

The complete mitochondrial genome sequence of the indigenous I pig (*Sus scrofa*) in Vietnam

Hieu Duc Nguyen¹, Tuan Anh Bui², Phuong Thanh Nguyen¹, Oanh Thi Phuong Kim¹, and Thuy Thi Bich Vo^{1,*}

* **Corresponding Author:** Thuy Thi Bich Vo
Tel: +84-4-32191174, **E-mail:** thuytbvo@igr.ac.vn

¹Institute of Genome Research, Vietnam Academy of Science and Technology, Hanoi 100000, Vietnam

²Institute of Forensic Science, Hanoi 100000, Vietnam

Submitted Aug 11, 2016; Revised Sept 30, 2016;
Accepted Dec 13, 2016

Objective: The I pig is a long nurtured longstanding breed in Vietnam, and contains excellent indigenous genetic resources. However, after 1970s, I pig breeds have become a small population because of decreasing farming areas and increasing pressure from foreign breeds with a high growth rate. Thus, there is now the risk of the disappearance of the I pigs breed. The aim of this study was to focus on classifying and identifying the I pig genetic origin and supplying molecular makers for conservation activities.

Methods: This study sequenced the complete mitochondrial genome and used the sequencing result to analyze the phylogenetic relationship of I pig with Asian and European domestic pigs and wild boars. The full sequence was annotated and predicted the secondary tRNA.

Results: The total length of I pig mitochondrial genome (accession number KX094894) was 16,731 base pairs, comprised two rRNA (12S and 16S), 22 tRNA and 13 mRNA genes. The annotation structures were not different from other pig breeds. Some component indexes as AT content, GC, and AT skew were counted, in which AT content (60.09%) was smaller than other pigs. We built the phylogenetic trees from full sequence and D loop sequence using Bayesian method. The result showed that I pig, Banna mini, wild boar (WB) Vietnam and WB Hainan or WB Korea, WB Japan were a cluster. They were a group within the Asian clade distinct from Chinese pigs and other Asian breeds in both phylogenetic trees (0.0004 and 0.0057, respectively).

Conclusion: These results were similar to previous phylogenic study in Vietnamese pig and showed the genetic distinctness of I pig with other Asian domestic pigs.

Keywords: Complete Mitochondrial Genome; Genetic Distance; Mitochondrial DNA; I Pig; Phylogenetic Relationships

INTRODUCTION

The I pig originated from Nam Dinh province, Vietnam and before 1970s it was common in Red River delta. I pig has some characteristics as black skin, short, sparse and black hair, big and short mouth, short body, sagging back, hanging down belly, ten breasts, and weak legs. There are several remarkable characteristics of I pig, e.g., they can adapt to poor quality feed, harsh climate, disease resistance, good quality meat, high nutrition and small string. This breed is early maturity (3 to 4 months for female and 6 to 8 months for male), from 110 to 115 days of pregnant period and births 8 to 10 offsprings in once. However, I pig size is quite small, the body weight at four months is less than 14 kg, mature pigs weight less than 70 kg and percentage of lean meat is less than 40%, so it is gradually replaced by Mong Cai pig, which has higher productivity. Nowadays, there are approximately 20 individuals locate in Thanh Hoa province [1,2]. In 2005, Vietnam Ministry of Agriculture and Rural Development put I pig on the list of rare breeding animal genetic resources need to be conserved and also showed the importance of I pig genetic resources.

The characteristic mitochondrial DNA (mtDNA) structure and function is suitable to become

a molecular maker to evaluate the molecular evolution, classification, population genetic analysis, and relationship identification. Firstly, mtDNA is a small, less than 20 kb circular sequence [3]. Secondly, it is haploid, maternally inherited and no recombination [4]. Thirdly, the mutation is primarily single base pair substitutions [5] and the evolution rate of mtDNA is more rapid and constant than nuclear DNA [3,6]. Animal mtDNA has one control region and codes for 13 protein subunits, 2 rRNAs, and 22 tRNAs [7]. The control region as known as *D loop*, which regulates both replication and transcription of mtDNA, has used in several previous phylogenetic studies [8,9].

In the present study, we report the full sequence of I pig mitochondrial genome. The full sequence and *D loop* region sequence are used to estimate the genetic relationships between I pig and Asian and European pigs. The phylogenetic results are useful for classification and identification of genetic origin in I pig breed conservation activities and further studies.

MATERIALS AND METHODS

Sample collection and DNA extraction

All animals used for sample isolation were handled in accordance with the National Institute of Animal Sciences (NIAS) (Hanoi, Vietnam) care and management guidelines. Five milliliter (5 mL) of blood was collected with anticoagulant from jugular vein of the I pigs reared in the Thanh Hoa province (Vietnam) according to I pig breed conservation project of NIAS. Total DNA extraction was performed by standard phenol-chloroform method [10].

Amplification by polymerase chain reaction

The full sequence of mtDNA was amplified by using 30 primer pairs into 30 overlap segments (Table 1). Each 25 µL polymerase chain reaction (PCR) reaction consist of 12.5 µL GoTaq Green Master Mix, 1.0 µL DNA template, 0.5 µL of each primer (10 ppm), and 10.5 µL sterile deionized water. The PCR thermocycler was an initial denaturation at 94°C for 5 min, followed by 25 cycles of 94°C for 30 s, 50°C to 55°C based on the primer sequences for 30 s, and 72°C for 50 min, with a final extension at 72°C for 10 min. The product sequence length was checked by 1% agarose gel electrophoresis. PCR products were purified by QIAquick PCR Purification Kit (Qiagen, Hilden, Germany).

Sequencing and structure analysis

The PCR products were sequenced using ABI3500 DNA sequencer. The DNA sequences were assembled using DNA Dragon v1.6.0 software and exported to FASTA file. The complete sequence was annotated using MITOS and DOGMA. The mitochondrial circle structure was built using GenomeVx and the tRNAs structure was downloaded from online report of MITOS analysis result. The C, G, A, T frequencies and A+T-content of the entire mitochondrial genome were calculated using DAMBE v6.3.17 [11]. After that, skewness for each protein coding gene was evaluated

Table 1. Primer pairs for I pig complete mitochondrial DNA

Primer pair	Sequence	
	Forward	Reverse
1	AGGAGACTAACTCCGCCAT	GCGGATACTTGATGTGT
2	ACTAAGTCAATGCCTATTCTG	CAAATGTATGAAACCTCAG
3	CTACACAATAACCTCCATA	TGGCACGAGATTACCAACT
4	GCTCATAACGCCTTGCTC	ATTCTTTCATCTTCCCTT
5	CACAACCATGCAAGAAGAGACA	ACAACCAGCTATCACCAGGC
6	CCGTAAGGGAAAGATGAAAG	TATGGTTATTTTGACTGGT
7	CCGTGCAAAGGTAGCATA	CCAACATCGAGGTCGTAA
8	TGGGGTGACCTCGGAGTAC	AATATGGCGAAAGTCCGG
9	CGAGCAGTAGCCAAACA	GGTCGTATCGGAATCGTG
10	GTATCAGGCTTTAACGTAGA	TGGTAATACTGCTGCATTC
11	CACAGAAGCAGCCACAAA	ATGGGATAGGGATAAAGT
12	ACATAGGATGAATGACAGC	TGGTGAAAGTAGTCAGAAAC
13	GCACTGCCTTGAGCCTAC	GTGTTCAGGTTGCCGTCT
14	CCCATTATGATTGGGGGTTT	TGCTGTGTATGCGTCAGGAT
15	CACCTTTGTAATCATATTCGTAG	TAGTTGGAAAGGGTAAGC
16	TTCATCTACTAACAGCAG	TTGAGTTCGGTTGATTCTG
17	GCTTCATGCCATTGTAC	TTATAGCGGAATCCTGTG
18	GCAAGCCCAGAATCAACCG	CGAGGAGGATTGAGGTGT
19	ATACCACATAGTAAACCCAA	CCTGTAGCCACAAAAGAAA
20	CTAAACACCTCAATCCTCC	TTGGACGTAATCGGTACCCG
21	CCTTGCAGGGTTACTTAT	TTCCGGTTGTGGTTCTT
22	CGGTACCGATTACGTCCAA	CCGATTAGATTGATGGATG
23	ACCAGCTCTATCTGCTTA	GAGGCTTTGATGTTGTTA
24	ATGATGACTAATAGCAAGCC	GGGATGTAGTCCGAATTG
25	CATCGGAGACATTGGATT	AGTTGGCTTGAAGTTGAG
26	CCTACTCCTAGCTGCAGCAG	ATTATGGAGATTACTCGTGG
27	TCCGCATCATCATTACTA	TTTATGGTGGACTTGGGT
28	TAATTACCACGAGTAATCTC	TTCTACGAGGTCTGTTCCG
29	GGAGCATCCATATCTTT	GGGTAGTTGTCTGGGTCT
30	TCGTAGAATGAATCTGAGG	GGTGATACGCATGTTGACTG

using the following formulas: GC skew = (G-C)/(G+C) and AT skew = (A-T)/(A+T).

Phylogenetic analysis

The *D loop* and full sequence of I pig were aligned with other pig sequences using MUSCLE algorithm [12] of MEGA6 [13]. The best model for full mitochondrial genome group was HKY+G and TN93+I. We used Bayesian phylogenetic analysis in BEAST v1.8.3 [14] with Yule process and MCMC 10000000 to calculate the alignment data and Figure Tree v1.4.2 to build the phylogenetic trees.

RESULTS

Gene annotation

The total length of I pig mitochondrial genome was 16,731 base pairs (bp) (Figure 1). Similar to other *Sus scrofa* mitogenomes [15,16], there were 2 ribosome RNA genes, 13 protein coding genes, 22 transfer RNA genes and a non-coding *D loop* region. The base composition was 34.66% for A, 26.24% for C, 13.35%, for G and 25.75% for T, the A+T content was 60.41%. The *D loop*

less Cytosine bias was 16S rRNA (-0.11), whereas 12S rRNA gene had the most Adenine bias and *COX1* gene had Cytosine content equal Guanine. The only one gene coded in L strand was *ND6* had positive GC skew (0.56) and negative AT skew (-0.35) (Table 2). Because the tandem repeat motif “CGTGCGTACA” could cause heteroplasmy [17], we removed them out the mitogenome sequence before analysis. The component of I pig full sequence after remove motif was 34.83%, 26.19%, 13.15%, and 25.82% for A, C, G, and T, respectively. Compared the AT content, GC and AT skew between I pig sequence and other pig breeds, the range of AT content, GC skew and AT skew of published full sequence were (60.19; 60.82), (-0.34; -0.30) and (0.12; 0.15), respectively. The range of AT content, GC skew, and AT skew of published D loop sequence were (60.73; 61.80), (-0.38; -0.31); (0.15; 0.19), respectively. We found that the AT content of I pig D loop (60.09%) was smaller than other pigs, whereas its full sequence (60.65%)

was larger than Bihu pig (60.19%). The D loop of I pig also had less Cytosine bias than most pig breeds, excepted WB Yunnan (-0.31). GC and AT skew of full genome and AT skew of D loop did not have too much different (Table 3).

The secondary structure of tRNAs was predicted by MITOS online software (Figure 2). Twenty one tRNAs excepted for tRNA *Ser-1* had the typical cloverleaf structure. The tRNA *Ser-1* dihydrouridine arm did not form a stable structure.

Phylogenetic analysis

We built the phylogenetic trees of the I pig full sequence and *D loop* sequence with the published mitogenomes of pig breeds from three different geographical regions: European Countries, Mekong region and China. In both phylogenetic trees, most of the pigs divided into two clades: Asian clade and European clade, the genetic distances between them in full sequence and *D loop*

Table 3. Pig breeds, their geographic definitions and sequence components

Geographic definitions	SeqName	Genbank ID	Full sequence			D loop sequence			
			%AT	GC skew	AT skew	%AT	GC skew	AT skew	
North East Asia	I-Vietnam	KX094894	60.65	-0.33	0.15	60.09	-0.32	0.16	
	Jeju	AY879785.1	-	-	-	61.45	-0.35	0.17	
	Korean native pig	AY879794.1	-	-	-	61.24	-0.35	0.15	
	WB-China northeast	EU333163.1	60.75	-0.33	0.15	61.80	-0.34	0.15	
	WB-Japan	AB015085.1	-	-	-	60.91	-0.33	0.16	
Yellow river valley	WB-Korea	AY574047.1	60.81	-0.33	0.15	61.41	-0.33	0.16	
	Bamei	EF545583.1	60.79	-0.33	0.15	61.31	-0.34	0.16	
	Huzu	EF545588.1	60.80	-0.33	0.15	61.31	-0.34	0.16	
European country	Berkshire	AY574045.1	60.72	-0.33	0.15	61.53	-0.34	0.16	
	Duroc	AY337045.1	60.68	-0.33	0.15	61.34	-0.35	0.16	
	Hampshire	AY574046.1	60.67	-0.33	0.15	61.73	-0.36	0.16	
	Iberian	FJ236994.1	60.71	-0.34	0.15	61.63	-0.35	0.16	
	Landrace	AF034253.1	60.72	-0.33	0.15	61.53	-0.35	0.15	
	Large White	KC250275	60.73	-0.33	0.15	61.53	-0.35	0.16	
	Pietrain	KC469587	60.72	-0.34	0.15	61.48	-0.35	0.16	
	WB-European	FJ237000.1	60.72	-0.34	0.15	61.44	-0.35	0.17	
	Mekong region	Banna mini	GQ220328.1	60.72	-0.33	0.15	60.84	-0.34	0.16
		Dahe	GQ220329.1	60.76	-0.33	0.15	61.08	-0.33	0.16
Thailand indigenous pig		FM244493.1	-	-	-	61.30	-0.33	0.16	
WB-Malaysia		EF545592.1	60.82	-0.33	0.15	61.41	-0.38	0.19	
WB-Vietnam		EF545584.1	60.74	-0.33	0.15	61.31	-0.34	0.16	
South China	WB-Yunnan	EF545573.1	60.72	-0.33	0.15	60.73	-0.31	0.15	
	Lantang	KC250274	60.78	-0.33	0.15	61.43	-0.34	0.16	
	Lanyu	DQ518915.2	60.75	-0.30	0.12	61.45	-0.34	0.16	
	WB-Fujian	EF545569.1	60.81	-0.33	0.15	61.29	-0.33	0.16	
	WB-Hainan	EF545572.1	60.81	-0.33	0.15	61.32	-0.33	0.15	
Yangtze River Region	Aba	EF545578.1	60.79	-0.33	0.15	61.31	-0.34	0.16	
	Bihu	EF545591.1	60.19	-0.31	0.14	61.31	-0.34	0.16	
	Jinhua	KC469586	60.78	-0.33	0.15	61.54	-0.34	0.16	
	Kele	EF536857.1	-	-	-	61.32	-0.34	0.16	
	Taoyuan	AM040653.1	-	-	-	61.34	-0.34	0.16	
	WB-Jiangxi	EF545579.1	60.78	-0.33	0.15	61.41	-0.34	0.16	
	Wei	EF545577.1	60.80	-0.33	0.15	61.31	-0.34	0.16	
	Xiang pig	KC250273	60.82	-0.33	0.15	61.41	-0.33	0.16	

-, non published full sequences; WB, wild boar.

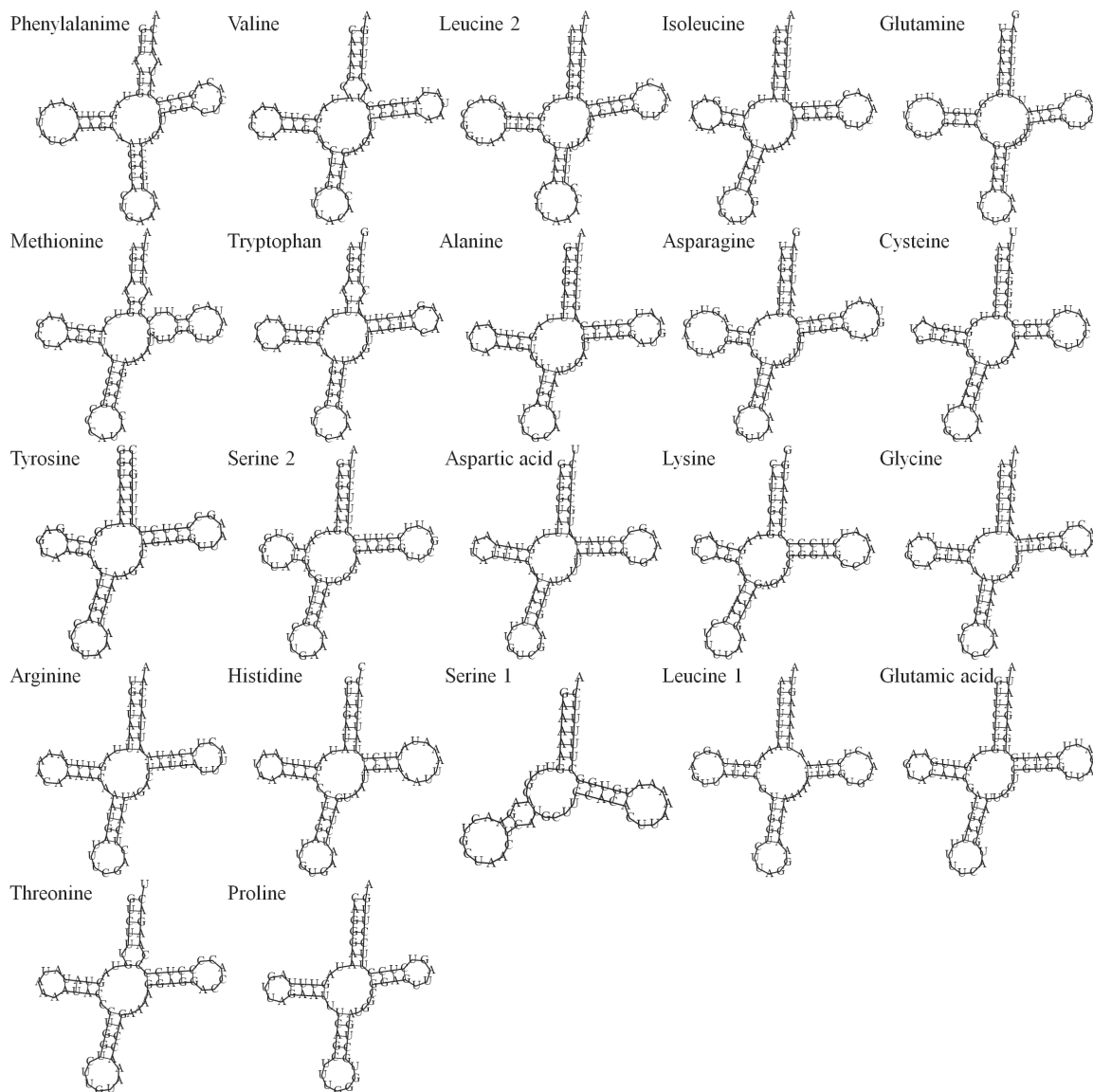


Figure 2. Secondary structure of 22 tRNAs encoded by the I pig mitogenome.

sequence phylogenetic trees were 0.008 and 0.0137, respectively.

The twenty-seven public reference sequences were used to build the full sequence phylogenetic tree (Figure 3). There were two large groups in the Asian clade. Although the Berkshire and WB Korea were in Asian clade they felt outside of both large groups, the distance from WB Korea was 0.0045 and Berkshire was further (0.0053). Genetic distance between two groups was 0.0004. The I pig, Banna mini pig, and WB Vietnam and WB Hainan clustered a group. I pig was a sister taxon with the Banna mini (0.0002), close was the WB Vietnam (0.0016) while the WB Hainan had most distant relationship (0.0022).

There were 33 sequences in the *D loop* phylogenetic tree and this tree showed the different distribution (Figure 4). The WB Japan and WB Korea joined to the group of I pig. Therefore, this group could be separated into two subgroups: I pig and Banna

mini continued to be cluster (0.0006) and WB Vietnam, WB Korea and WB Japan were another subgroup. The genetic distance between the two subgroups was 0.0065. WB Hainan felt out of this group, it and Taoyuan pig became sister taxons (0.0014). The genetic distance between the two big groups of the Asian clade was 0.0057.

DISCUSSION

This study was the first molecular research in I pigs. The result produced an overview of I pig mitogenome and compared it with other published *Sus scrofa* mitochondrial sequences. Although the I pig mitochondrial structure was similar to the pig breed mitochondrial structure in previous researches, there were some differences in sequence components of I pig especially in *D loop*

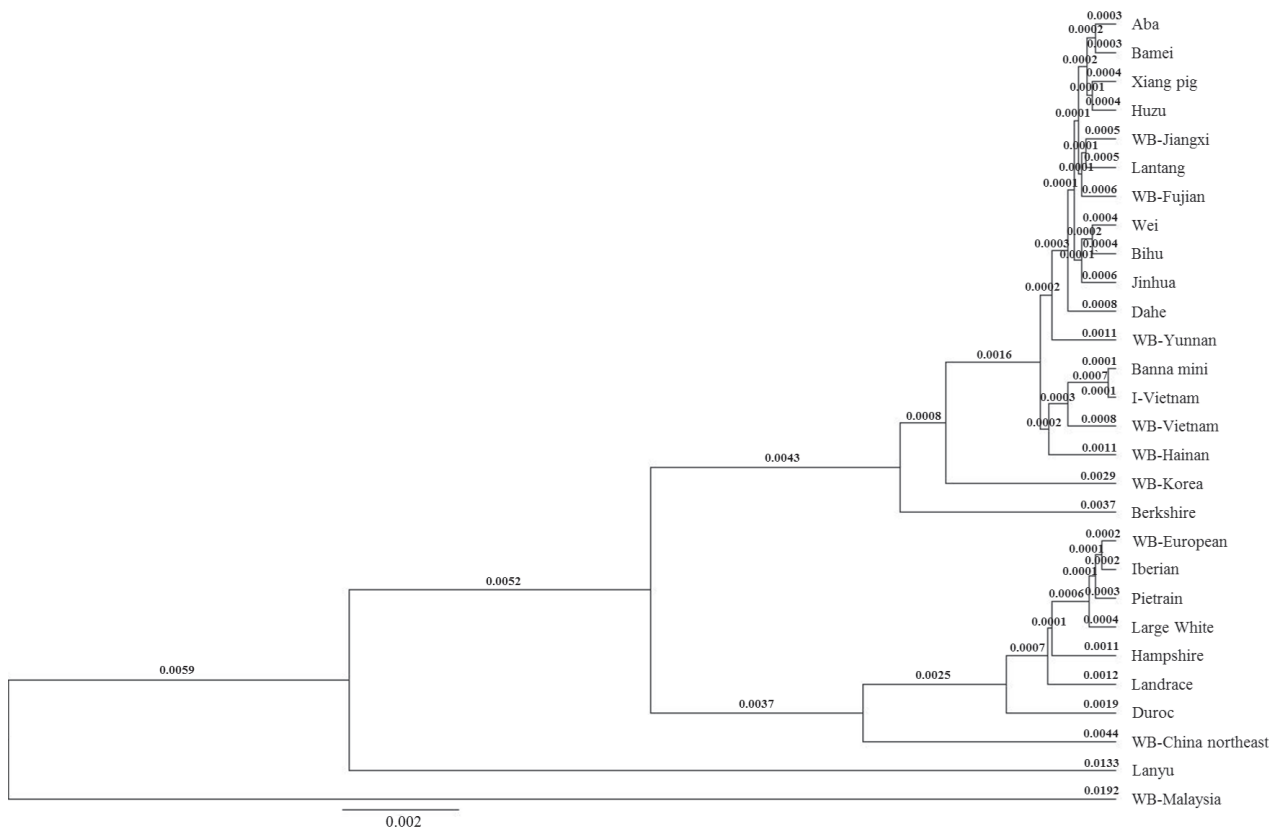


Figure 3. Full sequence phylogenetic tree using Bayesian method. WB, wild boar.

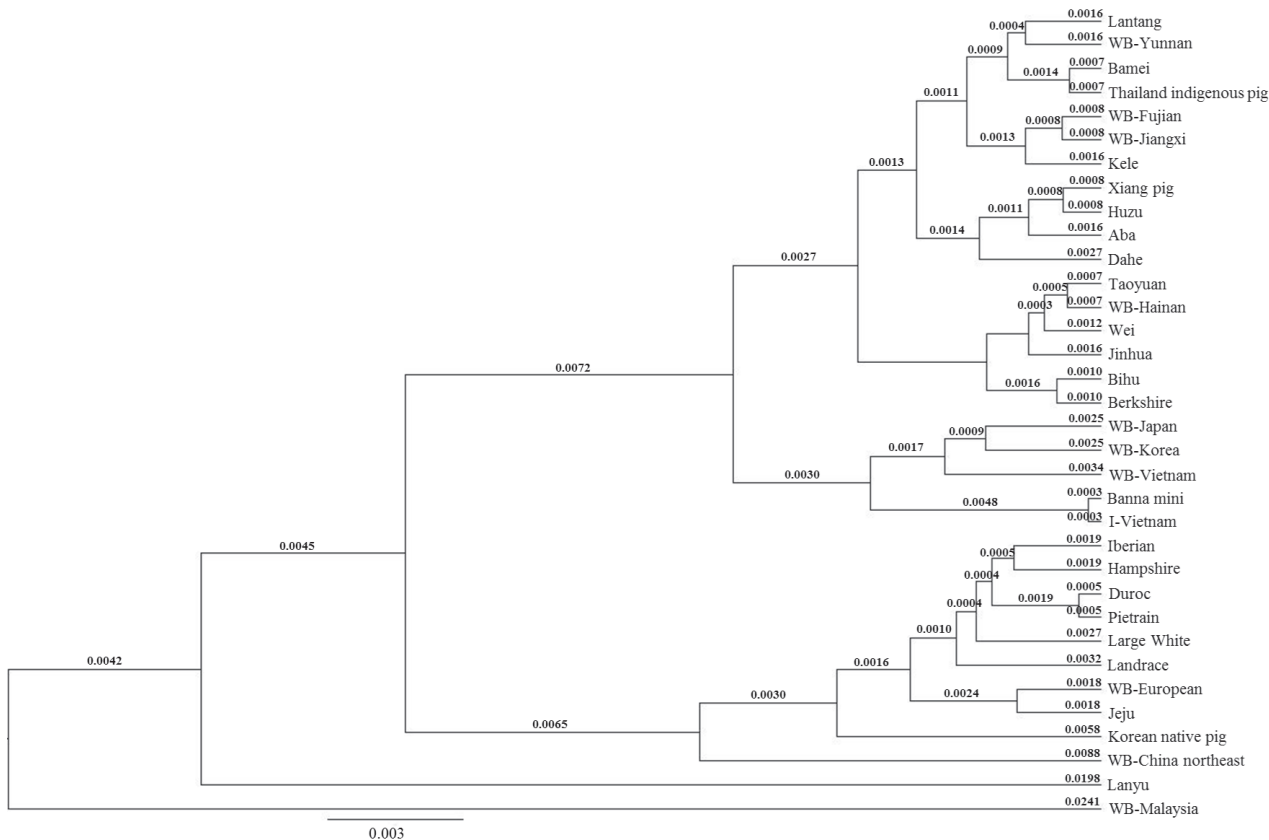


Figure 4. D loop sequence phylogenetic tree using Bayesian method. WB, wild boar.

region. AT content, GC skew, and AT skew were used to study the nucleotide-compositional behavior of mitochondrial genomes, and related to phylogenetic [18,19]. So that, the difference of I pig component indexes showed the genetic variation of I pig with pig breeds taxa.

The phylogenetic trees showed the difference between Asian and European pigs. This result identified the divergence point of Asian and European domestic pigs had occurred [20] or 746,000 years before present [21]. And, phylogenetic trees illustrated the contribution of Asian and European pig breeds to the development of each other, which were similar to the genetic evidence Giuffra found for introgression in the late 18th and early 19th centuries [20].

The previous researches reported the close genetic distance between Vietnamese pig and Asian pig. Both of phylogenetic trees in our study showed similar results, namely I pig had a close genetic relationship with Banna mini, a Mekong pig breed, WB Vietnam and other Asian wild boars. While the most of Chinese pig and some breeds of the Mekong region and Berkshire pig became a separate large group. This result agreed with a study of Lan and Shi which showed that Vietnamese wild boar and Chinese domestic pig were separated by a large genetic distance [22]. Moreover, a later study of Hongo et al [23] compared the 574 bp mitochondrial DNA sequences of 30 Vietnamese pigs with Asian domestic pigs and wild boars revealed some Vietnamese domestic pigs were closely related to wild boars and genetically distinct from East Asian domestic pig groups. However, there are a need for additional studies to ensure the domestic origin of I pig in Vietnam.

In present research, we reported the mitochondrial full genome sequence of I pig (KX094894) in NCBI genbank. The sequence was used to predict and analyze the component and structure of I pig mitogenome. The result of the phylogenetic analysis was conducted to illustrate the genetic separation between I pig and most of Asian domestic pigs.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number: 106-NN.05-2014.66.

REFERENCES

1. Dang-Nguyen TQ, Tich NK, Nguyen BX, et al. Introduction of various Vietnamese indigenous pig breeds and their conservation by using assisted reproductive techniques. *J Reprod Dev* 2010;56:31-5.

2. Phung TV. Livestock technique of reproductive sow. 1st Hanoi, Vietnam: Labour Publishing House; 2005.
3. Bruford MW, Bradley DG, Luikart G. DNA markers reveal the complexity of livestock domestication. *Nat Rev Genet* 2003;4:900-10.
4. Giles RE, Blanc H, Cann HM, Wallace DC. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA* 1980;77:6715-9.
5. Wolstenholme DR. Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol* 1992;141:173-216.
6. Brown WM, George M, Wilson AC. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA* 1979;76:1967-71.
7. Boore JL. Animal mitochondrial genomes. *Nucleic Acids Res* 1999;27:1767-80.
8. Kim KI, Lee JH, Li K, et al. Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. *Anim Genet* 2002;33:19-25.
9. Toro MA, Rodríguez J, Silio L, Rodríguez C. Genealogical analysis of a closed herd of black hairless Iberian pigs. *Conserv Biol* 2000;14:1843-51.
10. Sambrook J, Russell DW. *Molecular cloning: a laboratory manual*. 3rd ed. New York, US: ColdSpring-Harbour Laboratory Press; 2001.
11. Xia X, Xie Z. DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 2001;92:371-3.
12. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792-7.
13. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013;30:2725-9.
14. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 2012;29:1969-73.
15. Sun JL, Zhang B, Ma QY, et al. Analysis of mtDNA Partial D-loop Sequences in Luchuan Pig. *Zhongguo Xu Mu Shou Yi* 2010;6:035.
16. Li M, Jin L, Ma J, et al. Detecting mitochondrial signatures of selection in wild Tibetan pigs and domesticated pigs. *Mitochondrial DNA Part A DNA Mapp Seq Anal* 2016;27:747-52.
17. Ghivizzani SC, Mackay SL, Madsen CS, Laipis PJ, Hauswirth WW. Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. *J Mol Evol* 1993;37:36-47.
18. Hassanin A, Léger N, Deutsch J. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. *Syst Biol* 2005;54:277-98.
19. Wei S-J, Shi M, Chen X-X, et al. New views on strand asymmetry in insect mitochondrial genomes. *PLoS ONE* 2010;5:e12708.
20. Giuffra E, Kijas J, Amarger V, Carlborg Ö, Jeon J-T, Andersson L. The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics* 2000;154:1785-91.
21. Fernández A, Alves E, Ovilo C, Rodríguez M, Silió L. Divergence time estimates of East Asian and European pigs based on multiple near complete mitochondrial DNA sequences. *Anim Genet* 2011;42:86-8.

22. Lan H, Shi L. The origin and genetic differentiation of native breeds of pigs in southwest China: an approach from mitochondrial DNA polymorphism. *Biochem Genet* 1993;31:51-60.
23. Hongo H, Ishiguro N, Watanobe T, et al. Variation in mitochondrial DNA of Vietnamese pigs: relationships with Asian domestic pigs and Ryukyu wild boars. *Zool Sci* 2002;19:1329-35.