UDC 577.2
DOI: 10.18413/2409-0298-2016-2-3-45-49

Quan Ke Thai,[1]
Van-Hieu Huynh,[2]
Anh-Dung Chung,[3]
Hoang-Dung Tran[4]

EVALUATION OF GENETIC DIVERSITY OF VIETNAMESE DOGS BASED ON MITOCHONDRIAL DNA HYPERVARIABLE-1 REGION

1) Lecturer, Researcher, PhD candidate, Saigon University, 273 An Duong Vuong, Ward 4, District 5, Ho Chi Minh city, 700000, Vietnam, *E-mail: quan.tk@cb.sgu.edu.vn*
2) Nguyen Tat Thanh University
3) Department of Biotechnology, Institute of Agricultural Science for Southern Vietnam
4) Department of Biotechnology, Institute of High-Tech NTT, Nguyen Tat Thanh University

Abstract
Haplogroup E were detected with high frequency in the population of Phu Quoc ridgeback dogs suggesting that this Vietnamese valuable dog breed originated from domestic dogs harbouring haplogroup E from Vietnam's mainland or from East Asia where the presence of haplogroup E was reported. Evaluation of the genetic diversity of the Vietnam's mainland dogs and haplogroup E screening would support us in tracking out the origin of Phu Quoc ridgeback dogs. Analysing the 582 bp region of mtDNA HV1 of 30 Vietnam's mainland dogs, 27 polymorphics sites (including 1 new insertion) and 16 haplotypes (including 3 new haplotypes) were identified, showing the higher genetic diversity than those of Phu Quoc ridgeback dogs. The lack of haplogroup E in Vietnam's mainland dogs suggested that the haplogroup E in the Phu Quoc ridgeback dogs probably originated from East Asia.
Key words: control region; genetic diversity; haplotype E; Phu Quoc ridgeback dog; Vietnam's mainland dog

Phu Quoc ridgeback dog is one of famous dog breeds in Vietnam. The intelligence and its ability to swim, climb and co-ordinate with others makes it a valuable hunting and guard dog [5]. Having the same characteristic of a ridge of hair running along the back in the opposite direction from the rest of the coat, Phu Quoc ridgeback dog was believed to originate from Thai ridgeback dog although no genetic evidence has been provided to prove their relationship. Up to date, Federation Cynologique Internationale (FCI) has not approved Phu Quoc ridgeback dog as a dog breed because of lacking information.

Many recent studies use mitochondrial DNA (mtDNA) as a source for revealing of the evolutionary history of the domesticated dog [13, 23] as well as the origin of some dog breeds such as Madagascan dog [1], native American dog breeds [22], or Tibetan Mastiff [10]. mtDNA genome of a domestic dog *(Canis lupus familiaris)* is about 16.727 bp long with the control region (CR) spanning positions 15.458 – 16.727. This 1270 bp non-coding DNA segment contains two hypervariable regions, HV1 and HV2 which are separated by a variable numbers of tandem repeats of 5'-GTACACGT(G/A)C-3' motif [3, 7]. HV1 has the higher mutation rate making it a common region of analysis to search for mtDNA variation. A 582 bp region in the mtDNA HV1 exhibited high variation has become one of the molecular markers used to determine the maternally inherited relationships between individuals or populations of the same species [13, 18]. Analyzing the phylogenetic relationship of this region of 1576 dogs, Pang *et al.* (2009) found that all sequences were grouped into six distinct haplogroups A, B, C, D, E, F [13]. Three haplogroups A, B and C are common and have a widespread distribution. 71,3% of HV1 sequences were found harbouring haplogroup A; 17,26% harbouring haplogroup B, and 7,80% harbouring haplogroup C. The distribution of haplogroups D, E, F are more geographically specific which can be found only in particular region such as Turkey, Spain and Scandinavia (haplogroup D); Japan, China and South Korea (haplogroup E); Japan and Siberia (haplogroup F) [13]. In the mentioned analysis, only 1,78% of HV1 sequences were found harbouring haplogroup D, 0,63% harbouring haplogroup E and 0,19% harbouring haplogroup F. These haplogroups are rare and believed to be the result of postdomestication wolf-dog hybridization. Li and Zang (2012) analyzed 582 bp sequences in the HV1

of 50 Tibetan Mastiffs to study the origin and evolution of them and the relationship with other dog breeds across the world. Imes *et al.* (2012) sequenced the whole ~16,7 kb canine mitochondrial genome of 100 unrelated domestic dogs that exhibited 35 haplotypes, all of which clustered within one of four previously described haplogroups (A, B, C, D). Of these, 23 haplotypes were previously observed, 11 were novel, and one was ambiguous. The ten most commonly observed haplotypes (A2, A11, A16, A17, A18, A19, A22, A26, B1, C3) composed approximately 53% of the sequences (Imes *et al.*, 2012). The study on Madagascan dogs also revealed that 100% examined samples harboured haplogroups A, B, C. Thai ridgeback dogs were also included in some previous studies [12, 13] which showed that all of their HV1 sequences (13 samples) belonged to the haplogroup A and B.

At the beginning in the framework of seeking the origin of Phu Quoc ridgeback dogs, we analyzed the 582 bp sequences of 30 Phu Quoc ridgeback dogs living in the Phu Quoc island and found the high proportion (16,67%) of individuals harbouring haplogroup E in the population. The analysis left a question whether the haplogroup E in the population originated from other dogs from Vietnam's mainland or dogs from East Asia [15, 16]. In this study, Vietnam's mainland dogs were collected for analyzing the 582 bp region to check the haplotype diversity, and especially, to screen the haplogroup E in the population. The haplotype diversity would provide information about the population and, more important, contribute to clear the origin of Phu Quoc ridgeback dogs.

Materials and methods

Sample collection and preparation. Hair roots were collected from 30 Vietnamese normal dogs (hereafter called "normal dogs") in rural areas of Ho Chi Minh city and adjacent provinces with low influx of foreign dogs. To avoid the relatedness dogs with maternal bond were excluded. Each sample of dog hair root was stored in a plastic zip-lock bag at -30°C until being used. Samples were numbered from VD1 through VD30.

DNA extraction. DNA was extracted using the simple and cost-effective method [15, 16]. 20 hair roots of each dog in lysis buffer (Tris-HCl 10 mM, pH 8; Triton X100 1%; SDS 1%; EDTA 10 mM) were incubated at 50°C for 1 hour with 5 µl proteinase K (20 mg/ml) added. Released DNA in the solution is separated with other contents (protein, cell debris) by phenol:chloroform:isoamylalcohol mixture, and precipitated in absolute ethanol. Extracted DNA is then dissolved in distilled water and stored at -30 °C. The concentration, yield and purity of DNA samples were determined using optical density and agarose gel electrophoresis methods.

Polymerase Chain Reaction. A CR fragment of 1267 bp was amplified using a pair of primers (15412F- CCACTATCAGCACCCAAAG and 16625R-AGACTACGAGACCAAATGCG) suggested by Ding *et al.* (2007) [6]. Template DNA were added to the PCR reaction mixture consisting of Prime Taq Premix (2X) (0,1U/µl Taq DNA Polymerase; 0,4 mM of each dNTP; 0,4 mM $MgSO_4$; 20 mM KCl; 16 mM $(NH_4)_2SO_4$; 20 mM Tris-HCl, pH8.8) and 10 µM of forward (15412F) and reverse (16625R) primer in a total volume of 20 µl. The amplification program consisted of a pre-denaturation step (95°C, 5 min), followed by 35 cycles of denaturation at 95°C for 30 sec, primer annealing (49°C, 30 sec), extension (72°C, 1 min), and final extension at 72°C for 5 min. PCR reactions were performed in a Techne TC-3000 (USA).

DNA sequencing. PCR products were sequenced with two primers (15412F-CCACTATCAGCACCCAAAG and 16114R-CCTGAAACCATTGACTGAATAG) using the Sanger method [17] by NICEM, South Korea. Sequencing chromatograms were evaluated using FinchTV 1.4.0. Manual editing was used to resolve any inconsistencies between forward and reverse sequences. All sequences were aligned and trimmed using ClustalW [9] to produce 582 bp fragments.

Haplotyping. The 582 bp segment extracted from the first complete mtDNA sequence of the dog (GenBank accession number: U96639.2 [7]) was used as reference sequence. Each 582 bp sequence of dog was aligned with reference sequence using ClustalW program. If the sequence is identical to a sequence with known haplotype, the haplotype of query sequence is defined accordingly; otherwise it would be defined as a novel haplotype. Nucleotide positions are numberred as suggested by Pereira *et al.* (2004) [14].

Phylogenetic analysis. Sixty HV1 sequences in this project, 252 sequences of previously described haplotypes (GenBank accession number: U96639, KF002306, AF531654 – AF531741, AY656703 – AY656710, EU816457 – EU816557, GQ896338 – GQ896345, HQ452433, HQ452439 – HQ452465, HQ452424 – HQ452430, HQ452435 – HQ452438, D83611, D83637, AB007381) [2, 7, 8, 11, 18, 21,

22] and 6 representative coyotes *(Canis latrans)* (GenBank accession number: NC_008093, KF661096, DQ480509 – DQ480511, EU789789) [4, 13, 20] were used for phylogenetic analysis. The phylogenetic tree for the 582 bp sequences was inferred using the Neighbor–Joining method and the Maximum Likelihood method with the T92+I+G model and 1000 repeats bootstrapping.

Results and discussion

DNA extraction, PCR and sequencing. DNA extracted from 30 dog hair samples was good enough for further experiments. The DNA bands which have similar brightness can be seen by exposure of the gel agarose 1% to ultraviolet light. DNA concentrations from 5,5 to 9,1 ng/µl with A_{260}/A_{280} ratio of 1,8-2,0. Sequencing chromatograms showed the high consistency between forward and reverse sequences. After manually edited, 582 bp sequences were produced for further analysis.

mtDNA diversity. Based on the difference among sequences, 30 sequences of normal dogs were classified into 16 groups, in which, 13 groups could be recognized as already known haplotypes, and 3 groups were newly observed haplotypes. In the phylogenetic tree, VD1 and VD9 were grouped in the haplogroup A, and VD22 were in the haplogroup C, so that the haplotype An1 was assigned to VD1, An2 to VD9 and Cn1 to VD22. All of these 16 haplotypes belong to haplogroups A, B and C with the frequencies 63,33%, 20%, and 16,67% respectively, while no representative of haplogroups D, E, F were detecte (Table 1)

Table 1

Haplotypes and their frequencies. N: number of sequences; UT: universal type haplotype

Haplotype	Sample	N	Frequency (%)
A2 (UT)	VD16	1	3,33
A8	VD9/VD17	2	6,67
A9	VD20/VD21	2	6,67
A11 (UT)	VD2/VD3/VD4/VD8/VD14/VD24	6	20
A18 (UT)	VD10	1	3,33
A24	VD6	1	3,33
A29	VD11/VD28	2	6,67
A44 (UT)	VD23	1	3,33
A200	VD26	1	3,33
An1	VD1	1	3,33
An2	VD19	1	3,33
B1 (UT)	VD7/VD18/VD25/VD29/VD30	5	16,67
B6 (UT)	VD5	1	3,33
C2	VD12/VD13/VD15	3	10
C3 (UT)	VD27	1	3,33
Cn1	VD22	1	3,33

The high frequency of haplogroups A, B and C in the population is quite similar to the haplogroup proportions obtained in other published studies [13, 19]. Sixteen out of 30 (53,33%) sequences harbour universal haplotypes (A2, A11, A18, A44, B1, B6, C3). The most common haplotypes were A11 and B1 which could be observed in 11 individuals (36,67%). Ten haplotypes were observed once in the sample set.

There were totally 27 polymorphic sites detected (4,62% of nucleotide sites, i.e the sequence including indel sites), in which, 24 sites were observed with transition, 1 site with transversion and 3 sites with indel (Table 2). In these 27 sites, 18 sites were reported as highly polymorphic sites in canine mtDNA. Interestingly, there was one new insertion after nucleotide at position 15523 that had never reported so far. The insertion was checked carefully by eye and confirmed with a clear single peak in the electropherogram. Pairwise differences between sequences varied from 1 (0,17%) to 18 (3,09%) polymorphic sites.

Table 2

Twenty-seven polymorphic sites in HV1 sequences of normal dogs were found in this study. Nucleotide substitutions are shown relatively to the reference sequence. A blank cell indicates a match to the reference sequence and a dash (-) indicates an indel

	15464.1	15483	15508	15523.1	15526	15529	15557	15595	15611	15612	15613	15627	15630	15632	15639	15643	15650	15652	15710	15800	15815	15912	15938	15955	15959	16003	16025
	-	C	C	-	C	C	T	C	T	T	A	A	C	C	T	A	T	G	C	T	T	C	G	C	C	A	T
A2		T										G			A												
A8												G			A										T		C
A9												G			A												
A11															A												C
A18																											
A24												G			G												C
A29												G			A			A									
A44															A			A									
A200							C					G			A			A									
An1		T		T		T						G			A												
An2												G	T	T	A												C
B1					T			T		C				T	G	G		A		C	C	T		T		G	
B6	C				T					C				T	G	G		A		C	C	T		T		G	
C2			T		T				C						G		C			C		T	-	T		G	
C3			T		T				C						G		C		T	C		T	-	T		G	
Cn			T		T				C		G				G		C			C		T	-	T		G	

The phylogenetic analysis showed two new haplotypes A (An1 and An2) and one new haplotype C (Cn1) which was confirmed with SNP analysis. An1 sequence showed 2 mutations aparted from haplotype A2 (a T insertion after position 15523 and a substitution of a C at 15529 by a T), An2 showed 1 mutation aparted from haplotype A22 (a substitution of a C at 15630 by a T), and Cn1 showed a substitution of a T at 15639 by a G in comparison with haplotype C2). While the mutation profile of An2 and Cn1 were the combination of known substitutions, the mutaion profile of An1 included one insertion at a new polymorphic site. Base composition, with an relative values of 27,43% C, 29,82% T, 26,97% A and 15,79% G, showed that the sequences were comparable to 582 bp sequences from other dog studies, and also, together with sequence comparison, supported the mitochondrial origin of the examined sequences.

Although the number of polymorphic sites in the 582 bp region of normal dogs was slightly lower than those of Phu Quoc ridgeback dogs, a Vietnamese valuable dog breed (27 and 28, respectively) [15, 16], the discovery rate for haplotypes in the population of normal dogs was remarkably higher. There were only 11 haplotypes detected in the population of Phu Quoc ridgeback dogs (the haplotype discovery rate is 36,67%), while 16 haplotypes were detected in the latter (the haplotype discovery rate is 53,33%). Probably, on the isolated Phu Quoc island, the inbreeding had occurred commonly resulting the haplotype diversity of Phu Quoc ridgeback dogs was low, whereas the normal dogs living on the mainland had more chances to mate with individuals harbouring different haplotypes. Six common haplotypes (A11, A18, A24, B1, C2, C3) were shared by both groups. In which, 4 haplotypes A11, A18, B1, C3 were universal type haplotypes which

presented with high frequency in dogs all over the world. Interestingly, there was no haplotype E detected in the population of normal dogs while the frequency of the haplotype E in Phu Quoc ridgeback dogs was quite high (16,67%) although haplotypes D, E, F were distributed less than 3% individual dogs across the world [13]. This result showed no influence from Vietnamese normal dogs on the origin of Phu Quoc ridgeback dogs. Probably, Phu Quoc ridgeback dogs harbouring haplogroup E is originated from domestic dogs harbouring haplotype E migrated to Phu Quoc island via trading boats from East Asia, not from Vietnam's mainland.

Conclusion

Analysing 582 bp mtDNA HV1 sequences of 30 Vietnamese normal dog living in rural areas of Ho Chi Minh city and adjacent provinces, we identified 27 polymorphic sites including a new insertion and 16 haplotypes belonged to 3 haplogroups A, B, C which showed the high genetic diversity of the population. The lack of haplogroup E showed no influence from Vietnamese normal dogs on Phu Quoc ridgeback dogs with high frequency of haplogroup E, supporting us in tracking out the origin of Phu Quoc ridgeback dogs.

References

1. Ardalan A., Oskarsson M.C.R., van Asch B., Rabakonandriania E., Savolainen P. African origin for Madagascan dogs revealed by mtDNA analysis // Royal Society Open Science 2. 2015.

2. Ardalan A., Kluetsch C.F., Zhang A. B., Erdogan M., Uhlen M., Houshmand M., Tepeli C., Ashtiani S. R., Savolainen P. Comprehensive study of mtDNA among Southwest Asian dogs contradicts independent domestication of wolf, but implies dog-wolf hybridization // Ecology and evolution. 2011. V. 1. Pp. 373-385.

3. Bekaert B., Larmuseau M. H., Vanhove M. P., Opdekamp A., Decorte R. Automated DNA extraction of single dog hairs without roots for mitochondrial DNA analysis // Forensic science international. Genetics. 2012. V. 6. Pp. 277-281.

4. Bjornerfeldt S., Webster M.T., Vila C. Relaxation of selective constraint on dog mitochondrial DNA following domestication // Genome research. 2006. V. 16. Pp. 990-994.

5. Chiem N.H., Bien N.V., Thanh V.C., Then N.H., Hong T.T.K., Đuoc N.V., Quoc L.T., Khai L.T.L., Phung B.T.M., Truyen N.Đ. Điều tra nghiên cứu bảo tồn gen động vật: chó Phú Quốc, tỉnh Kiên Giang. 2004.

6. Gundry R.L., Allard M.W., Moretti T.R., Honeycutt R.L., Wilson M.R., Monson K.L., Foran D.R. Mitochondrial DNA analysis of the domestic dog: control region variation within and among breeds // Journal of forensic sciences. 2007. V. 52. Pp. 562-572.

7. Kim K.S., Lee S.E., Jeong H.W., Ha J.H. The complete nucleotide sequence of the domestic dog (Canis familiaris) mitochondrial genome // Molecular phylogenetics and evolution. 1998. V. 10. Pp. 210-220.

8. Klutsch C.F., Seppala E.H., Fall T., Uhlen M., Hedhammar A., Lohi H., Savolainen P. Regional occurrence, high frequency but low diversity of mitochondrial DNA haplogroup d1 suggests a recent dog-wolf hybridization in Scandinavia // Animal genetics. 2011. V. 42. Pp. 100-103.

9. Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. Clustal W and Clustal X version 2.0 // Bioinformatics. 2007. V. 23. Pp. 2947-2948.

10. Li Q., Liu Z., Li Y., Zhao X., Dong L., Pan Z., Sun Y., Li N., Xu Y., Xie Z. Origin and phylogenetic analysis of Tibetan Mastiff based on the mitochondrial DNA sequence // Journal of genetics and genomics = Yi chuan xue bao. 2008. V. 35. Pp. 335-340.

11. Okumura N., Ishiguro N., Nakano M., Matsui A., Sahara M. Intra- and interbreed genetic variations of mitochondrial DNA major non-coding regions in Japanese native dog breeds (Canis familiaris) // Animal genetics. 1996. V. 27. Pp. 397-405.

12. Oskarsson M. Analysis of the origin and spread of the domestic dog using Y-chromosome DNA and mtDNA sequence data. KTH Royal Institute of Technology, Stockholm. 2012.

13. Pang J.F., Kluetsch C., Zou X.J., Zhang A.B., Luo L.Y., Angleby H., Ardalan A., Ekstrom C., Skollermo A., Lundeberg J., Matsumura S., Leitner T., Zhang Y.P., Savolainen P. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves // Molecular biology and evolution. 2009. V. 26. Pp. 2849-2864.

14. Pereira L., Van Asch B., Amorim A. Standardisation of nomenclature for dog mtDNA D-loop: a prerequisite for launching a Canis familiaris database // Forensic science international. 2004. V. 141. Pp. 99-108.

15. Quan T.K., Tu N.V., Hieu H.V., Cong N.T., Dung T.H. A simple protocol for DNA extraction from dog hairs // Tap chi Sinh hoc. 2016. V. 38. Pp. 124-132.

16. Quan T.K., Tu N.V., Trinh T.N., Hieu H.V., Dung C.A., Dung T.H. Evaluation of genetic diversity of Phu Quoc ridgeback dogs based on mitochondrial DNA Hypervariable-1 region // Tap chi Cong nghe Sinh hoc. 2016. V. 14. Pp. 245-253.

17. Sanger F., Nicklen S., Coulson A.R. DNA sequencing with chain-terminating inhibitors // P Natl Acad Sci USA. 1977. V. 74. Pp. 5463-5467.

18. Savolainen P., Zhang Y.P., Luo J., Lundeberg J., Leitner T. Genetic evidence for an East Asian origin of domestic dogs // Science. 2002. V. 298. Pp. 1610-1613.

19. Suárez N.M., Betancor E., Fregel R., Pestano J. Genetic characterization, at the mitochondrial and nuclear DNA levels, of five Canary Island dog breeds // Animal genetics. 2013. V. 44. Pp. 432-441.