
Volume 90, Issues 1–3, May 2010 ISSN 0921-4488
ELSEVIER
SMALL
RUMINANT
RESEARCH
The Official Journal of the International Goat Association

Small Ruminant Research 90 (2010) 88–94

Contents lists available at ScienceDirect

Small Ruminant Research

journal homepage: www.elsevier.com/locate/smallrumres

Genetic diversity of Chinese indigenous sheep breeds inferred from microsatellite markers

T. Zhong[a,1], J.L. Han[a,b], J. Guo[c], Q.J. Zhao[a], B.L. Fu[a], X.H. He[a], J.T. Jeon[d], W.J. Guan[a], Y.H. Ma[a,*]

[a] *CAAS-ILRI Joint Laboratory on Livestock and Forage Genetic Resources, Institute of Animal Science, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China*
[b] *International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi 00100, Kenya*
[c] *Poultry Institute, Chinese Academy of Agricultural Sciences (CAAS), Yangzhou 225003, China*
[d] *Division of Applied Life Science, Gyeongsang National University, Jinju 660-701, Republic of Korea*

ARTICLE INFO

Article history:
Received 3 September 2009
Received in revised form 30 January 2010
Accepted 2 February 2010
Available online 15 March 2010

Keywords:
Chinese sheep
Microsatellite
Genetic diversity and differentiation
Bayesian clustering

ABSTRACT

To determine the genetic diversity and phylogenetic relationships among Chinese sheep, 10 indigenous breeds and one introduced breed were genotyped for 19 microsatellite loci. The mean number of alleles per breed ranged from 5.44 (Guide Black Fur sheep) to 9.13 (Ujumqin sheep and Hulunbeier sheep), the expected heterozygosity varied from 0.623 (Guide Black Fur sheep) to 0.737 (Zhaotong sheep), and the allelic richness ranged from 5.169 (Guide Black Fur sheep) to 7.610 (Zhaotong sheep). The deviation from Hardy–Weinberg equilibrium (HWE) was statistically significant ($P < 0.05$) at three loci (*SRCRSP5*, *OarAE129* and *DYMS1*) in most of the breeds. Chinese sheep breeds had maintained a high level of within-population genetic differentiation (95.23%), with the remainder explained by differentiation among populations (4.77%). The genetic differentiation pattern and genetic relationships among Chinese sheep breeds displayed a high consistency with the traditional classification. Both the Bayesian cluster and principal component analyses showed a reliable clustering pattern, which revealed three major clusters in Chinese indigenous sheep (Mongolian sheep, Kazakh sheep and Tibetan sheep), except Zhaotong and Guide Black Fur sheep. There were probably caused by different breeding history, geography isolation and different levels of inbreeding. This study will help to interpret the genetic characters of Chinese indigenous sheep and benefit to the future conservation programs.

1. Introduction

Sheep were one of the earliest livestock species to be domesticated, which occurred 9000 years ago mainly in the Fertile Crescent. It is proposed that sheep originated from at least three ancestral subspecies of the wild Mouflon, as inferred by multiple mitochondrial lineages (Guo et al., 2005; Chen et al., 2006; Chessa et al., 2009). Later with human migration, climate change, ecological environment and selections for different objectives, over 2351 sheep breeds have been developed worldwide today (DAD-IS, 2010). China has a great diversity of ecological conditions and an abundance of small ruminant resources. There are 31 indigenous sheep breeds which are classified into three groups: Mongolian sheep, Kazakh sheep and Tibetan sheep (Zheng, 1988). These native sheep breeds are distributed in different environmental areas such as pastoral area,

---

* Corresponding author. Tel.: +86 10 62813463; fax: +86 10 62895351.
E-mail addresses: zhongtao04@yahoo.com.cn (T. Zhong), hanjl@iascaas.net.cn (J.L. Han), guojun71@yahoo.com.cn (J. Guo), zqj126@yahoo.com.cn (Q.J. Zhao), bl@iascaas.net.cn (B.L. Fu), xhhe@iascaas.net.cn (X.H. He), jtjeon@gnu.ac.kr (J.T. Jeon), wjguan86@iascaas.net.cn (W.J. Guan), yuehui.ma@263.net (Y.H. Ma).

[1] Present address: Division of Applied Life Science, Gyeongsang National University, Jinju 660-701, Republic of Korea.

0921-4488/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.smallrumres.2010.02.001

Table 1
Summary statistics of mean number of alleles (MNA), observed (H_O) and expected (H_E) heterozygosities, allelic richness (r) in 11 sheep populations analyzed by 16 microsatellite loci.

Popultion	Code[a]	N[b]	MNA	H_O	H_E	r[c]	Group and morphology
Ujumqin sheep	UQ	45	9.13	0.684	0.693	7.383	Mongolia, short fat-tailed
Tan sheep	TAN	48	8.25	0.705	0.706	6.704	Mongolia, short fat-tailed
Hu sheep	HU	43	6.69	0.648	0.691	5.941	Mongolia, short fat-tailed
Hulunbeier sheep	HLBR	44	9.13	0.685	0.723	7.458	Mongolia, fat-rumped
Small Tailed Han sheep	STH	48	8.25	0.678	0.703	6.815	Mongolia, short fat-tailed
Anduo sheep	AD	47	7.31	0.673	0.680	6.234	Tibet, thin-tailed
Gangba sheep	GB	44	6.69	0.662	0.675	5.871	Tibet, thin-tailed
Guide Black Fur sheep	GBF	31	5.44	0.631	0.623	5.169	Tibet, thin-tailed
Zhaotong sheep	ZT	46	8.19	0.674	0.737	6.961	Tibet, thin-tailed
Hazakh sheep	HZK	40	8.94	0.720	0.732	7.610	Kazakh, fat-rumped
Poll Dorset	PD	48	5.25	0.632	0.610	4.574	Australian, fat-rumped

[a] Code: the abbreviation of breed name.
[b] n: the number of animals in the sampled population.
[c] Allelic richness per population is based on a minimum sample size of 19 individuals.

agro-pastoral and agricultural region, and they have many advantages, such as good environmental adaptations, high reproductive rates, abundant products and disease resistance. Detail information about the breeds can be drawn from the website (http://www.cdad-is.org.cn/).

However, several decades ago, many exotic breeds were imported into China to initiate a country-wide genetic improvement with local breeds for productive performance (wool) and weight gain. Indeed, the crossbreeding and selection methods had increased production successfully, but also aggravated a dramatic loss of genetic variability without enough attention on conservation (Ma et al., 2006). The current state of sheep genetic resources in China is represented by a crossbreeding among local breeds and introduced breeds. Especially, some local breeds are mainly under smallholder production systems in countryside, which tend to be affected by human and economic factors strongly. It was reported that local sheep populations were subject to annual diminishment, as for the environmental change and habitat deterioration, and about 15% sheep breeds had been identified as in danger of extinction (Ma et al., 2002).

To investigate the genetic resources of Asian small ruminants, a co-operation was launched by the International Atomic Energy Agency (IAEA), including Pakistan, Sri Lanka, Vietnam, Bangladesh, Iran, Indonesia and China, to assess genetic variability by microsatellite markers. In this study, to assess the degree of genetic diversity and differentiation among Chinese indigenous sheep breeds, 19 microsatellite markers were genotyped in 10 Chinese local breeds and one introduced breed. This study will be helpful to determine the current and future breeding programs, breed management and conservation prioritization.

2. Materials and methods

2.1. Sample collection and preparation

Ear tissue samples were randomly collected from 484 genetically unrelated individuals representing 11 breeds: Ujumqin sheep (UQ, 45), Tan sheep (TAN, 48), Hu sheep (HU, 43), Hulunbeier sheep (HLBR, 44), Small Tailed Han sheep (STH, 48), Anduo sheep (AD, 47), Gangba sheep (GB, 44), Guide Black Fur sheep (GBF, 31), Zhaotong sheep (ZT, 46), Kazakh sheep (HZK, 40) and Poll Dorset (PD, 48). The 10 Chinese indigenous breeds were mainly distributed in 8 provincial administrative regions (Inner Mongolia, Ningxia, Zhejiang, Shandong, Tibet, Yunnan, Xinjiang and Qinghai), basically represented the main sheep husbandry systems in China (Table 1 and Fig. 1). Poll Dorset sheep were from the research farm affiliated with the Institute of Animal Science, Chinese Academy of Agricultural Sciences in Beijing. DNA was extracted following the well-established phenol-chloroform method (Sambrook et al., 1989).

2.2. Microsatellite analysis and genotyping

Nineteen microsatellite markers were used in this study, and all of them were from the panel recommended

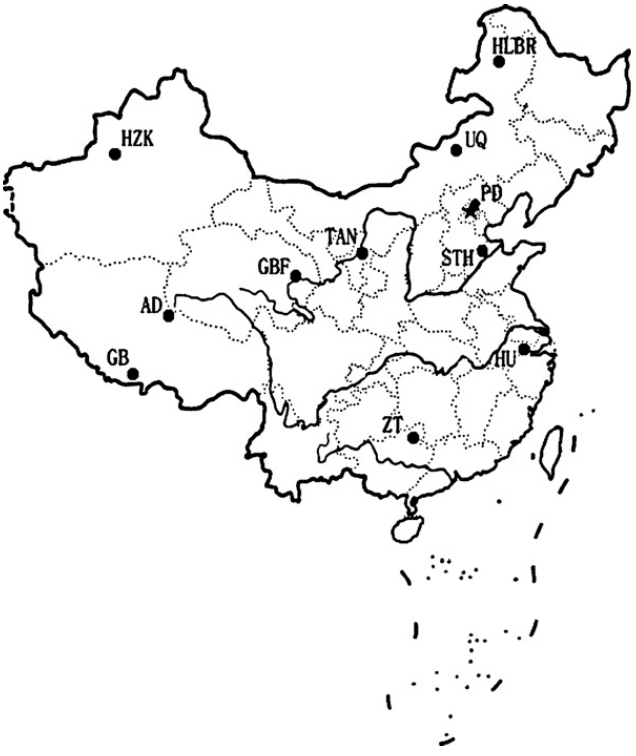

Fig. 1. Geographic distribution of 11 sheep populations in this study. For breed abbreviations, see Table 1.

by ISAG/FAO for sheep diversity studies (http://dad.fap.prg/). Forward primers were end-labeled with fluorescent dyes (6-FAM, VIC, PET or NED, Table S1). PCR was carried out in a total volume of 12 μL system containing 1 μL of template DNA (50 ng/μL), 1.2 μL of 10 × PCR buffer, 1 μL of 20 mM dNTPs, 0.25 μL of each 10 μM primer and 1 units DNA polymerase (Takara, Bio Inc.). The cycling conditions included an initial activation step at 95 °C for 5 min, 30 cycles of 94 °C for 30 s, annealing at 50–65 °C (Tm in Table S1) for 30 s and extension at 72 °C for 30 s, and a final extension at 72 °C for 10 min. PCR was performed on a PTC-200 thermocycler (MJ Research, Inc.). PCR products were genotyped using an ABI PRISM 3130XL DNA Analyzer (Applied Biosystems, USA). Data collection, extraction and allele size analysis were generated from Data Collection Software version 1.0 and GeneMapper^{TM} software Version 3.7 (Applied Biosystems, USA).

2.3. Statistical analyses

An exact test for deviation from Hardy–Weinberg equilibrium (HWE) was performed using the Cervus package (available at http://www.fieldgenetics.com). To assess within-population genetic diversity, the mean number of alleles (*MNA*), the observed (H_O) and the expected heterozygosity (H_E, Nei's unbiased gene diversity) were calculated using the Microsatellite Toolkit (available at http://animalgenomics.ucd.ie/sdepark/ms-toolkit/). Allelic richness (r) was estimated using the FSTAT program version 2.9.3 (Goudet, 1995).

Using the variance-based method of Weir and Cockerham (1984), F-statistics (F_{IS}, F_{IT} and F_{ST}) for calculating overall genetic differentiation among populations and between pairs of populations were performed and tested using the FSTAT with 1000 permutations. Analysis of molecular variance (AMOVA) was computed using the ARLEQUIN package version 3.11 (available at http://cmpg.unibe.ch/software/arlequin3/).

In addition, three methods were used to detect the genetic relationships and population structure among the 11 sheep breeds. Firstly, Nei's D_A genetic distances were calculated using the Dispan program (available from http://iubio.bio.indiana.edu/soft/molbio/ibmpc/dispan), which has a more discriminatory power for closely related groups, and it is the most appropriate genetic distance measure for loci with different rate of neutral mutation (Nei et al., 1983). The neighbor-joining algorithm was used to construct a phylogenetic tree with the Dispan. Tree robustness was evaluated by bootstrapping over loci (1000 replicates). Secondly, as the phylogenetic reconstruction may not really take into account the effect of admixture between populations, Principal Components Analysis (PCA) analysis was further performed as an alternative approach to understand the genetic relationship among populations. A PCA based on allelic frequencies was carried out according to the procedures described by Cavalli-Sforza et al. (1994) using the SPSS package version 13.0 (SPSS Inc.). Thirdly, population structure and degree of admixture were performed using the Bayesian clustering-model program STRUCTURE (Pritchard et al., 2000). The program estimates, using multi-loci genotypes to assign individuals to populations, individual admixture proportions and infers the number of parental populations (K) for a given sample. As suggested by several authors, the analysis involved an admixture model with correlated allele frequencies (Pritchard et al., 2000; Vicente et al., 2008; Zuccaro et al., 2008). The parameter of Ln Pr($X|K$) is an indicator of the posterior probability of the K clusters, as suggested by Pritchard et al. (2000). To obtain a representative value of K, we preformed 50 independent runs for each K ($1 \leq K \leq 11$) and the burn-in time and replication number were both set to 100,000 (Legaz et al., 2008). The estimated individual membership coefficients were graphically visualized using a computer program DISTRUCT (Rosenberg, 2004).

3. Results

3.1. Genetic variability

A total of 257 alleles at 19 microsatellite markers were detected in the 11 sheep breeds. The *MNA* per locus ranged from 7 at *OarCP38*, *SRCRSP5* and *OarAE129* to 30 at *DYMS1*, with an average of 13.53 alleles per locus. The highest polymorphism was detected at *DYMS1* locus (0.900) and the lowest polymorphism was 0.506 at the *SRCRSP9* locus (Table S1).

The H_O values ranged from 0.631 (GBF) to 0.720 (HZK). However, ZT had a higher H_E than HZK (Table 1). Most of the 19 loci were under HWE in all breeds, except *SRCRSP5*, *OarAE129* and *DYMS1* that deviated from HWE in most of the breeds. These three loci were then excluded from the further genetic differentiation analysis. *SRCRSP5* and *OarAE129* displayed heterozygote deficits in this study, which were consistent with Swiss sheep (Glowatzki-Mullis et al., 2009). It was noted that *SRCRSP5* and *OarAE129* had been shown to produce null alleles in other studies (Peter et al., 2005, 2007). While null alleles frequencies of *SRCRSP5* and *OarAE129* were 0.155 and 0.217 in this study, respectively. Unexpectedly, *DYMS1* locus showed an odd motif repeat unit, which was different from others with an even motif repeat unit. Generally, all the microsatellite loci revealed high polymorphisms in the 10 local breeds, and the Mongolian sheep breeds had a little higher genetic variation than the Tibetan sheep breeds.

3.2. Genetic differentiation

The mean estimates of F-statistics obtained by jackknifing over loci were: $F_{IS} = 0.024 \pm 0.020$ (within-population inbreeding estimate), $F_{IT} = 0.087 \pm 0.020$ (total inbreeding estimate), $F_{ST} = 0.065 \pm 0.005$ (estimate of population differentiation). The value of overall F_{IS} was low, but highly significant ($P < 0.001$), probably due to inbreeding within population. All the 16 loci notably influenced the values of F-statistics (Table 2). The average genetic differentiation between all breeds (F_{ST}) was 6.5%, which was significantly different from zero ($P < 0.001$, Table 2). In addition, the AMOVA analysis revealed a high variance component of among populations (95.23%) vs. within populations (4.77%) (Table 3).

Table 2
Wright's *F*-statistics for 16 microsatellite loci in the 11 sheep breeds.

Locus	F_{IS}	F_{IT}	F_{ST}
OarCP38	0.015(0.038)***	0.082(0.034)***	0.070(0.043)***
SRCRSP9	−0.019(0.026)***	0.046(0.051)***	0.064(0.045)***
MAF214	−0.041(0.015)***	0.018(0.025)***	0.056(0.018)***
OarCP34	−0.041(0.023)***	−0.003(0.022)***	0.037(0.011)***
OarVH72	−0.008(0.027)***	0.085(0.032)***	0.093(0.037)***
OarFCB128	0.118(0.063)***	0.202(0.071)***	0.094(0.029)***
OarHH47	0.052(0.017)***	0.083(0.020)***	0.033(0.013)***
MCM527	0.056(0.053)***	0.122(0.044)***	0.070(0.028)***
ILSTS5	−0.041(0.023)***	0.021(0.033)***	0.060(0.040)***
MAF209	−0.021(0.022)***	0.039(0.031)***	0.059(0.031)***
OarJMP29	−0.003(0.030)***	0.075(0.032)***	0.079(0.031)***
OarFCB226	−0.021(0.027)***	0.050(0.046)***	0.070(0.045)***
ILSTS28	0.250(0.041)***	0.299(0.050)***	0.065(0.027)***
BM8125	−0.007(0.038)***	0.024(0.028)***	0.031(0.014)***
ILSTS11	0.025(0.027)***	0.106(0.043)***	0.084(0.040)***
OarFCB304	0.013(0.019)***	0.069(0.022)***	0.057(0.017)***
Muti-locus	0.024(0.020)***	0.087(0.020)***	0.065(0.005)***

*** $P < 0.001$; the numbers in brackets are the standard error.

Table 3
Analysis of molecular variance (AMOVA) of the 11 sheep breeds based on 16 microsatellite DNA variations.

Structure	Source of variation	Degree of freedom	Sum of squares	Squared value	Percentage of variation
11 breeds	Among populations	10	200.677	0.196	6.44
	Within populations	957	2725.176	2.848	93.56
	Total	967	2925.853	3.044	
10 Chinese indigenous breeds	Among populations	9	138.583	0.144	4.77
	Within populations	862	2477.551	2.874	95.23
	Total	871	2616.134	3.018	

3.3. Genetic relationship and population structure analysis

The pair-wise genetic differentiations (F_{ST}) between breeds are shown in Table 4. F_{ST} values ranged from 0.0024 for UQ-HLBR pair to 0.2384 for GBF-PD pair. The Nei's distances (D_A) ranged from 0.0526 between UQ and HLBR to 0.3374 between GBF and PD. HZK had a lower distance with Mongolian group (<0.1) than with Tibetan group (especially in HZK-GBF pair, 0.2263). PD, an introduced breed from Australia, showed the largest genetic distance in comparison with Chinese indigenous breeds, as what we expected.

The neighbor-joining tree showed clearly defined clusters (Fig. 2). Nevertheless, a majority of bootstrapping values was low (< 50%). AD, GB and GBF were found in the same branch with a higher value (78%). The result was well supported by the traditional classification (Tibetan group) and their close distribution in the neighboring areas of Tibetan and Qinghai provinces. A less significant clustering of the Mongolian breeds was observed, and a relatively close relationship was also found in UQ-HLBR pair. The Mongolian group was well further supported by PCA and STRUCTURE analysis.

The groupings identified by PCA were similar to the NJ tree. The first two components in the PCA loosely separated the introduced breed (PD) and one local breed (GBF) from the other breeds (Fig. 3A). Among the Chinese local breeds, the PC1 and PC2 separated the Mongolian group (HU, TAN, UQ, HLBR and STH) from the Tibetan group (AD, GB and GBF) (Fig. 3B). Furthermore, ZT tended to be separated from the Tibetan group. It seemed possible that there was a population subdivision within the Tibetan group. In

Table 4
Pair-wise F_{ST} and genetic distance among the 11 tested breeds in this study.

	UQ	TAN	HU	HLBR	STH	AD	GB	GBF	ZT	HZK	PD
UQ		0.0272	0.0300	0.0024	0.0209	0.0307	0.0345	0.1521	0.0384	0.0156	0.1036
TAN	0.1076		0.0364	0.0234	0.0288	0.0353	0.0438	0.1344	0.0475	0.0304	0.1305
HU	0.0898	0.1129		0.0212	0.0354	0.0392	0.0449	0.1552	0.0533	0.0224	0.1437
HLBR	0.0526	0.0898	0.0835		0.0217	0.0322	0.0369	0.1383	0.0324	0.0060	0.1068
STH	0.0898	0.1126	0.1025	0.0905		0.0400	0.0454	0.1452	0.0420	0.0166	0.1198
AD	0.1104	0.1130	0.1124	0.0873	0.0951		0.0429	0.1319	0.0478	0.0304	0.1331
GB	0.1446	0.1264	0.1400	0.1210	0.1311	0.1047		0.1437	0.0572	0.0387	0.1537
GBF	0.2459	0.2329	0.2432	0.2204	0.2377	0.1883	0.2180		0.1234	0.1396	0.2384
ZT	0.1407	0.1734	0.1689	0.1156	0.1470	0.1523	0.1946	0.2619		0.0325	0.1038
HZK	0.0751	0.0998	0.0812	0.0567	0.0838	0.0957	0.1200	0.2263	0.1222		0.1031
PD	0.2122	0.2420	0.2258	0.1843	0.2088	0.2095	0.2495	0.3374	0.2114	0.1884	

F_{ST} estimates above the diagonal are all significant at $P < 0.05$. Nei's genetic distances (D_A) are below the diagonal.

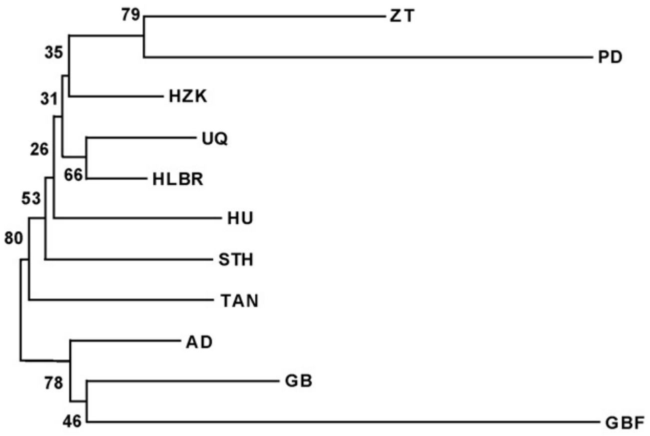

Fig. 2. Neighbor-Joining (NJ) phylogeny based on D_A genetic distances for 11 breeds. Numbers at the nodes are bootstrapping values, based on 1000 permutations. For breed abbreviations, see Table 1.

addition, the heterozygosity of ZT was higher than those of the other Tibetan breeds (H_E or r, Table 1).

In structure analysis, the Ln Pr($X|K$) increased distinctly from $K=2$ to $K=4$ and reached a "plateau" at $K=4$. While the Ln Pr($X|K$) did not show a significant fluctuation from $K=5$ to 11, as illustrated in Fig. 4. Our data set of 11 breeds was well described by $K=4$ clusters. The results are shown in Fig. 5 for K ranging from 2 to 5. Analysis at $K=2$ clusters revealed an obvious difference between introduced and indigenous breeds, which was also consistent with the PCA results, though it was not the most probable number of subpopulations considering the log-likelihood of K. The Tibetan group (AD, GB and GBF) remained the same subpopulation for $K=3$, while ZT was separated from the Tibetan group, indicating a probable introgression of the Mongolian group. At $K=4$ the Tibetan group was further subdivided into two clusters: AD and GB in one cluster and GBF in one independent cluster. GBF had the lowest heterozygosity comparing with the other Tibetan breeds. For $K=5$, the breeds remained with homogeneous contributions of the original populations, therefore the clustering was similar to those observed for $K=4$. Apart from genetic clusters, GBF, PD and ZT remained the most genetically pure breeds comparing with the others. This

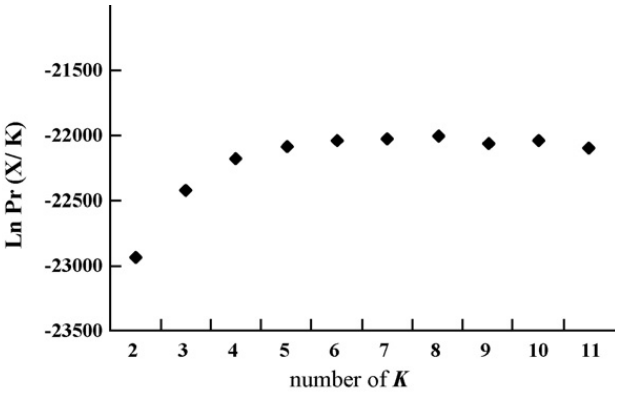

Fig. 4. Estimated posterior probabilities [Ln Pr($X|K$)] for different numbers of inferred clusters (K).

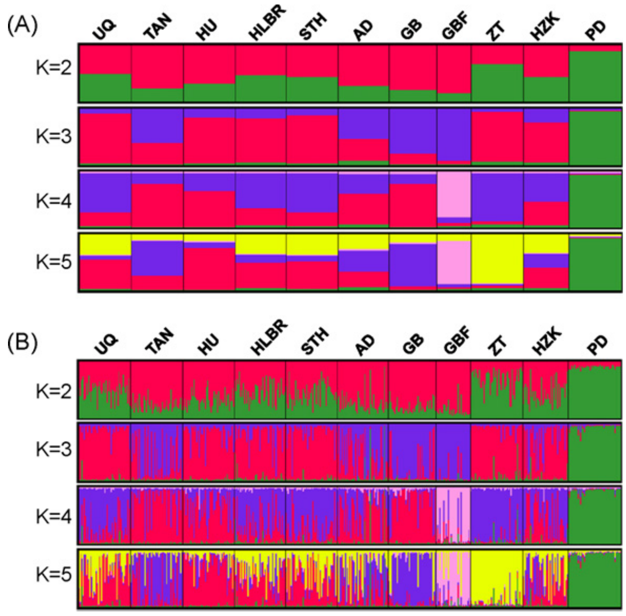

Fig. 5. STRUCTURE estimated population structure of the 11 tested sheep breeds for each of the inferred K clusters. (A) Estimated population structure displayed with population Q-matrix. (B) Estimated population structure displayed with individual Q-matrix. Black lines separated the different populations. For breeds abbreviations, see Table 1.

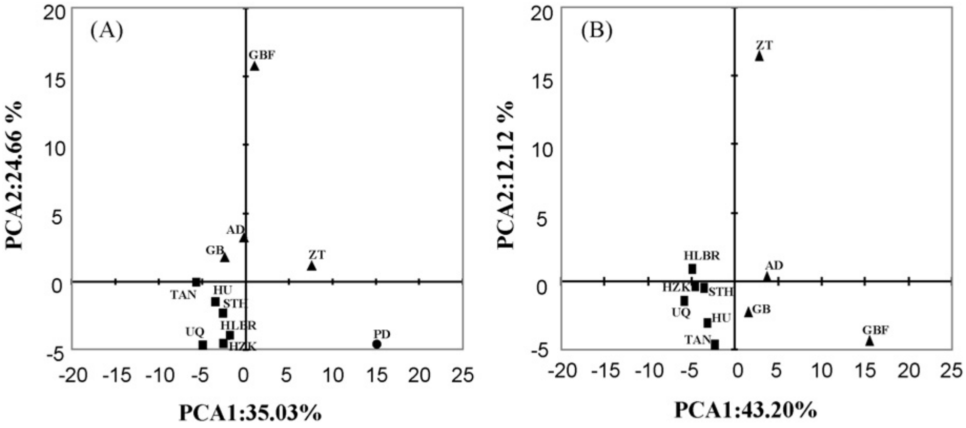

Fig. 3. Principal component analysis. The principal components were extracted from averaged covariance matrices based on allele frequencies. (A) PCA analysis of all the 11 breeds. Projection on PC 1 and PC 2, which cumulatively explained 30.61% of the total inertia contained in the data set. (B) PCA analysis of the 10 Chinese indigenous breeds. For breeds abbreviations, see Table 1.

may have been achieved through different breeding history leading to the reduced genetic diversity in BGF and PD, while ZT might be caused by geographical isolation and different domestication history. Admixtures were found between HU-TAN and AD-GB, and several GBF individuals contained Mongolian breeds' multi-loci genotypes (Fig. 5B).

4. Discussion

The investigated 10 Chinese indigenous breeds were distributed in the main sheep habitats, and represented the main sheep types in China. The genotype data indicated that significant amounts of genetic variations are still maintained in these indigenous breeds (Table 1). The mean number of alleles and heterozygosities in this study were similar to the values reported for other domestic sheep populations (Tapio et al., 2005; Calvo et al., 2006). A clear decrease of genetic diversity from the North to the South of China with the lowest diversity for the Tibetan breeds (AD, GB and GBF) was observed, with ZT being a clear exception. The smallest genetic distance was observed in the UQ-HLBR pair. Both of UQ and HLBR are the main breeds reared in the prairie of Inner Mongolia. HLBR, which is fat-rumped and different from the other Mongolian breeds (short fat-tailed), is a newly developed breed for meat, whereas it has a high genetic diversity ($MNA = 9.13$, $HE = 0.723$ and $r = 7.458$). Although a high genetic variation was observed, the genetic differentiation (F_{ST}) among the 10 Chinese breeds only amounted to 4.8%, which indicated that genetic variation was mainly present within -populations. The result was a little lower than those from Northwestern Indian (8.3%) (Sodhi et al., 2003), Austrian (8%) (Brenneman et al., 2007) and European & Middle Eastern sheep (5.7%) (Peter et al., 2007), while higher than that of Spanish sheep (3.7%) (Calvo et al., 2006). Allelic richness was considerably higher (5.169–7.610) than those in northern European sheep breeds (2.62–6.26) (Tapio et al., 2005).

In terms of morphology (tail type), the 10 indigenous breeds are divided into three groups: short fat-tailed (UQ, TAN, HU and STH), fat-rumped (HLBR, HZK) and thin-tailed (AD, GB, GBF and ZT) (Table 1). Interestingly, the distribution of tail types was largely in accordance with genetic relationships between the 10 breeds (Fig. 2). The result was also consistent with the traditional classification proposed earlier by some experts, which indicated that breed classification should rely on the phenotype first, rather than geographic location (Zhang, 1986).

However, classification only based on the phenotypic information, breed history or geographic distribution, might not reflect the real underlying genetic relationship between the sheep populations. Clusters derived from genetic study were not only consistent with the current classification and morphological characters, but also contributed to exploit the differences among breeds and hidden substructures within populations. In the NJ tree, the thin-tailed sheep (AD, GB and GBF) clustered together, while another thin-tailed sheep (ZT) was divided into another group with PD. In the Bayesian analysis, ZT revealed more Mongolian genetic background than the other Tibetan breeds did (Fig. 5). It could be explained by geographic isolation with other Tibetan breeds (Himalayas) and crossbreeding improvement for semi-fleece as early as 1960s with Mongolian breeds. In addition, GBF having a unique black fur coat was highly differentiated from other breeds (Table 4, Figs. 3 and 5). This scenario could be explained by geography isolation, different breeding history and no introgression or genetic admixture from other breeds. Although HU, TAN and STH were distributed in different provinces in China, they belonged to Mongolian breeds sharing the same tail type (short fat-tailed), and all originated from the ancient Mongolian sheep. HU reared in the Zhejiang province located in the downstream of the Yangtze River, it is well adapted to local environmental conditions (humid climate). Famous of being a genetically stable breed for fur and its meat is also of good quality. The breed's history could be traced back to Tang dynasty (618-907 AD) and its products were famous tributes in Song (960–1132 AD) and Ming (1368–1644 AD) dynasty (Animal Bureau of Ningxia Province, 1984; Xie, 1985). STH was migrated with ancient minorities like Khitan, Nuchen and Mongolian to the middle land of ancient China in Song dynasty, developed into a breed of fur and meat of good quality (Animal Bureau of Shandong Province, 1999).

In conclusion, the phylogenetic clusters, PCA and structure analysis pointed out a complicated genetic structure of Chinese sheep breeds under the effects of different breeding histories, geographic distributions and ecological factors. This study contributes to the knowledge of the genetic structure and molecular characterization of Chinese local sheep populations. To protect the indigenous sheep breeds effectively, both the within- and between-populations variability should be included in conservation decisions.

Acknowledgements

We thank the anonymous reviewers for their comments on an early version of the manuscript. This work was supported in part by the Natural High Technology Research and Development Program (863) of China (grant no. 2006AA10Z198) and International Atomic Energy Agency (130071RO).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.smallrumres.2010.02.001.

References

Animal Bureau of Ningxia Province, 1984. Breeds of Domestic Animal and Poultry in Ningxia (in Chinese).

Animal Bureau of Shandong Province, 1999. Breeds of Domestic Animal and Poultry in Shandong. Shenzhen Haitian Publishers (in Chinese).

Brenneman, R.A., Chase Jr., C.C., Olson, T.A., Riley, D.G., Coleman, S.W., 2007. Genetic diversity among Angus, American Brahman, Senepol and Romosinuano cattle breeds. Anim. Genet. 38, 50–53.

Calvo, J., Bouzada, J., Jurado, J., Serrano, M., 2006. Genetic substructure of the Spanish Manchega sheep breed. Small Rumin. Res. 64, 116–125.

Cavalli-Sforza, L.L., Menozzi, P., Piazzi, A., 1994. The History and Geography of Human Genes. Princeton University Press, Princeton.

Chen, S.Y., Duan, Z.Y., Sha, T., Xiangyu, J., Wu, S.F., Zhang, Y.P., 2006. Origin, genetic diversity, and population structure of Chinese domestic sheep. Gene 376, 216–223.

Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, I., Kao, R.R., Pemberton, J.M., Beraldi, D., Stear, M.J., Alberti, A., Pittau, M., Iannuzzi, L., Banabazi, M.H., Kazwala, R.R., Zhang, Y.P., Arranz, J.J., Ali, B.A., Wang, Z., Uzun, M., Dione, M.M., Olsaker, I., Holm, L.E., Saarma, U., Ahmad, S., Marzanov, N., Eythorsdottir, E., Holland, M.J., Ajmone-Marsan, P., Bruford, M.W., Kantanen, J., Spencer, T.E., Palmarini, M., 2009. Revealing the history of sheep domestication using retrovirus integrations. Science 324, 532–536.

Glowatzki-Mullis, M.L., Muntwyler, J., Baumle, E., Gaillard, C., 2009. Genetic diversity of Swiss sheep breeds in the focus of conservation research. J. Anim. Breed Genet. 126, 164–175.

Goudet, J., 1995. Fstat (version 1.2): a computer program to calculate Fstatistics. Journal of Heredity 86, 485–486, Version 2.9.3 available from http://www.unil.ch/izea/softwares/fstat.html.

Guo, J., Du, L.X., Ma, Y.H., Guan, W.J., Li, H.B., Zhao, Q.J., Li, X., Rao, S.Q., 2005. A novel maternal lineage revealed in sheep (*Ovis aries*). Anim. Genet. 36, 331–336.

Legaz, E., Álvarez, I., Royo, L.J., Fernández, I., Gutiérrez, J.P., Goyache, F., 2008. Genetic relationships between Spanish Assaf (Assaf.E) and Spanish native dairy sheep breeds. Small Rumin. Res. 80, 39–44.

Ma, Y.H., Rao, S.Q., Lu, S.J., Hou, G.Y., Guan, W.J., Li, H.B., Li, X., Zhao, Q.J., Guo, J., 2006. Phylogeography and origin of sheep breeds in Northern China. Conserv Genet. 7, 117–127.

Ma, Y.H., Xu, G.F., Wa, D.Y., Liu, H.L., 2002. Study on dynamic information of animal genetic resources in China. Sci. Agric. Sin. 35, 552–555.

Nei, M., Tajima, F., Tateno, Y., 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J. Mol. Evol. 19, 153–170.

Peter, C., Bruford, M., Perez, T., Dalamitra, S., Hewitt, G., Erhardt, G., 2007. Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. Anim. Genet. 38, 37–44.

Peter, C., Prinzenberg, E.M., Erhardt, G., 2005. Null allele at the OarAE129 locus and corresponding allele frequencies in German sheep breeds. Anim. Genet. 36, 92.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Rosenberg, N.A., 2004. Distruct: a program for the graphical display of population structure. Mol. Ecol. Notes 4, 137–138.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbour, Cold Spring Harbour.

Sodhi, M., Mukesh, M., Arora, R., Tantia, M.S., Bhatia, S., 2003. Genetic structure of Garole—a unique Indian micro sheep assessed using microsatellite markers. Indian J. Dairy Sci. 56, 167–173.

Tapio, M., Tapio, I., Grislis, Z., Holm, L.E., Jeppsson, S., Kantanen, J., Miceikiene, I., Olsaker, I., Viinalass, H., Eythorsdottir, E., 2005. Native breeds demonstrate high contributions to the molecular variation in northern European sheep. Mol. Ecol. 14, 3951–3963.

Vicente, A.A., Carolino, M.I., Sousa, M.C., Ginja, C., Silva, F.S., Martinez, A.M., Vega-Pla, J.L., Carolino, N., Gama, L.T., 2008. Genetic diversity in native and commercial breeds of pigs in Portugal assessed by microsatellites. J. Anim. Sci. 86, 2496–2507.

Weir, B.S., Cockerham, C.C., 1984. Estimating *F*-Statistics for the analysis of population structure. Evolution 38, 1358–1370.

Xie, C.X., 1985. History of Raising Cattle, Sheep and Goat in China (Attached History of Raising Deer). Agricultural Publishing House of China (in Chinese).

Zhang, Z.G., 1986. Pig Breeds in China. Shanghai Scientific and Technical Publishers, Shanghai (in Chinese).

Zheng, P.L., 1988. Sheep and Goat Breeds in China. Shanghai Scientific & Technology Publishers, Shanghai (in Chinese).

Zuccaro, A., Bordonaro, S., Criscione, A., Guastella, M., Perrotta, G., Blasi, M., D'Urso, G., Marletta, D., 2008. Genetic diversity and admixture analysis of Sanfratellano and three other Italian horse breeds assessed by microsatellite markers. Animal 2, 991–998.