Journal of
Animal Breeding and Genetics

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

The phylogeographic system survey of native sheep breeds in the eastern and southern Central Asia

W. Sun[1,2], H. Chang[2], K. Tsunoda[3], H.H. Musa[2,4], Z.P. Yang[2], Y.H. Ma[1] & W.J. Guan[1]

1 Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China
2 College of Animal Science and Technology, Yangzhou University, Yangzhou, China
3 Showa University School of Medicine, Hatanodai, Tokyo, Japan
4 Faculty of Veterinary Science, University of Nyala, Nyala, Sudan

Keywords
Eastern and Southern of Central Asia; genetic differentiation; native sheep; phylogenetic group.

Correspondence
Prof. Y. Ma, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100094, China. Tel: 86-10-62813463; Fax: 86-10-62813463; E-mail: yuehui.ma@263.net

Prof. H. Chang, College of Animal Science and Technology, Yangzhou University, Yangzhou 225009, China. Tel:86-514-87997203; Fax:86-514-87350440; E-mail: dkxmsunwei@163.com

Received: 02 November 2008;
accepted: 18 September 2009

Summary

The genetic diversity and phylogenetic survey of native sheep breeds in the eastern and southern Central Asia were assessed in the present study. The clustering, principal components, structure and *F* statistics all demonstrate that the native sheep breeds in these regions be classified into two genetic groups: Mongolia-Tibetan sheep group and South-Southeast Asia sheep group. The Mongolia sheep group and the Tibetan sheep group had a certain degree of gene communication from the ancient times. In the present study we demonstrated that the Chinese native sheep populations belonged to Mongolia-Tibetan sheep group. However, the relationships among the sheep populations in Mongolia sheep group in China were not closely related to the geographical distance among sheep populations.

Introduction

The maintenance of domestic animal diversity has been emphasized in several reports (Hall & Bradley 1995; Kantanen *et al.* 1999). Domestic animal diversity is critical for food security and essential to meet unpredictable future requirements. To evaluate genetic diversity of domestic animal breeds, statistical measures derived from Wright's *F* statistics (Wright 1951) or phylogenetic techniques based on genetic distances estimated from polymorphic markers (Hall & Bradley 1995) have been the methods of choice. Recently, Bayesian model-based clustering methods have been proposed, which allow for the inference of population structure and the assignment of individuals to populations (Pritchard *et al.* 2000; Corander *et al.* 2003). Domestic sheep (*Ovis aries*) have played important roles in diverse human societies as a source of food, hide and wool, and are one of the major components of agro-pastoral societies since the Neolithic. The phylogenetic study of sheep breeds provides useful information on the evolution of breeds, gene pool development and magnitude of genetic differentiation (Kantanen *et al.* 2000).

The definition of a breed, as applied by the Food and Agriculture Organization of the United Nations (FAO), is based on the homogeneity of external characteristics, or on a generally accepted identity of animals of a geographically or culturally separated group (FAO 1998). China has a centuries-old history of breeding sheep, with more than 40 local sheep breeds, which serve as important genetic resources

  doi:10.1111/j.1439-0388.2009.00843.x

for the sustainable development of animal production and the preservation of biological diversity (Wang *et al.* 2006). These breeds are distributed from the high Qinghai–Tibet Plateau to the lowland of East China (Tu *et al.* 1989). Based on the morphology and the distribution of wild sheep populations of *O. ammon*, it was hypothesized that Chinese domestic sheep derived from wild sheep populations of *O. orientalis* and *O. ammon* (Tu 1989). However, molecular phylogenetic analyses have excluded *O. ammon* as the maternal ancestor of domestic sheep (Hiendleder *et al.* 1998, 2002; Wu *et al.* 2003). The polymorphic analysis of sheep populations in East- and South- Asia was done based on morphology and blood protein, and this provided important data to elucidate the phylogenetic group of sheep populations from these regions (Tsunoda *et al.* 1988, 1992, 1993, 1995, 1998a,b, 1999, 2004).

This study is important for decisions concerning breed conservation; breeds with a unique evolutionary history could potentially have a value in the maintenance of genetic diversity at the species level (Hall & Bradley 1995). We used Hu sheep, Tong sheep, Tan sheep, small-tailed Han sheep and Wadi sheep as subjects to explore the genetic structure and relationships within the Mongolia sheep group and put forward further rational information for the utilization and protection of these breeds. And this study may also put forward some information for the sustainable development of sheep production and the preservation of biological diversities of native sheep breeds in the eastern and southern Central Asia.

Materials and methods

Materials

Blood samples from 337 domestic sheep representing five sheep populations were collected from different location throughout China according to the simple random sampling method. These breeds were (Hu sheep, n = 63; Tong sheep, n = 65; Small-tailed Han sheep, n = 60; Tan sheep, n = 73 and Wadi sheep, n = 76). Blood samples were transferred to the Animal Genetics Laboratory, College of Animal Science and Technology, Yangzhou University, China, and stored at −20°C. In addition, 10 Asian native sheep populations detected by foreign scholars were used as a reference; the origin of population is illustrated in Table 1. Geographical distribution of 15 native sheep breeds in the Eastern and Southern Central Asia is shown in Figure 1.

Blood protein and non-protein typing

Eighteen blood protein and two non-protein loci include: albumin (Alb), gc-protein (Gc), transferrin (Tf), alkaline phosphatase (Alp), leucine aminopeptidase (Lap), arylesterase (Ary-Es), haemoglobin-α (Hb-α) and haemoglobin-*β* (Hb-*β*), x-protein (X-p), carbonic anhydrase (CA), catalase (Cat), malate dehydrogenase (MDH), slow-α2-macroglobulin (α2-M), ceruloplasmin (Cp), nadh-diaphorase I (Dia-I) and nadh-diaphorase-II (Dia-II), glucosephosphate isomerase (GPI), and esterase D (EsD), lysine (Ly) and potassium (Ke) were determined electrophoretically and ion-densitometrically. The starch-gel, polyacrylamide-gel, cellulose acetate film electrophoresis and polyacrylamide-gel isoelectric focusing were used. The staining procedures for separating and detecting the blood protein and non-protein components were conducted according to the standards universally accepted in the neighbouring countries of China (Tsunoda *et al*. 1988, 1990, 1992, 1995, 1998a, 1999, 2004; Seiki *et al.* 1989; Gahne *et al.* 1977).

Statistical analysis

Gene frequencies of polymorphic loci were computed by two methods, the square root method for dominant locus: Ary-Es, Alp, Lap, Ly, X-p and Ke loci, and gene counting method for co-dominant locus: Alb, Tf, Hb-*β*, MDH, Cat, EsD, Gc, CA, Dia-I, Dia-II, Hb-α, Cp, α_2-M loci (Tsunoda *et al.* 1990). The reliability was calculated using the following formula, which ensures that the estimate does not deviate from the true value more than 0.5 times (β), and the relative deviation when the reliability reaches 0.9545(η) (Chang 1995).

$$\beta = \int_0^{\lambda} \frac{2e^{-\lambda^2}}{\sqrt{2\pi}} d\lambda \qquad \eta = 2\left[V(p)^{\frac{1}{2}}\right] P^{-1}$$

In this formula, P and $V(p)$ represents gene frequency and its gene variance, respectively; λ is the standard error of deviation of the estimate, calculated based on the following formula.

$$\lambda = 0.5 \div [V(p)]^{\frac{1}{2}}$$

Genetic differentiation analysis was carried out by four methods:

First, Nei genetic distance (Nei 1978) was calculated among 15 sheep populations, and consensus

Table 1 Summary of 15 native sheep populations in the region of East and South of central Asia

Population	Abbreviation	Number	Sampling location	Number of individuals	Reference
Kharkhorin sheep	Kh	1	The suburbs of Kharkhorin in Central Mongolia	99	Tsunoda *et al.* 1999
Ulaanbaatar sheep	Ub	2	The suburbs of Ulaanbaatar in Central Mongolia	97	Tsunoda *et al.* 1999
Small-tailed Han sheep	Han	3	Liangshan, Shandong Province of China	60	This study, 2008
Hu sheep	Hu	4	Huzhou, Zhejiang Province of China	63	This study, 2008
Tong sheep	Tong	5	Baishui, Shanxi Province of China	65	This study, 2008
Tan sheep	Tan	6	Yanchi, Ningxia Province of China	73	This study, 2008
Wadi sheep	WD	7	Dongying, Shandong Province of China	76	This study, 2008
Yunnan sheep	Yunnan	8	Lufeng and Lunan,Yunnan Province of China	35	Tsunoda *et al.* 1995
Vietnamese Cham sheep	Cham	9	The Ninh Son district in Ninh Thuan Province of Vietnam	34	Tsunoda *et al.* 1998b
Bangladeshi sheep	Ban	10	Jessore, Khulna, Mymensingh and Noakhali area of Bangladesh	76	Tsunoda *et al.* 1988
Bhyanglung sheep	Bhy	11	Kathmandu city and Ghara area of Nepal	41	Tsunoda *et al.* 1992, 1993, 1995
Baruwal sheep	Bar	12	Solu area, Kathmandu city and Ghara are of Nepal	43	Tsunoda *et al.* 1992, 1993, 1995
Kagi sheep	Kag	13	Kathmandu city and Chitlang farm of Nepal	41	Tsunoda *et al.* 1992, 1993, 1995
Lampuchre sheep	Lam	14	Narayangath, Somnath-parasi and Lumbini area of Nepal	22	Tsunoda *et al.* 1992, 1993, 1995
Myanmar sheep	c-Mya	15	Bagan, Marway and Mandalay of Myanmar	124	Tsunoda *et al.* 2004

tree UPGMA based on Bayesian clustering model was reconstructed and was implemented in NTSYS-PC software package version 2.1 (Exter Software, Setauket, NY, USA) (Rohlf 2000). Second, a principal component analysis (PCA) based on the covariance matrix of gene frequencies was performed using SAS statistical package version 6.12 (SAS Institute Inc., Cary, NC, USA). This analysis absorbs the information of allele frequencies into small number of synthetic variables; it was performed because

Figure 1 Geographical distribution of 15 native sheep breeds in the eastern and southern Central Asia.

genetic distances do not take into account the effect of reticulation between branches. The principal component values were used to draw two dimensions (2D) and three dimension (3D) scatter plot charts implemented in NTSYS-PC software package version 2.1 (Rohlf 2000).

Third, STRUCTURE version 2.0 (Department of Human Genetics, University of Chicago, Chicago, IL, USA) (Pritchard *et al.* 2000) was implemented to determine the most likely number of partitions in the dataset and independence of breed assignment. To estimate the number of clusters (K), two independent runs of K = 2 to 15 were carried out at 10 000 MCMC repetitions and 1 000 000 burn-in period with the admixture model. The most likely number of K was estimated by comparing the log-likelihood of each K-value. Populations were then assigned to a subpopulation based on a membership probability $q \geq 50\%$. The output of STRUCTURE was visualized by the DISTRUCT program (http://rosenberglab.bioinformatics.med.umich.edu/distruct.html).

Fourth, for dominant locus, total heterozygosity (H_T), genetic diversity within populations (H_S), and genetic differentiation among populations (G_{ST}) were estimated according to Nei (1973), all the assessed genetic parameters (H_T, H_S and G_{ST}) were computed using DISPAN software packages (Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, PA , USA) (Ota 1993). For co-dominant locus, population subdivision was examined using Weir and Cockerhams unbiased estimator of Wright's fixation index (Weir & Cockerham 1984), all *F*-statistics (F_{IS}, F_{IT} and F_{ST}) were computed using POPGENE version 1.31 software packages (Department of Renewable Resources, University of Alberta, Edmonton, Canada) (Yeh *et al.* 1999).

Results

Gene frequencies and their reliability and precision

The estimates of gene frequencies, reliability and precision for the five sheep populations are presented in (Supplementary Tables S1–S5). Among the 20 loci tested, Al, Cp, Ca, EsD, α2M, Hb-α, Dia-II and GPI loci were found non-polymorphic.The most polymorphic loci were detected in Hu sheep (13 loci) and Tong sheep (12 loci). Gene frequency was significantly variable among the five sheep populations studied as indicated in supplementary tables (Tables S1–S5).Generally, using simple random sampling methods in typical sheep colony, the estimates of gene frequencies had a good reliability; hence the samples could be used to evaluate the breed features.

Genetic distance and phylogenetic reconstruction

The UPGMA consensus tree was constructed among 15 sheep populations based on Nei genetic distance (Figure 2). It could be seen from the results that Ub and Kh sheep were clustered first, and Tong sheep was clustered with Tan sheep followed by Hu sheep, while Han sheep showed a close relationship with WD sheep. The clustering of five main sheep breeds in China with Ub and Kh sheep, demonstrated that they had a close relationship with Mongolia sheep.

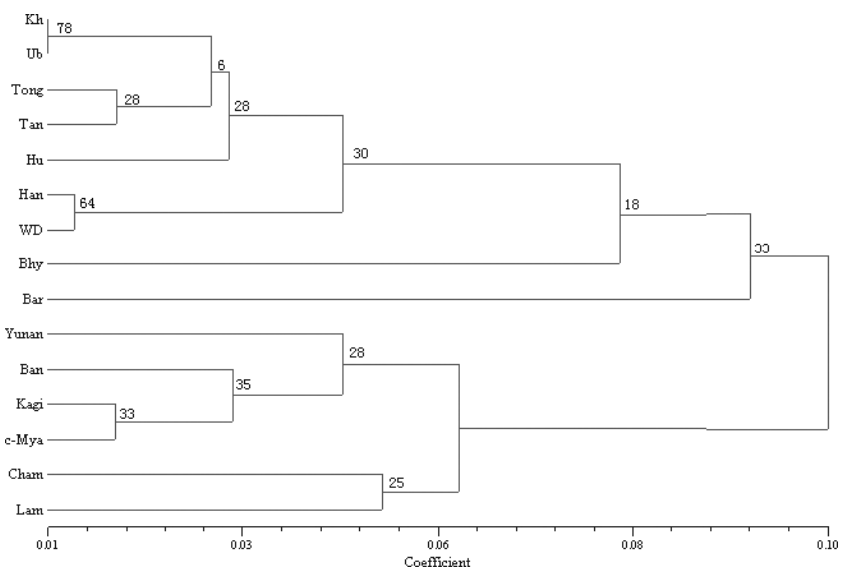

Figure 2 The dendrogram showing the genetic relationships among 15 native sheep populations in the eastern and southern Central Asia. Numbers at the nodes are the bootstrap values of the 20 blood protein and non-protein genotyped.

Our results proved that the 15 sheep populations in the eastern and southern Central Asia could be classified into two groups: Mongolia-Tibetan sheep group (Ub, Kh, Hu, Han, Tong, Tan, WD, Bhy and Bar sheep), and South-Southeast Asia sheep group (Yunnan, Ban, Kagi, C-Mya, Cham and Lam). The Mongolia sheep group had a closer relationship with Tibetan sheep group, because they belong to the same ancestor compared with South -and Southeast Asia sheep.

Principal components analysis of sheep populations

Eigenvalue and accumulative contribution rate were calculated by principal components analysis based on the Nei genetic distance among 15 populations.

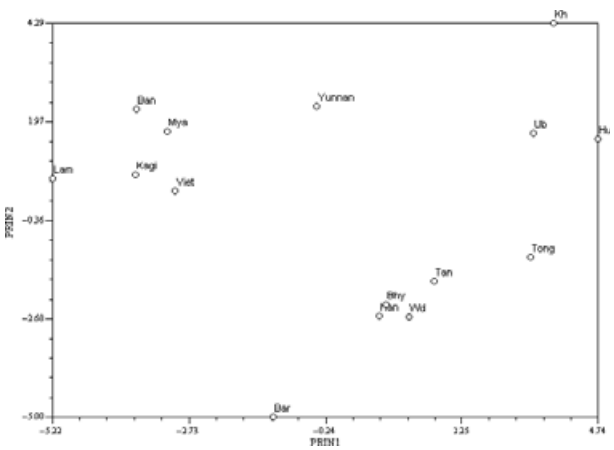

Figure 3 2D Scatter plot showing the first two principal components (PRIN1 and PRIN2) of 15 sheep populations in the eastern and southern Central Asia.

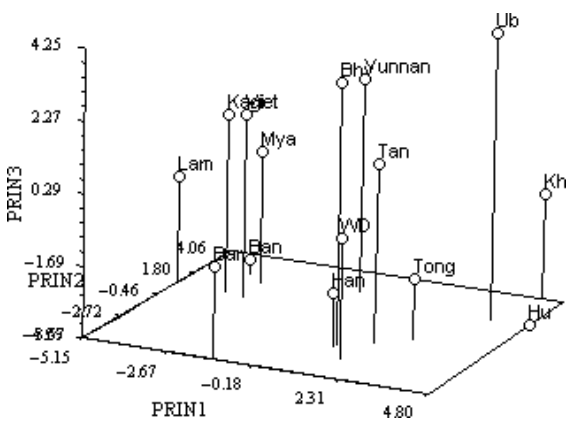

Figure 4 3D Scatter plot showing the first three principal components (PRIN1, 2 and 3) of 15 sheep populations in the eastern and southern Central Asia.

The accumulative contribution rate of the first eight principal components was up to 86%, among which the first three principal components were 50%. The first, the second and the third explained 23, 14 and 12% of total variation, respectively. The principal components of 15 sheep populations were calculated, and then the 2D and 3D scatter plot charts were drawn based on the principal component values (Figures 3 and 4).

The analysed 15 sheep populations were classified into two groups: South-Southeast Asia sheep group (Yunnan, Ban, Kagi, C-Mya, Cham and Lam) and Mongolia-Tibetan sheep group (Ub, Kh, Hu, Han, Tong, WD, Bhy and Bar sheep) (Figure 2). From the 2D chart, we could not differentiate Mongolia sheep from Tibetan sheep, which indicated that Mongolia sheep was closer to Tibetan sheep compared with South-Southeast Asia sheep. While from 3D chart in Figure 4 we could find that the first principal component separated Ban, Kagi, C-Mya, Cham and Lam from other sheep populations, and also separated Ub, Kh, Han, Tong, WD, Bhy and Tan sheep from other sheep populations, while Bar and Yunnan sheep were in the middle of both groups but closer to the latter. The second principal component separated Bar, Han, Tong, WD, Bhy and Tan sheep from other sheep populations. The third one separated Hu, Han, Tong and Ban sheep from other sheep populations and also separated Ub, Bhy and Yunnan sheep from other sheep populations and cluster the

Table 2 The estimated membership probabilities of 15 sheep populations to different clusters

	Inferred clusters		Inferred clusters			
Give pop	1	2	1	2	3	Number of individuals
Hu	0.254	0.746	0.727	0.170	0.103	63
Tong	0.339	0.661	0.601	0.239	0.160	65
Han	0.322	0.678	0.395	0.452	0.154	60
Tan	0.217	0.783	0.472	0.435	0.093	73
WD	0.287	0.713	0.239	0.632	0.129	76
C-Mya	0.855	0.145	0.121	0.082	0.798	116
Ub	0.257	0.743	0.557	0.322	0.120	100
Kh	0.300	0.700	0.454	0.387	0.159	100
Cham	0.773	0.227	0.069	0.325	0.606	34
Yunnan	0.493	0.507	0.286	0.438	0.276	35
Ban	0.838	0.162	0.143	0.081	0.776	76
Bar	0.311	0.689	0.267	0.578	0.155	43
Bhy	0.248	0.752	0.067	0.878	0.055	41
Kagi	0.807	0.193	0.084	0.196	0.720	41
Lamp	0.867	0.133	0.046	0.149	0.805	22

Numbers in bolds indicated the maximum values.
For the abbreviations of sheep populations refer to Table 1.

others in the middle. The 2D and 3D scatter plot charts reflect the genetic diversity and relationships among sheep populations. In the present study the relationships among sheep populations was not consistent with cluster results in Figure 1, because the first three principal components contributed only 50.172% of the total variations.

Breed assignment

The results of the estimated membership probabilities of 15 sheep populations to different clusters were calculated. The genetic composition of the 15 populations was studied using STRUCTURE package with K = 2, 3,..., 15. The results of the estimated membership probabilities of 15 sheep populations to two clusters and three clusters are shown in Table 2 and Figure 5. The 15 populations were classified into two group when K = 2: Mongolia-Tibetan sheep group (Yunnan, Ub, Kh, Hu, Han, Tong, WD, Bhy and Bar sheep) and South-Southeast Asia sheep group (Ban, Kagi, C-Mya, Cham and Lam), the classification is same as in Figure 2. Based on STRUCTURE analysis, Yunnan sheep was classified to Mongolia-Tibetan sheep when the probability was 0.5107 and to South-Southeast Asia sheep when the probability was 0.493. However, when K = 3, the 15 populations were classified into three groups, Mongolia sheep group, Tibetan sheep group South-Southeast Asia sheep group. Among the Mongolia sheep group, the probabilities of Tong and Hu (0.73 and 0.60, respectively) were higher than Tan (0.47), indicating the close relationship between Tong and Hu. Yunnan sheep was classified to Mongolia-Tibetan sheep when the probability was 0.51 and to South-Southeast Asia sheep when the probability was 0.49, and the similarity of both values inferred a complex bloodline of Yunnan sheep. When K = 3, Ban, Kagi, C-Mya, Cham and Lam were grouped together, but the Mongolia-Tibetan sheep diverged: Bar, Bhy and Yunnan sheep were grouped together, while Hu, Tong, Tan sheep were clustered together, Kh with Ub sheep were clustered, and Han with WD sheep were clustered, then the three clusters were grouped together. Yunnan sheep was grouped together with Bar and Bhy sheep when K = 3, indicating that the bloodline of Yunnan sheep was influenced by Tibetan sheep.

F-Statistics of co-dominant and dominant locus in sheep population

The *F*-statistics of the eight co-dominant loci across populations showed that the 18% of the variations

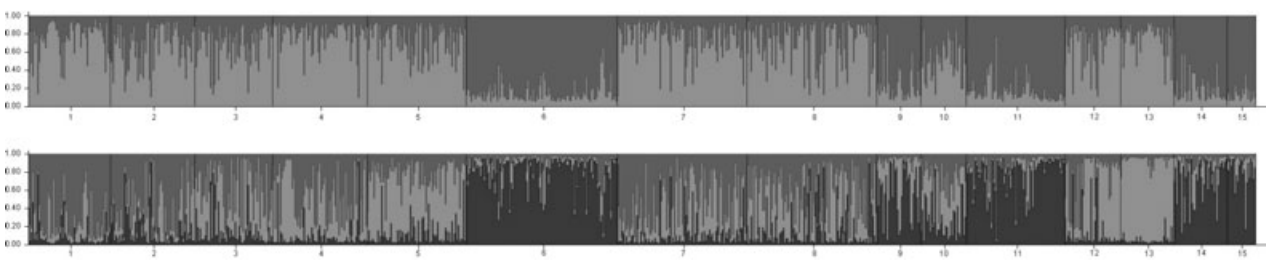

Figure 5 structure analysis of five Chinese indigenous sheep populations in the eastern and southern Central Asia, with K = 2 and 3.

Table 3 The estimated *F*-Statistics of co-dominant and dominant locus of sheep populations

Loci								
	Co-dominant locus				Dominant locus			
Group	Sample size	Mean *F*is	Mean *F*it	Mean *F*st	Sample size	Mean H_T	Mean H_S	Mean G_{ST}
Mongolia sheep group	1062	0.0895	0.1414	0.0571	529	0.3227	0.3044	0.0568
Tibetan sheep group	168	0.1914	0.3807	0.2341	84	0.2167	0.1960	0.0957
South-Southeast Asia sheep group	648	0.0779	0.2052	0.1381	324	0.3261	0.2882	0.1162
Multiple population	1878	0.0937	0.2600	0.1834	937	0.3334	0.3099	0.0703

*F*is, inbreeding coefficient at the population level; *F*it, inbreeding coefficient at the total sample level; *F*st, proportion of differentiation among populations.

H_T, total genetic diversity; H_S, mean diversity within populations; G_{ST}, amount of gene diversity within the populations and between populations.

was inter-population, while 82% of the all variations were due to intra-population genetic variations. Among the three groups the highest (*F*st = 0.23) gene divergence coefficient was found in Tibetan sheep group, while the lowest (*F*st = 0.06) was found in Mongolia sheep group (Table 3). The gene divergence level of the 15 sheep populations was generally low (G_{ST} = 0.07) when *F*-statistics for six dominant loci across populations are estimated. The highest inter-population variation was (G_{ST} = 0.12) among South-Southeast Asia sheep group, while the lowest was (G_{ST} = 0.06) among Mongolia sheep group (Table 3). Combining the data from both loci, we found that the inter-population variation of Mongolia sheep group was smaller, while the inter-population variation was higher among South-Southeast Asia sheep and Tibetan sheep group.

Discussion

The phylogeny of sheep populations in the East and South Asia

In the present study sheep populations in Eastern and Southern Central Asia were grouped into Mongolia-Tibetan sheep group and South-Southeast Asia sheep group. Tsunoda *et al.* (1990) grouped the sheep of these regions into Mongolian group, Tibetan group and Indo-Pakistan group. Later they detected four sheep breeds including Bhy, Bar, Lam, and Kagi, in which Bhy and Bar were described as 'Tibetan Sheep group', while Lam and Kagi were described as 'Indo-Pakistan sheep group'(Tsunoda *et al.* 1993). In the present study we described Lam and Kagi as a member of 'South-Southeast Asia sheep group', because the Indo-Pakistan sheep group proposed by Tsunoda *et al.* (1993) did not include Myanmar and Vietnam sheep. Moreover, the further investigation also revealed that their population data lacked the Mongolia sheep group of China. Further, they suggested that at least two large and phylogenetically different groups existed in East Asia, located in the boundary of Himalayas (Tsunoda *et al.* 2006). In previous study we showed that South Asia group contained both Tibetan sheep and South-Southeast Asia sheep group (Sun *et al.* 2002), while in the present study we found that the Mongolia sheep group and Tibetan sheep group had a certain degree of gene communication. *F*-statistic analysis also showed that there were higher inter-population variation between South-Southeast Asia sheep and Tibetan sheep. Therefore, it is more reasonable to classify the sheep populations in the eastern and southern region of Asia into Mongolia -Tibetan sheep group and South-Southeast Asia sheep group.

Geng *et al.* (2007) demonstrated that the 15 native sheep populations in East and South Asia can be divided into two groups using 13 blood protein and non-protein loci. Lu *et al.* (2005) indicated that small-tailed Han sheep and Hu sheep were from Mongolian sheep. Hu sheep, small-tailed Han sheep, and Cham Tribe sheep were decreasingly affected by Mongolian sheep bloodline (Lu *et al.* 2005), he detected only Han sheep with 12 blood protein and non-protein loci, and other sheep populations from coastal areas in East Asia were regarded as referenced population. In the previous study we found that Mongolia sheep was the origin of Tong sheep (Sun *et al.* 2007).

All sheep breeds in China except those from Qingzang plateau and its adjacent region originated from ancient Mongolian sheep (Xie 1985). The previous studies (Sun 2006; Zou *et al.* 1994; Lei 1999; Animal Bureau of Shandong Province 1999; Li 1993; Animal Bureau of Ningxia Province 1984; Animal Bureau of Gansu Province 1986; Animal Bureau of Shandong Province 1999; Ran *et al.* 1998; Shan 1983) support our findings that the five sheep populations (Hu sheep, Tong sheep, Tan sheep, small-tailed Han sheep and Wadi sheep) were separated and originated on different stages. Thereafter, in the process of natural and artificial selection in different ecological environments and through genetic differentiation, the relationships among five sheep populations were not closely linearly correlated with their geographical distribution. In agreement to our results Chen *et al.* (2006) used mtDNA of 19 populations from 13 geographical regions, they found that the multidimensional scaling of population pairwise *F*st genetic distances among populations were not in harmony with their geographical locations, and there was no significant geographical structuring of mtDNA variation among Chinese sheep populations (Chen *et al.* 2006). The five sheep populations used in the present study are genetically importance compared to the other sheep breeds, and our findings could be extended to all local sheep breeds in the Mongolian sheep group.

Cluster and principal component scatter plot analysis

STRUCTURE software is regarded as an ideal analytical tool for studying population genetic structure with many successes (Collins-Schramm *et al.* 2004; Heidi *et al.* 2004; Ibeagha-Awemu & Erhardt 2005; Schelling *et al.* 2005). The analysis presumes that every

individual in any group has the same ancestor and estimates the probability of individuals in the group. If a probability exceeds 80% the individual is assigned to the group (Pritchard *et al.* 2000; Falush *et al.* 2003). The cluster analysis based on genetic distance combines all genetic data into one value and may lose some useful information. The bloodline grouping of populations within species is dynamic and fuzzy and inter-population multidirectional gene flow still continues. Therefore, the distance clustering could not reflect well population breeding process. Principal component scatter plot analysis is an effective complement for phylogenic tree and it tends to support the phylogenic results. Cavalli-Sforza *et al.* (1994) indicated that the principal component scatter plot analysis would be better than phylogenic tree, because it provided genetic communication among geographically close populations. 2D and 3D principal component scatter plot charts would reveal the actual relationship among populations to a certain degree.

Conclusions

In the present study the native sheep breeds in the eastern and southern Central Asia were classified into two genetic groups: Mongolia-Tibetan sheep group and South Asia-Southeast Asia sheep group. Mongolia sheep group and the Tibetan sheep group had a certain degree of gene communication.We found that the relationships among the sheep populations in Mongolia sheep group in China were not closely related to the geographical distance among sheep populations. This study may put forward some information for the sustainable development of sheep production and the preservation of biological diversities of native sheep breeds in the east and south region of Central Asia and the Mongolia sheep group of China.

Acknowledgements

This work was supported by the State Scientific Basic Research platform Program (No. 2005DKA21101), International Cooperation Item of the National Natural Science Foundation of China (30410103150), China Postdoctoral Science Foundation funded project (No. 20080430470), Natural Science Foundation of Jiangsu Province of China (BK2007556), Basic Natural Science Foundation for Colleges and Universities Jiangsu Province (NK051039), National High Technology Research and Development Program of China (863 Program) (No. 2006AA10Z198), Support Foundation of China during the 11th Five-Year Plan Period (No. 2006BAD13B08), Qing Lan Project of Colleges and Universities Jiangsu Province, Jiangsu Government Scholarship for Overseas Studies Project, and the New Century Talent Project of Yangzhou University in China.

Supporting information

Additional Supporting information may be found in a online version of this article:

Table S1 The estimates of gene frequencies and their reliability and precision of Hu sheep

Table S2 The estimates of gene frequencies and their reliability and precision of Tong sheep

Table S3 The estimates of gene frequencies and their reliability and precision of Han sheep

Table S4 The estimates of gene frequencies and their reliability and precision of Tan sheep

Table S5 The estimates of gene frequencies and their reliability and precision of WD sheep

Please note: Wiley-Blackwell are not responsible for the content or functionality of any Supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding for the article.

References

Animal Bureau of Gansu Province. (1986) Breeds of Domestic Animal and Poultry in Gansu, Shenzhen Haitian Publishers, People Publishing House of Gansu, Lanzhou. 69–71.

Animal Bureau of Ningxia Province. (1984) Editorial Committee of the Breeds of Domestic Animal and Poultry in Ningxia. Breeds of Domestic Animal and Poultry in Ningxia, Animal Bureau of Ningxia Province, Yinchuan, 1–5.

Animal Bureau of Shandong Province. (1999) Editorial Committee of the Breeds of Domestic Animal and Poultry in Shandong. Breeds of Domestic Animal and Poultry in Shandong, Shenzhen Haitian Publishers, Shenzhen, 51–53.

Cavalli-Sfporza L.L., Menozzi P., Piazza A. (1994) The history and geography of human genes. Princeton University Press, Princeton, NJ.

Chang H. (1995) Essentials of Animal Genetic Resources Science, lst edn. Agricultural Publishing House of China, China, 99–109.

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.

Collins-Schramm H.E., Chima B., Morii T., Wah K., Figueroa Y., Criswell L.A., Hanson R.L., Knowler W.C.,

Silva G., Belmont J.W., Seldin M.F. (2004) Mexican American Ancestry-Informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians. *Hum. Genet.*, **114**, 263–271.

Corander J., Waldmann P., Sillanpaa J. (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367–374.

Falush D., Stephens M., Pritchard J.K. (2003) Inference of population structure: extensions to linked loci and correlated allele frequencies. *Genetics*, **164**, 1564–1587.

FAO (1998) Second Guidelines for Development of National Farm Animal Genetic Resources Management Plans Measurement of Domestic Animal Diversity (MoDAD): Original Working Group Report, FAO, Rome, Italy.

Gahne B., Juneja R.K., Grolmus J. (1977) Horizontal polyacrylamide gradient gel electrophoresis for the simultaneous phenotyping of transferrin, post-transferrin, albumin and post-albumin in the blood plasma of cattle. *Anim. Blood Groups Biochem. Genet.*, **8**, 127–137.

Geng R., Chang H., Wang L., Tsunoda K., Yang Z.P., Sun W., Ji D.J., Li Y. (2007) Genetic differentiation of native sheep populations in East and South Asia. *Biochem. Genet.*, **45**, 263–279.

Hall S.J.G., Bradley D.G. (1995) Conserving livestock breed diversity. *Trends Ecol. Evol.*, **10**, 267–270.

Heidi G.P., Lisa V.K., Nathan B.S., Scott C., Travis D.L., Tiffany B.M., Gary S.J., Hawkins B.D., Elaine A.O., Leonid K. (2004) Genetic structure of the purebred domestic dog. *Science*, **304**, 1160–1164.

Hiendleder S., Mainz K., Plante Y., Lewalski H. (1998) Analysis of mitochondrial DNA indicates that domestic sheep are derived from two different ancestral maternal sources: no evidence for contributions from urial and argali sheep. *J. Hered.*, **89**, 113–120.

Hiendleder S., Kaupe B., Wassmuth R., Janke A. (2002) Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc. R. Soc. Lond. B*, **269**, 893–904.

Ibeagha-Awemu E.M., Erhardt G. (2005) Genetic structure and differentiation of 12 African *Bos indicus* and *Bos Taurus* cattle breeds, inferred from protein and microsatellite polymorphisms. *J. Anim. Breed. Genet.*, **122**, 12–20.

Kantanen J., Olsaker I., Adalsteinsson S., Sandberg K., Eythorsdottir E., Pirhonen K., Holm L.E. (1999) Temporal changes in genetic variation of North European cattle breeds. *Anim. Genet.*, **30**, 6–28.

Kantanen J., Olsaker I., Holm L.E., Lien S., Vilkki J., Brusgaard K., Eythorsdottir E., Danell B., Adalsteinsson S. (2000) Genetic diversity and population structure of 20 North European cattle breeds. *J. Hered.*, **91**, 446–457.

Lei Z.Q. (1999) Tong sheep in Shaanxi. *J. Anim. Sci. Vet. Medi.*, **18**, 35–36.

Li Z.N. (1993) Sheep Science in China, Agricultural publishing House, Beijing, 44–55.

Lu S., Chang H., Du L., Tsunoda K., Sun W., Yang Z., Chang G., Ji D. (2005) Phylogenetic relationships of sheep populations from coastal areas in East Asia. *Biochem. Genet.*, **43**, 251–260.

Nei M. (1973) Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA*, **70**, 3321–3323.

Nei M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590

Ota T. (1993) DISPAN: Genetic Distance and Phylogenetic Analysis. Pennsylvania State University, University Park, PA.

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

Ran R.J., Li G.L., Xu D.F., Li L.C. (1998) Discuss on the breeding history for the Wadi sheep. *China sheep Science*, **3**, 13–14.

Rohlf J.F. (2000) NTSYSpc: Numerical Taxonomy and Multivariate Analysis Group. Version 2.1, Users Guide. Exeter Software, Setauket, New York.

Schelling C., Gaillard C., Dolf G. (2005) Genetic variability of seven dog breeds based on microsatellite markers. *J. Anim. Breed. Genet.*, **2**, 71–77.

Seiki W., Hiroyasu S., Kou N. (1989) Relation between free amino acid type and potassium type of red blood cells in sheep. *Jpn. J. Zootech. Sci.*, **52**, 431–437.

Shan N.S. (1983) Primary discuss of history for Han sheep. *China Semi Fine Wool*, **3**, 15–17.

Sun W. (2006) Population Genetics of Different Ecological Type Sheep Breeds in the region to the East and South of Central Asia, PhD thesis, Yangzhou University, Yangzhou.

Sun W., Chang H., Yang Z.P., Geng R.Q., Lu S.X., Chang G.B., Xu W., Wang H.Y. (2002) Study on genetic relationship of sheep populations from east and south of central Asia Asian-Aus. *J. Anim. Sci.*, **15**, 1398–1402.

Sun W., Chang H., Yang Z.P., Geng R., Tsunoda K., Ren Z., Chen H., Hussein M.H. (2007) Analysis on the origin and phylogenetic status of Tong sheep using 12 blood protein and nonprotein markers. *J. Genet. Genomics*, **34**, 1097–1105.

Tsunoda K., Amano K., Nozawa K., Hasnath M.A. (1988) Morphological characters and blood protein polymorphism of sheep in Bangladesh and genetic relationship with European sheep breeds. *Rep. Soc. Res. Native Livestock*, **12**, 161–185.

Tsunoda K., Amano T., Nozawa K., Hasnath M.A. (1990) Genetic characteristics of Bangladeshi sheep as based on biochemical variations. *Jpn. J. Zootech. Sci.*, **61**, 54–66.

Tsunoda K., Doge K., Yamamoto Y., Kurosawa Y., Shotake T., Nishida T., Rajbhandry H.B. (1992) Morphological traits and blood protein variation of the native Nepalese sheep. *Rep. Soc. Native Livestock*, **14**, 155–183.

Tsunoda K., Doge K., Yamamoto Y., Namikawa T., Amano T., Kurosawa Y., Shotake T., Nishida T., Rajbhandry H.B. (1993) Biochemical polymorphism of Nepalese sheep. *Breeds Anim. Sci. Technol.*, **6**, 1051–1059.

Tsunoda K., Nozawa K., Okamoto S., Zhu J., Hashiguchi T., Liu A., Lin S., Xu W.B., Shi L.M. (1995) Blood protein variation of native sheep population in Lufeng and Lunan in Yunnan province of China. *Rep. Soc. Res. Native Livestock*, **15**, 119–129.

Tsunoda K., Nozawa K., Hasnath M.A. (1998a) Genetic polymorphism of plasma vitamin D-binding protein (Gc) in some Asian sheep Asian-Aus. *J. Anim. Sci.*, **11**, 318–322.

Tsunoda K., Okabayashi H., Amano T., Kuroki K., Namikawa T., Yamagata T., Yamamoto Y., Xun V.T., Loc C.B. (1998b) Morphologic and genetic characteristic of sheep raised by the cham tribe in Vietnam. *Rep. Soc. Res. Native Livestock*, **16**, 63–73.

Tsunoda K., Nozawa K., Madeda Y., Tanabe Y., Taserenbatin T., Rajbhandry H.B. (1999) External morphological characters and blood protein and non-protein polymorphism of native sheep in central Mongolia. *Rep. Soc. Res. Native Livestock*, **17**, 63–82.

Tsunoda K., Amano K., Nozawa K., Hasnath M.A. (2004) Morphological traits and biochemical polymorphisms of Myanmar sheep. *Rep. Soc. Res. Native Livestock*, **21**, 155–169.

Tsunoda K., Chang H., Sun W., Hasnath M.A., Nyunt M.M., Rajbhan-dary H.B., Dorji T., Tumennasan H., Sato K. (2006) Phylogenetic relationships among indigenous sheep populations in East Asia based on five informative blood protein and nonprotein polymorphisms. *Biochem. Genet.*, **44**, 287–306.

Tu Y. (1989) Sheep and Goat Breeds in China. Shanghai Scientific and Technical Publishers, Shanghai. (In Chinese).

Wang X., Ma Y.H., Chen H., Guan W.J. (2006) Genetic and phylogenetic studies of Chinese native sheep breeds (*Ovis aries*) based on mtDNA D-loop sequences. *Small Rumin. Res.*, **72**, 232–236.

Weir B.S., Cockerham C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1369.

Wright S. (1951) The genetical structure of populations. *Ann. Eugen.*, **15**, 323–354.

Wu C.H., Zhang Y.P., Bunch T.D., Wang S., Wang W. (2003) Mitochondrial control region sequence variation within the argali wild sheep (*Ovis ammon*): evolution and conservation relevance. *Mammalian*, **67**, 109–118.

Xie C.X. (1985) History of Raising Cattle, Sheep and Goat in China (Attached History of Raising Deer). Agricultural Publishing House of China, Beijing, 144–150, 173–179.

Yeh F.C., Yang R.C., Boyle T. (1999) Popgene version 1.31 (available at: http://www.ualberta.ca/-fyea/fyea; last accessed 10 May 2008).

Zou J.H., Wang M.N., Niu J.P. (1994) The history of animal and veterinary in ancient time of China. Agricultural Science and Technical Publishers of China, Beijing, 113–118.

