Journal of Animal and Veterinary Advances 11 (21): 3984-3990, 2012
ISSN: 1680-5593

Genetic Diversity and Population Structure of Tibetan Pig Inferred from Mitochondrial DNA Sequence Analysis

Mingwang Zhang, Mingzhou Li, Long Jin, Yan Li and Xuewei Li
Institute of Animal Genetics and Breeding, College of Animal Science and Technology,
Sichuan Agricultural University, 625014 Ya'an, Sichuan, P.R. China

Abstract: Tibetan pig is an endangered plateau type pig breed which distributed mainly in Qinghai-Tibetan Plateau. In present study, researchers analyzed 218 prior published partial mitochondrial D-loop sequences (415 bp) by using phylogenetic, network, mismatch distribution and Analysis of Molecular Variance (AMOVA) analysis to clarify the population differentiation and explore the relationships among populations from different geographical locations of Tibetan pigs. The results revealed a total of 43 haplotypes in 218 samples from seven Tibetan pig populations in which only one common Haplotype (H1) was shared by the seven populations. Haplotype diversity was high (0.889) whereas nucleotide diversity among all individuals was low (0.00534). It is notable that Hezuo Tibetan pigs have 19 unique haplotypes in total 24 haplotypes of 53 individuals with the highest nucleotide diversity (0.00765). Bayesian tree and median joining network procedures carried out on the data showed that seven Tibetan pig haplotypes were essentially indistinguishable except some unique haplotypes from Hezuo. The AMOVA analysis indicated that the genetic variation mainly occurred within populations (85.41%) but variance among populations was only 14.59%. The unimodal mismatch distribution of haplotypes, together with significant negative values of Fu's F_S and negative values of Tajima's D suggest that a recent population expansion of Tibetan pig has been occurred in Diqing, Hezuo and total population.

Key words: Tibetan pig, mitochondrial DNA, genetic diversity, population structure, mismatch

INTRODUCTION

Tibetan pig is a typical miniature pig breed which is mainly distributed in the Qinghai-Tibetan Plateau of China and largest continuous high elevation ecosystem in the world with an average elevation of >4000 m (Zhang *et al*., 1986; Yang *et al*., 2011). Based on classification of pig breeds in China, Tibetan pig was classified as only one of Plateau type (Zhang *et al*., 1986). However, with the growing demand for pig meat and bettering the life for the local people, high input-high output commercial pig breeds have been introduced and crossed with Tibetan pigs which led the pure Tibetan pigs face a risk of extinction (Jiao *et al*., 2009). Accordingly, Gong'gyamda Tibetan Pig National Nature Reserve has been established in Gong'gyamda County in 2011, for the purpose of protecting and utilization germplasm resource of Tibetan pig.

Clearly, research on the genetic diversity and population structure of this breed is very important for the formulation of adequate protection and management strategies. Unfortunately, there has been no large scale study of the population genetics of Tibetan pig, only a few studies have been carried out in population genetic variation based on mtDNA data (Jiao *et al*., 2009; Yang *et al*., 2011). The earlier study conducted by Jiao *et al*. (2009) involved limited sample size (80 individuals from 5 populations). A recent article about mtDNA D-loop variation within seven regional Tibetan pig populations has revealed local origin of the pig breed in the Tibetan Plateau (Yang *et al*., 2011), however it did not discuss about the genetic structure and diversity of this breed.

In this study, to expand limited geographic sampling of earlier researches (Jiao *et al*., 2009; Yang *et al*., 2011), researchers download 212 the published mtDNA D-loop sequence data from GenBank which represents 218 Tibetan pig individuals from seven geographical populations in Tibetan plateau. The objective of this study is to analyze the population structure and diversity of Tibetan pig based on the increase population coverage and sample size and to provide a theoretical reference for this breed conservation and sustainable use in China.

MATERIALS AND METHODS

DNA data: A total of 415 bp mtDNA D-loop fragment of 218 individuals of Tibetan pig was retrieved from

Corresponding Author: Xuewei Li, Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, 625014 Ya'an, Sichuan, P.R. China

GenBank as described by the earlier published literatures (Wu *et al*., 2007; Jiao *et al*., 2009; Larson *et al*., 2010; Yang *et al*., 2011). The samples span the entire known range of Tibetan pig from seven localities in China (25 individuals from Linzhi, 11 from Shannan, 17 from Changdu, 56 from Diqing of Yunnan, 34 from Aba and 22 from Ganzi of Western Sichuan, 53 from Hezuo of Gansu) (Table 1 and Fig. 1). Sample sites were shown in Fig. 1. In addition, one mtDNA control region sequence of African warthog (*Phacochoerus africanus*) was retrieved from GenBank (accession numbers DQ409327) and was used as outgroup for reconstruction the phylogenetic tree based on earlier study (Wu *et al*., 2007). All samples of Tibetan pig sequences used in this study with GenBank accession number and geographical origin is shown in supplementary Table 1.

Data analysis: All download sequences were aligned with Clustal X 1.83 (Thompson *et al*., 1997) with default parameters. Researchers estimated the numbers of haplotypes, values of haplotype diversity (h) (Nei, 1987) and nucleotide diversity (π) (Nei and Tajima, 1981) using DnaSP 5.1 (Librado and Rozas, 2009). The Tajima's

Table 1: Mitochondrial DNA haplotype variable sites and distribution of Tibetan pig

	Variable nucleotides sites								
	--								
	11111 1122222222 2223333333 4								
	479913456 8923667778 8890137778 0								
Haplotype	7433519184 7741252674 5764051238 1	LZ	SN	CD	DQ	AB	GZ	HZ	Total
H1	CCCCACTGCC CTCGCCCCGC ACAACCAATT A	15	5	2	5	10	17	4	58
H2	....C. .	1	-	1	-	-	-	-	2
H3	....T......T	2	-	-	-	-	1	-	3
H4	....T.......	3	1	2	9	3	-	3	21
H5	....C .	1	-	1	2	-	-	-	4
H6	.... .C..T..... ....C .	2	-	1	3	-	-	-	6
H7	....T.....	1	2	-	-	11	-	6	20
H8	....A......	-	2	-	-	4	-	17	23
H9	.... .C........	-	1	-	-	1	-	-	2
H10	....T.T.....	-	-	3	17	-	-	2	22
H11	....T..... ....C. .	-	-	3	3	-	-	-	6
H12	.......A.. ....C .	-	-	1	1	-	-	-	2
H13	....T.T..... ......G... .	-	-	2	3	-	-	-	5
H14	....C T	-	-	1	-	-	-	-	1
H15	....T.T..... .....G.... .	-	-	-	1	-	-	-	1
H16	.........T	-	-	-	1	-	-	-	1
H17	......C... .C........ ....C .	-	-	-	1	-	-	-	1
H18	....T.T..... ....C .	-	-	-	6	-	-	-	6
H19	....C..... ..T.......	-	-	-	2	-	-	-	2
H20	.... .C........ ....C .	-	-	-	2	-	-	-	2
H21	....T.T..... ....C. .	-	-	-	-	4	-	-	4
H22	....G....... .	-	-	-	-	1	-	-	1
H23	.... .CT....... ...G...... .	-	-	-	-	-	2	-	2
H24	....A.	-	-	-	-	-	2	-	2
H25	....T..... ....C .	-	-	-	-	-	-	1	1
H26	....G.. .	-	-	-	-	-	-	1	1
H27	....T..G..	-	-	-	-	-	-	2	2
H28	G.GG...... ....TA.G..	-	-	-	-	-	-	1	1
H29	G..G...... ....T..G.. .A......C. .	-	-	-	-	-	-	1	1
H30	G..G...... ....T..G..	-	-	-	-	-	-	2	2
H31	G......... ....T..G.. C.C....... .	-	-	-	-	-	-	1	1
H32	....T..G.. ....C .	-	-	-	-	-	-	1	1
H33	G..G...... ..T.T..G..	-	-	-	-	-	-	1	1
H34	G..G.A.... ..T.T..G.. .A........ .	-	-	-	-	-	-	1	1
H35	....T..... ....A..... .	-	-	-	-	-	-	1	1
H36	....T..G.. .A..A..... .	-	-	-	-	-	-	1	1
H37	....AT.....	-	-	-	-	-	-	1	1
H38	G......... ...AT..G..	-	-	-	-	-	-	1	1
H39	G......... ....T..G..	-	-	-	-	-		1	1
H40	....A...G..	-	-	-	-	-	-	1	1
H41	GT.G....A. A...T.AG..	-	-	-	-	-	-	1	1
H42	G......... ....T..G.. .A..A..... .	-	-	-	-	-	-	1	1
H43	G....A.... ....T..G..	-	-	-	-	-	-	1	1
Sum		25	11	17	56	34	22	53	218

Haplotype in bold occurred in more than one population. Dots (.) indicate nucleotide identity as in haplotype H1, LZ, SN, CD, DQ, AB, GZ and HZ represent population of Linzhi, Shannan, Changdu, Diqing, Aba, Ganzi and Hezuo, respectively

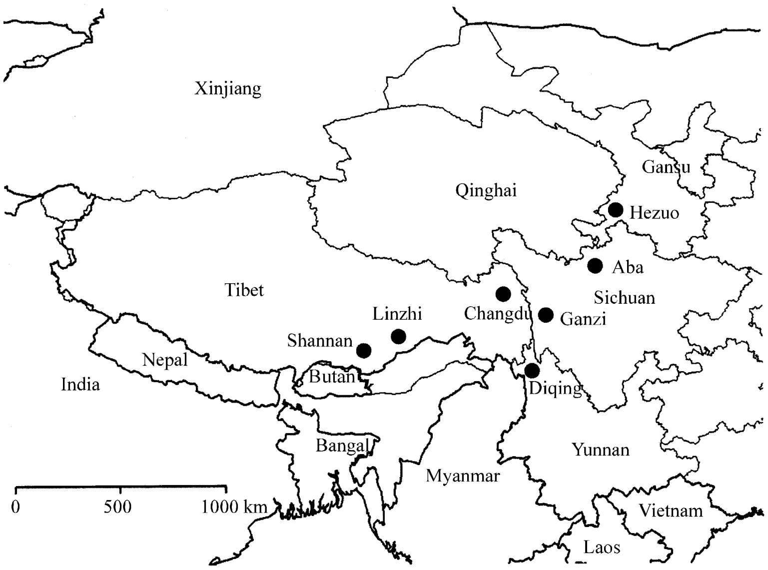

Fig. 1: The geographic location of the seven Tibetan pig populations in Tibetan plateau

D (Tajima, 1989), Fu's Fs (Fu, 1997) neutrality tests and mismatch distribution for Tibetan pigs were calculated also by DnaSP 5.1. Population average pairwise differences and Analysis of Molecular Variance (AMOVA) (Excoffier *et al.*, 1992) were calculated using Arlequin Software (Version 3.11) (Excoffier *et al.*, 2005).

Researchers used Software MODELTEST 3.7 (Posada and Crandall, 1998) and PAUP* 4.0b10 (Swofford, 2001) to find the best fit model of 56 models of nucleotide evolution under the Akaike Information Criterion (AIC) for the sequence data. Phylogenetic analysis using Bayesian Inference (BI) was implemented using MRBAYES Version 3.1.2 (Ronquist and Huelsenbeck, 2003) based on GTR+I+G Model. Four independent Markov Chain Monte Carlo (MCMC) analysis were run for 5×10^6 generations simultaneously to avoid entrapment in local optima with one tree save every 100 generations and with the first 25% discarded as burn-in after the average standard deviation of split frequencies <0.01. Researchers constructed the median-joining networks of haplotype data with the program NETWORK 4.6.1.0 (Bandelt *et al.*, 1999; Polzin and Daneschmand, 2003) (http://www.fluxusengineering.com) to investigate the possible relationships among haplotypes of Tibetan pigs.

RESULTS AND DISCUSSION

In total, 43 haplotypes were identified and no insertion or deletion was observed among the 218 Tibetan pig individuals in the 415 bp partial mtDNA D-loop fragment (Table 1). There were 31 variable sites of which 12 were parsimony informative and 19 were singleton sites in the Tibetan pig haplotypes. These 43 haplotypes are shown in Table 1 to show their distributions in populations, the number of haplotypes detected in each population varies from 4 (Ganzi) to 24 (Hezuo). Among these haplotypes, the most frequent Haplotype (H1) includes 58 (26.6%) individuals from all seven populations; haplotypes H8, H10, H4 and H7 include 23, 22, 21 and 20 individuals, respectively. About 19 haplotypes were shared by 2-58 samples while the remaining 24 haplotypes were found in only one individual. It is notable that Hezuo Tibetan pigs have 19 unique haplotypes in total 24 haplotypes from 53 individuals (Table 1), this result is consistent with prior studies (Jiao *et al.*, 2009; Yang *et al.*, 2011).

The genetic diversity of each population and all samples is shown in Table 2 and haplotype diversity values varies from 0.403 (Ganzi) to 0.934 (Changdu). Meanwhile, the Hezuo population had the highest nucleotide diversity value (0.00731) and Ganzi population had lowest one (0.00207) (Table 2). The total haplotype diversity was 0.8894 with nucleotide diversity 0.00534.

The Bayesian consensus tree for 43 mtDNA D-loop haplotypes showed a very shallow phylogenetic structure with low bootstrap support (Fig. 2). Interestingly, the tree showed that most of the branches did not correspond to the geographic location; in other words, researchers did not find the geographic distance in proportion to

Table 2: Summary of haplotype diversity (h), nucleotide diversity (π) and Fu's (Fs) and Tajima's D test for control region sequences

Population	Haplotype diversity (h)	Nucleotide diversity (π)	Fs	p-value	Tajima D	p-value
Linzhi	0.6330	0.00280	-2.540	0.050	-0.80752	>0.05
Shannan	0.7820	0.00245	-2.018	0.089	-0.93168	>0.05
Changdu	0.9340	0.00574	-4.518	0.009	0.02391	>0.05
Diqing	0.8640	0.00470	-5.627	0.002	-0.52932	>0.05
Aba	0.7950	0.00335	-1.409	0.110	-0.14538	>0.05
Ganzi	0.4030	0.00207	-0.264	0.235	-1.11483	>0.05
Hezuo	0.8817	0.00731	-15.525	0.000	-0.98518	>0.05
Total	0.8894	0.00534	-40.622	0.000	-1.66047	>0.05

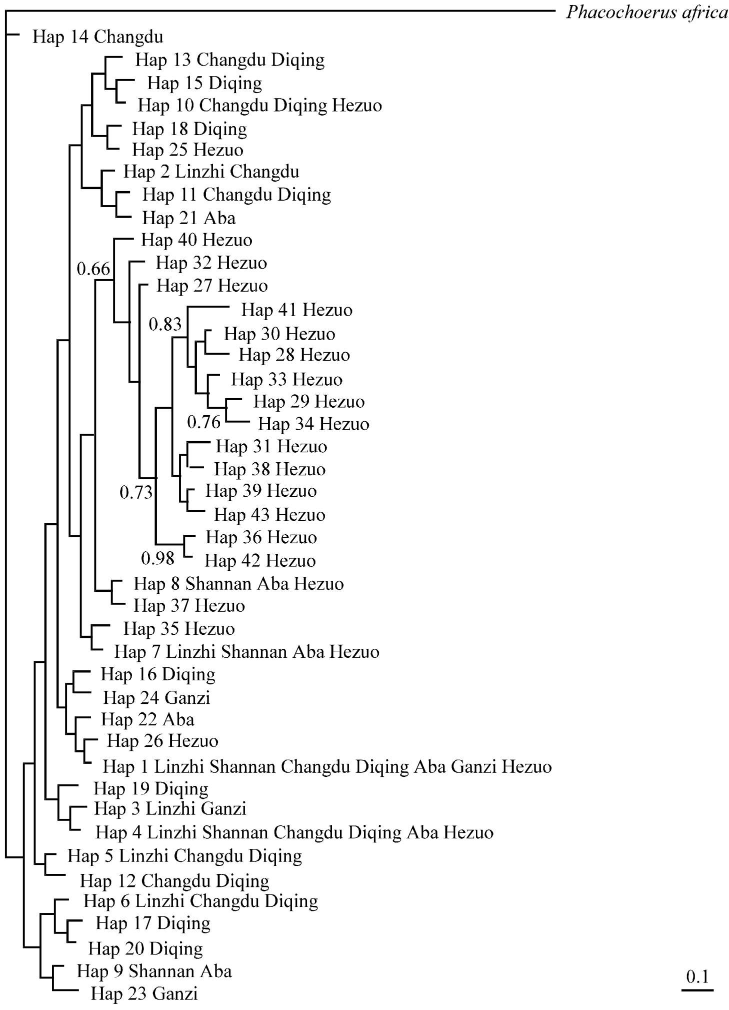

Fig. 2: Bayesian consensus tree of the observed 43 haplotypes of Tibetan pigs with African warthog (*Phacochoerus africanus*) as outgroup. Bayesian posterior probabilities are given only for those nodes which obtained values >0.60 with this method

population genetic similarity. This might be caused by gene flow among the Tibetan pig populations and the fact that different populations have low selection pressure. However, Fig. 2 clearly shows that most unique haplotypes (15 of 19) from Hezuo Tibetan pig population were clustered into a branch (BPP = 0.66) which is separated from other populations of this breed.

Similarly, the parsimony median-joining network of Tibetan pig haplotypes also showed that populations from different geographic regions intermingled as similar pattern as the Bayesian tree (Fig. 3). Haplotypes from each population did not cluster together and separate from other populations except some haplotypes from Hezuo. In addition, haplotypes from Hezuo almost distributed the whole network. These results indicated

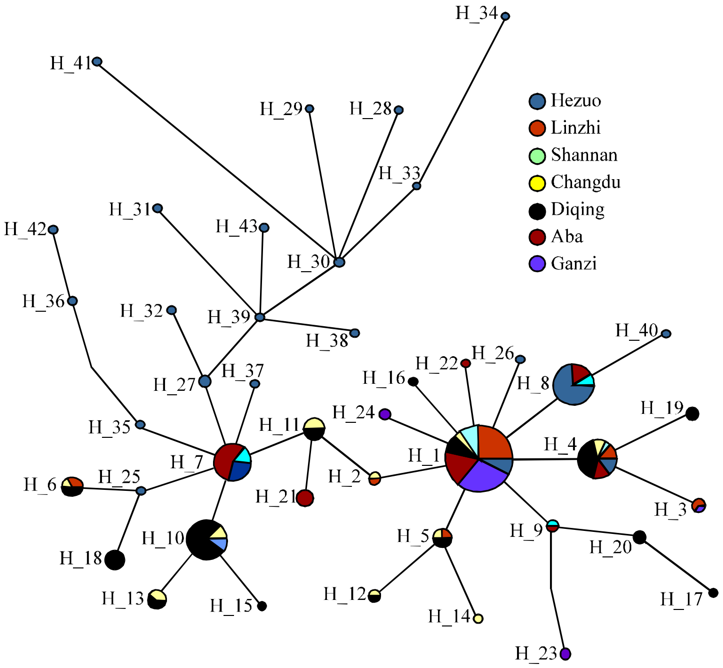

Fig. 3: Parsimony median-joining network profile of the mtDNA D-Loop region haplotypes in Tibetan pigs. The haplotypes are symbolized by circles and separated by one substitution step. Circle size is proportional to haplotype frequency and the length of each branch is proportional to the number of associated mutations

that there was no correspondence between the geographic region and different population. It is evident that the high-frequency haplotype H1 (58 individuals, 26.6% of the total 218 individuals) and the haplotypes around it showed a star-like distribution pattern that is a signal of exponential population expansion on the plateau, although not indicate any overt geographic structure (Fig. 2). H1 shared among individuals from the seven populations (15 individuals from Linzhi, 5 from Shannan, 2 from Changdu, 5 from Diqing, 10 from Aba, 17 from Ganzi and 4 from Hezuo) (Table 1 and Fig. 3). H1 is the core haplotype that has a series of one, two or greater than two mutation distance derivatives (derived haplotypes) detected in other haplotypes. Therefore, taken together, haplotypes within network might have originated from the center Haplotype (H1) as a result of domestication events followed by subsequent expansion.

Table 3: AMOVA analysis of genetic variation in Tibetan pig populations

Source of variation	df	Variance components	Percentage of variation	Fixation index
Among populations	6	0.18065	15.84	F_{ST} = 1.14046
Within populations	211	0.95982	84.16	
Total	217	1.14046		

df = degree of freedom

Researchers performed Analysis of Molecular Variance (AMOVA) to assess geographical structuring of genetic variability within population and among populations. The AMOVA analysis indicated that most of genetic variance is attributable within populations (84.16%, p<0.001) and only lower amount of genetic variance is among populations (15.84%, p<0.001) (Table 3). The Fixation index F_{ST} was 0.15840 (p<0.001). Therefore, the AMOVA analysis indicated that the genetic variation mainly occurred with populations instead of among populations. In other words, most of the observed molecular variance was due to genetic differences within these seven geographic populations indicating a relatively poor geographic structure for Tibetan pig.

Table 4 showed the pairwise genetic differentiation between and within the seven Tibetan pig populations. The average number of pairwise differences between populations is from 0.96330 (Ganzi and Shanan) to 3.17469 (Hezuo and Diqing) and within population between from 0.86249 (Ganzi) to 3.05765 (Hezuo). The genetic differentiation within population almost equal for that between populations also indicates genetic variation mainly occurred within populations as AMOVA analysis. The corrected average pairwise differences between the populations is from 0.00181 (Linzhi and Shannan) to 0.70721 (Diqing and Ganzi). Genetic differentiation was significant among all sampled populations with the

Table 4: Population average pairwise differences for control region sequences of Tibetan pig based on Kimura 2P Distance Method

Population	Linzhi	Shannan	Changdu	Diqing	Aba	Ganzi	Hezuo
Linzhi	**1.16675**	1.09608	1.98139	2.01067	1.38617	1.02416	2.56747
Shannan	0.00181	**1.02179**	1.98480	2.05784	1.24664	0.96330	2.30939
Changdu	0.19849**	0.27437**	**2.39906**	2.21674	1.96298	2.05037	3.14995
Diqing	0.44551**	0.56516**	0.03542	**1.96357**	1.99184	2.12024	3.17469
Aba	0.10398**	0.03694	0.06464	0.31124**	**1.39762**	1.34185	2.45529
Ganzi	0.00954	0.02115	0.41960**	0.70721**	0.21179**	**0.86249**	2.53204
Hezuo	0.45527**	0.26967*	0.42159**	0.66408**	0.22766**	0.57197**	**3.05765**

Above diagonal: average number of pairwise differences between populations (PiXY). Diagonal elements: average number of pairwise differences within population (PiX) in bold. Below diagonal: corrected average pairwise difference [PiXY−(PiX + PiY)/2]. Corrected average pairwise differences which are statistically different are indicated. $^{*}p<0.05$; $^{**}p<0.01$

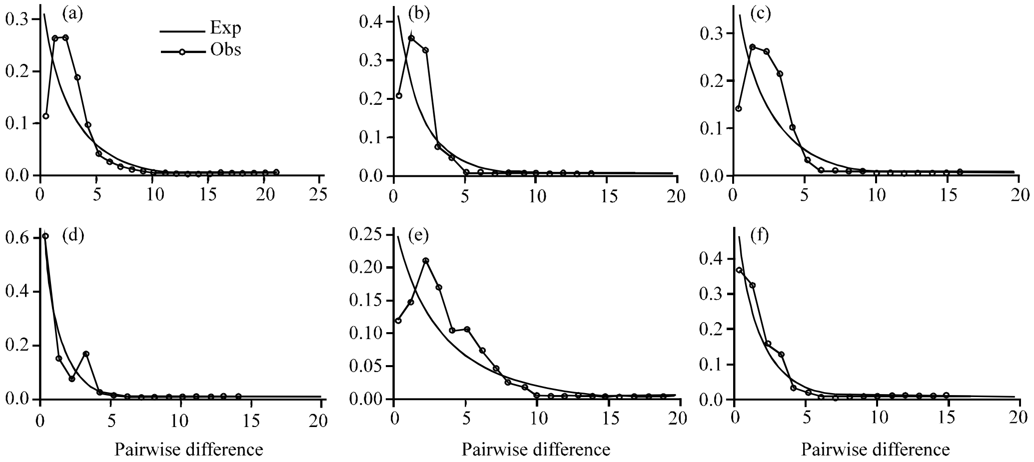

Fig. 4: Mismatch distribution of the Tibetan pig mtDNA control region for a) all samples, b) Aba, c) Diqing, d) Ganzi, e) Hezuo and f) Linzhi

exception of Ganzi vs. Linzhi and Shannan, Aba vs. Shannan and Changdu, Diqing vs. Changdu and Shannan vs., Linzh (Table 4).

Researchers used two methods to investigate population demographic history. First, researchers investigated the demographic history by comparing mismatch distributions for all the samples and each population (sample size >20) using DnaSP. Assuming the infinite sites model, the mismatch distribution is smooth and often unimodal as a result of population expansion whereas for stationary populations the distribution is ragged and often multimodal (Harpending *et al*., 1998). Second, researchers also calculated Fu's Fs neutrality test (Fu, 1997) and Tajima's D test to confirm the evidence of possible population expansion. Under the assumption of neutrality, a population expansion produces a large negative value of Fs (Fu, 1997).

Historical expansion of the Diqing, Huzuo and total population is suggested by biased unimodal mismatch distributions (Fig. 4), significant ($p<0.05$) negative values of Fs and negative values of Tajima's D (Table 2). However, non-significant values for Fu's Fs test and the ragged mismatch distribution supported sequence evolution consistent with the expectations of selective neutrality and stable demographic history of the Aba, Ganzi and Linzhi populations.

Tibetan pig is now facing the danger of extinction due to its original distribution areas are decreasing gradually (Li and Luo, 1993) and crossbreeding with exotic boars. It is well known that genetic diversity exists in animal population and is considered the raw material of evolution (Su *et al*., 2001).

The result shows that Tibetan pig harbors a considerable amount of genetic variation. Therefore, maintain the genetic diversity will contribute to strengthen conservation, management and sustainable use of this special plateau miniature pig germplasm resources.

CONCLUSION

The present study shows that all populations' samples were clustered together and there was no population-specific clustering pattern except Hezuo. Meanwhile, researchers also found substantial mtDNA diversity in Tibetan pigs. In particular, the nucleotide diversity of Hezuo Tibetan pigs (0.00731) was found to be the highest in the seven sampled populations (Table 2). Hence, conservation strategy of the Tibet pig should not only to protect this breed natural habitat but also the genetic diversity. Furthermore, special conservation attention should be paid to Hezuo Tibetan pig because of its 19 unique haplotype and the highest nucleotide diversity.

ACKNOWLDEGEMENTS

This project was supported by Shuang Zhi Ji Hua Foundation of Sichuan Agricultural University (01570709) and Open projects of State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences (GREKF11-01). Thanks are given to Dr. Chen SY and Wang LJ for their useful ideas. Special thanks to Dr. He K and Wang JS (Kunming Institute of Zoology, Chinese Academy of Sciences) for geographical information and maps.

REFERENCES

Bandelt, H.J., P. Forster and A. Rohl, 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol., 1: 37-48.

Excoffier, L., G. Laval and S. Schneider, 2005. Arlequin (version 3.01): An integrated software package for population genetics data analysis. Evol. Bioinform. Online, 1: 47-50.

Excoffier, L., P.E. Smouse and J.M. Quattro, 1992. Analysis of molecular variance inferred from metric distances among DNA haplo-types: Application to human mitochondrial DNA restriction data. Genetics, 131: 479-491.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147: 915-925.

Harpending, H.C., M.A. Batzer, M. Gurven, L.B. Jorde and S.T. Sherry, 1998. Genetic traces of ancient demography. Proc. Natl. Acad. Sci., 95: 1961-1967.

Jiao, T., S.G. Zhao, C. Wang, G. Danbaro, L.P. An and Y. Cai, 2009. Mitochondrial DNA D-Loop diversity of tibetan pig populations. Philippine Agri. Sci., 92: 362-369.

Larson, G., R. Liu, X. Zhao, J. Yuan and D. Fuller *et al*., 2010. Patterns of East Asian pig domestication, migration and turnover revealed by modern and ancient DNA. Proc. Natl. Acad. Sci., 107: 7686-7691.

Li, J.Y. and Z. Luo, 1993. Research on the habits and characteristics of tibet pigs on tibet plateau. Ecol. Domestic Anim., 14: 18-21.

Librado, P. and J. Rozas, 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25: 1451-1452.

Nei, M. and F. Tajima, 1981. DNA polymorphism detectable by restriction endonucleases. Genetics, 97: 145-163.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, USA.

Polzin, T. and S.V. Daneschmand, 2003. On Steiner trees and minimum spanning trees in hypergraphs. Oper. Res. Lett., 31: 12-20.

Posada, D. and K.A. Crandall, 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics, 14: 817-818.

Ronquist, F. and J.P. Huelsenbeck, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 12: 1572-1574.

Su, B., Y.X. Fu, Y.X. Wang, L. Jin and R. Chakraborty, 2001. Genetic diversity and population history of the red panda (Ailurus fulgens) as inferred from mitochondrial DNA sequence variations. Mol. Biol. Evol., 18: 1070-1076.

Swofford, D.L., 2001. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0, Sinauer Associates, Sunderland, MA., USA.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123: 585-595.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougi and D.G. Higgins, 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res., 25: 4876-4882.

Wu, G.S., Y.G. Yao, K.Y. Qu, Z.L. Ding and H. Li*et al*., 2007. Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. Genome Biol., 8: R245-245.

Yang, S., H. Zhang, H.M. Mao, D.W. Yan and S.X. Lu *et al*., 2011. The local origin of the Tibetan pig and additional insights into the origin of Asian pigs. PLoS One, 6: e28215-e28215.

Zhang, Z.G., B.T. Li, X. H. Chen, 1986. Pig Breeds in China. Shanghai Scientific and Technical Publishers, Shanghai, China, pp: 175-178.