bos taurus* and *Bos indicus* provide new insights into intra-species variation, taxonomy and domestication. Cytogenet. Genome Res. 120: 150-156.

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

Jia, S., Zhou, Y., Lei, C., Yao, R., Zhang, Z., Fang, X. and Chen, H. 2010. A new insight into cattle in six Asian countries. J. Genet. Genomics 37: 173-180.

Kiesow, A.M., Kasmarik, T. and Binstock, R.L. 2011. Detection of domestic cattle gene introgression in a small population of north American bison. Proc. South Docata Acad. Sci. 90: 79-85.

Kyu-II, K., Jun-Heon, L., Sung-Soo, L. and Young-Hoon, Y. 2003. Phylogenetic relationship of Northeast Asia cattle to other cattle populations determined using mitochondrial DNA D-loop sequence polymorphism. Biochem. Genet. 41: 91-98.

Loftus, R.T., MacHugh, D.E., Bradley, D.G. and Sharp, P.M. 1994. Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci. USA. 91:2757-2761.

Mateus, J.C., Penedo, M.C.T., Alves, V.C., Ramos, M. and Rangel-Figueiredo, T. 2004. Genetic diversity and differentiation in Portuguese cattle breeds using microsatellites. Anim. Genet. 35: 106-113.

Mohamad, K., Olsson, M., van Tol, H.T.A., Mikko, S., Vlamings, B.H., Andersson, G., Rodriguez-Martinez, H., Purwantara, B., Paling, R.W., Colenbrander, B. and Lenstra, J.A. 2009. On the origin of Indonesian cattle. Plos One 4(5): 1-6.

Posada, D. 2004. Collapse 1.2. A Tool for collaping sequence to haplotypes (Online). http://darwin.uvigo.es.

Rogerio, A.C., Ligia, S.L.S.M. and Antonio, C.S. 2007. mtDNA of Nellore and European x Nellore crossing cattle of high performance. Pesq. Agropec. Bras. Brasilia 42(8): 1203-1205.

Saitou, N. and Nei, M. 1987. The Neighbour-joining method: A new method for reconstructing phylogenetic tress. Mol. Biol. Evol. 4: 406-425.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. 2007. Molecular evolutionary genetic analysis (MEGA) software ver4.0. Mol. Biol. Evol. 24: 596-1599.

Tamura, K. and Nei, M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in human and chimpanzees. Mol. Biol. Evol. 10(3): 512-526.

Xuan, T.P., Georgescu, S.E., Manea, M.A., Hermenean, A.O. and Costache, M. 2010a. Genetic diversity and phylogenetic relationships of Romanian cattle breeds inferred from cytochrome b gene partial sequence. Romanian Biotechnological Letters 15(2): 5154-5158.

Xuan, T.P., Georgescu, S.E., Manea, M.A., Hermenean, A.O. and Costache, M. 2010b. Phylogenetic relationships of Romanian cattle to other cattle populations determined by using mitochondrial DNA D-Loop sequence variation. Romanian Biotechnological Letters 15(3): 5287-5292.

Zhang, B., Chen, H., Hua, L., Zhang, C., Kang, X., Wang, X., Pan, C., Lan, X. and Lei, C. 2008. Novel SNPs of the mtDNA ND5 gene and their associations with several growth traits in the Nanyang cattle breed. Biochem. Genet. 46: 362-368.