Biochem Genet (2010) 48:376–384
DOI 10.1007/s10528-009-9319-2

Analysis of Geographic and Pairwise Genetic Distances Among Sheep Populations

**Wei Sun · Hong Chang · Kenji Tsunoda ·
Hassan Musa · Yuehui Ma · Weijun Guan**

Received: 2 January 2009 / Accepted: 14 September 2009 / Published online: 8 January 2010
© Springer Science+Business Media, LLC 2010

Abstract This study assessed the usefulness of geographic and pairwise genetic distances in the characterization of five sheep populations using 15 microsatellite markers. The average F statistics across loci were $F_{\mathrm{IT}} = 0.523 \pm 0.140$, $F_{\mathrm{ST}} = 0.363 \pm 0.131$, and $F_{\mathrm{IS}} = 0.263 \pm 0.092$. The average heterozygosity was 0.716 ± 0.069, polymorphism information content was 0.691 ± 0.070, and effective number of alleles was 3.736 ± 0.998. Sheep populations clustered into group 1 (Hu and Tong breeds) and group 2 (small-tailed Han, Wadi, and Tan breeds). Reynolds' distance varied from 0.0062 to 0.0499, and the range of gene flow (N_{m}) was 4.8834–40.0726 among the sheep populations. The results showed that the genetic structure of the five populations was not consistent with their genetic distances, and the population genetic divergence was not linearly related to geographic distance as indicated by a Mantel test ($P = 0.7936$).

Keywords Geographic distance · Pairwise distance · Sheep population

---

W. Sun · Y. Ma (✉) · W. Guan
Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100094, China
e-mail: yuehui.ma@263.net

W. Sun
e-mail: dkxmsunwei@163.com

W. Sun · H. Chang · H. Musa
Animal Science and Technology College, Yangzhou University, Yangzhou 225009, China

K. Tsunoda
School of Medicine, Showa University, Tokyo 142-8555, Japan

H. Musa
Faculty of Veterinary Science, University of Nyala, Nyala 155, Sudan

Introduction

Assessment of genetic variability in domestic sheep is the first step toward the conservation of genetic resources for maintaining breeding options. Phylogenetic studies of population diversification have revealed the essential features of the evolutionary history of many species (Avise 2000). Population differentiation through reduced gene flow, selection, and genetic drift can change the scenario of geographic shifting under regional warming (Davis and Shaw 2001). Domestic animals are currently losing their genetic diversity through many mechanisms. First, highly productive breeds have recently been intensively selected for production traits, without adequate emphasis on the preservation of overall genetic diversity. Second, autochthonous breeds in marginal areas are also seriously endangered (Taberlet et al. 2008). The tools of choice for determining the genetic diversity of domestic animal breeds are statistical measures derived from Wright's F statistics (Wright 1951) or phylogenetic techniques based on genetic distances estimated from polymorphic microsatellite markers (Hall and Bradley 1995). Microsatellites provide vast amounts of genetic information based on individual genotypes, which contributes to establishing genetic relationships between populations (Arranz et al. 1998; Kantanen et al. 2000; Ruane 1999). Calculation of genetic distances from microsatellite data can provide useful information for monitoring and managing the genetic diversity of rare breeds. Therefore, the goal of our study is to assess the usefulness of geographic and pairwise genetic distances in the characterization of livestock breeds. We used a set of 15 microsatellite markers that had previously been shown to be useful in obtaining sound assessments of the genetic relationships among sheep breeds.

Materials and Methods

Experimental Animals

The Chinese sheep populations used in the present study included 63 Hu sheep from Huzhou, Zhejiang Province; 65 Tong sheep from Baishui, Shaanxi Province; 60 small-tailed Han sheep from Liangshan, Shandong Province; 73 Tan sheep from Yinchuan, Ningxia Hui Nationality Autonomous Region; and 76 Wadi sheep from Dongying, Shandong Province.

DNA Extraction and Polymerase Chain Reaction

Blood samples were randomly collected from the jugular vein of each animal. DNA was extracted from whole blood using the saturated salt method (Sambrook et al. 1989), quantified spectrophotometrically, and adjusted to 50 ng/μl. The 15 sheep microsatellite markers (Table 1), located on different chromosomes (Crawford et al. 1995), were synthesized by Shanghai Sangon Biological Engineering Company. The PCR solution comprised 2 μl 50 ng/μl genomic DNA, 0.4 μl 5 U/μl *Taq* DNA polymerase, 15 μl 25 mmol/l $MgCl_2$, 0.5 μl 10 mmol/l dNTP, and 1 μl 8 pmol/μl

Table 1 Genetic diversity across five sheep populations for 15 microsatellite loci

Marker	$F_{\mathrm{IT}} = F$	F_{ST}	$F_{\mathrm{IS}} = f$	H_e	Mean PIC	N_e
0arAE101	0.662***	0.453***	0.382**	0.726	0.699	3.650
OarFCB48	0.504**	0.271***	0.320*	0.682	0.669	3.144
MAF33	0.578***	0.441**	0245**	0.782	0.757	4.587
OarFCB11	0.329**	0.186***	0.176***	0.802	0.786	5.050
MAF70	0.282***	0.190***	0.114***	0.668	0.642	3.012
OarFCB304	0.488**	0.269***	0.300***	0.659	0.628	2.932
OarFCB128	0703***	0.555**	0.332**	0.708	0.667	3.424
OarFCB20	0.489**	0.357***	0.205***	0.817	0.796	5.464
OarHH35	0.302**	0.209***	0.118***	0.773	0.742	4.405
OarAE129	0.546**	0.346***	0.306***	0.650	0.617	2.857
BM1314	0.521**	0.356***	0.256***	0.690	0.672	3.226
BM757	0.663***	0.463***	0.372***	0.719	0.684	3.559
OarCP38	0.607**	0.470***	0.259***	0.662	0.637	2.959
OarCP34	0.459***	0.278***	0.251***	0.582	0.569	2.392
BM1329	0.1726**	0.602***	0.312***	0.814	0.795	5.376
Mean	0.523 ± 0.140 (0.012)***	0.363 ± 0.131 (0.008)***	0.263 ± 0.092 (0.010)**	0.716 ± 0.069 (0.082)	0.691 ± 0.070 (0.084)	3.736 ± 0.998 (0.081)

Note: F, total inbreeding estimate; F_{ST}, measure of population differentiation; f, within-population inbreeding estimate. H_e, heterozygosity. PIC, polymorphism information content. N_e, effective number of alleles. Mean estimates from jack-knife over loci; standard deviations in *parentheses*. Significance of F statistics was calculated using Bonferroni permutations based on 1,000 resamplings

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

forward and reverses primers. PCR conditions were an initial denaturation for 5 min at 94°C; 35 cycles at 94°C for 1 min, 53–66°C for 1 min, and 72°C for 1 min; and a final extension at 72°C for 5 min. A 6 μl aliquot of the PCR product was subjected to 1.5–3.0% agarose gel electrophoresis and 12–14% nondenatured PAGE. The length of the amplified fragment in relation to the pBR322/*Msp*I marker was calculated using Kodak Digital Science ID Image Analysis software.

Statistical Analysis

Gene and genotypic frequencies were estimated by direct gene counting. Heterozygosity (H_e) (Nei et al. 1983), polymorphism information content (PIC) (Botstein et al. 1980), and the effective number of alleles (N_e) (Kimura and Crow 1974) were estimated by Popgene version 1.31 (Yeh et al. 1997). The F statistic indices (Wright 1978) were estimated in terms of F, θ, and f. The sample-based estimators of these parameters were proposed by Weir and Cockerham (1984), and the indices were implemented by the Fstat program (version 2.9.3; Goudet 2002).The significance of the F statistics was determined from permutation tests in which the sequential Bonferroni procedure was applied over loci (Hochberg 1988). The F_{ST} values among pairs of breeds were calculated using Genepop (Raymond and Rousset 1995), and then isolation by distance was applied to the sheep populations (Rousset 1997).

Linear regression was used to estimate the coefficients $Y = F_{ST}/(1 - F_{ST}) = \alpha + \beta \ln(d)$, where d is the pairwise geographic distance between breeds. Gene flow between populations was defined as the number of reproductively successful migrants per generation (N_m), determined on the basis of the n island model of population structure. The estimate was based on the relationship $F_{ST} = 1/(4N_m + 1)$, where N is the effective population size, m is the migration rate, and F_{ST} is the mean F_{ST} value across all loci (Slatkin and Barton 1989). Reynolds' genetic distance between breeds was calculated on the basis of the F_{ST} values (Reynolds et al. 1983): $D_R = -\ln(1 - F_{ST})$. A consensus UPGMA tree was reconstructed using the NTsys-PC package (Rohlf 2000).

Results

Genetic Diversity

The average F statistics across loci were $F_{IT} = 0.523 \pm 0.140$, $F_{ST} = 0.363 \pm 0.131$, and $F_{IS} = 0.263 \pm 0.092$. The average heterozygosity (H_e) was 0.716 ± 0.069, PIC was 0.691 ± 0.070, and effective number of alleles (N_e) was 3.736 ± 0.998 (Table 1). In addition, the average values of F_{IS}, H_e, PIC, and N_e across the 15 loci were estimated for each breed (Table 2), and all breeds showed relatively high heterozygosity. Among the populations studied, the Han sheep had the highest H_e, PIC, and N_e values, and the Hu sheep had the lowest.

Table 2 Mean heterozygosity across 15 microsatellite loci in five sheep populations

Breed	F_{IS}	Mean H_e	Mean PIC	Mean N_e
Hu sheep	0.1560	0.708	0.692	3.425
Tong sheep	0.1326	0.725	0.703	3.636
Small-tailed Han sheep	0.0682	0.821	0.804	5.587
Tan sheep	0.0758	0.796	0.772	4.902
Wadi sheep	0.0814	0.756	0.732	4.098

Phylogenetic Analysis

The UPGMA tree divided the five sheep populations into two groups (Fig. 1). The Hu and Tong populations clustered in group 1 at the 0.26 coefficient level; the small-tailed Han and Wadi populations were clustered first in group 2, followed by the Tan population. The N_m value ranged from 4.8834 between the Hu and Tan breeds to 40.0726 between the Hu and Tong breeds (Table 3). Reynolds' genetic distances (D_R) between populations varied from 0.0062 between the Hu and Tong breeds to 0.0512 between the Hu and Tan breeds (Table 4).

Geographic and Pairwise Distances

The geographic distance ln(d) and the pairwise distance $F_{ST}/(1 - F_{ST})$ among the five sheep populations are also presented in Table 4. The application of Rousset's isolation by distance method, as implemented in the Genepop program, yielded the

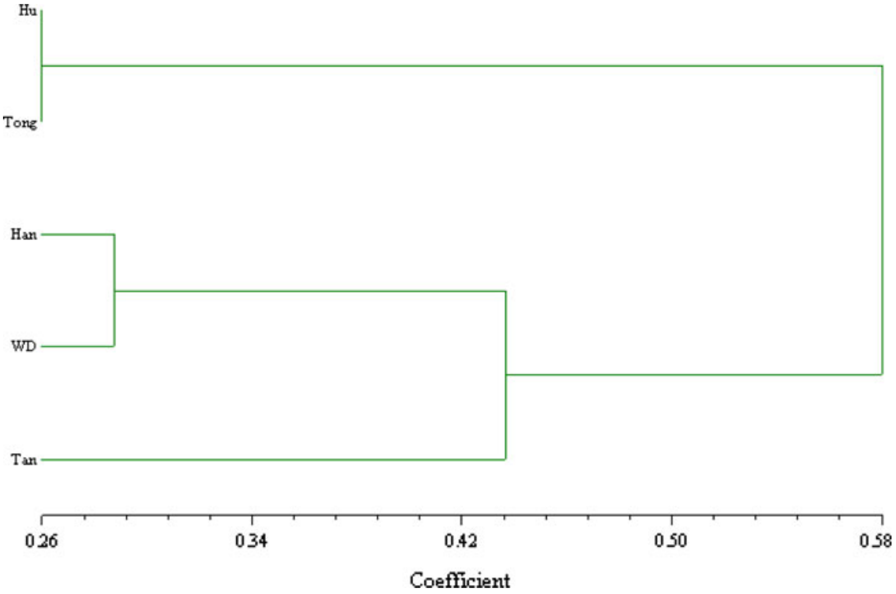

Fig. 1 UPGMA dendrogram for five sheep populations based on 15 microsatellite loci

Table 3 Gene flow between sheep populations

Population	Hu	Tong	Small-tailed Han	Tan	Wadi
Hu		40.0726	10.9107	4.8834	7.8145
Tong	0.0062		11.8273	5.8625	10.2542
Small-tailed Han	0.0227	0.0209		8.1112	12.0652
Tan	0.0499	0.0418	0.0304		7.1464
Wadi	0.0315	0.0241	0.0205	0.0333	

Above the diagonal: number of effective migrants per generation (N_m). Below the diagonal: F_{ST}. The estimate is in balance with genetic drift ($F_{ST} = 1/(4N_{em} + 1)$)

Table 4 Genetic distance between sheep populations

Population	Hu	Tong	Small-tailed Han	Tan	Wadi
Hu	0.0000	7.1507	6.6067	7.5858	6.8701
Tong	0.0062	0.0000	6.9007	6.4615	7.1107
Small-tailed Han	0.0229	0.0211	0.0000	7.2724	5.8749
Tan	0.0512	0.0426	0.0308	0.0000	7.2731
Wadi	0.0320	0.0244	0.0207	0.0350	0.0000

Above the diagonal: geographic distance ln(d). Below the diagonal: pairwise $F_{ST}/(1 - F_{ST})$ distance

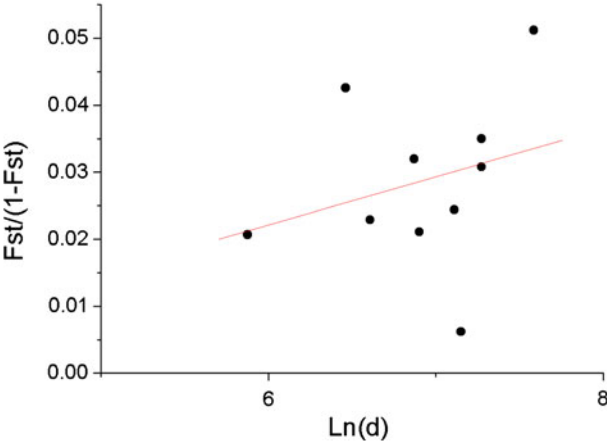

Fig. 2 Plot of relationship between geographic distance ln(d) and pairwise distance $F_{ST}/(1 - F_{ST})$ for all sheep populations. The *line* corresponds to the equation $F_{ST}/(1 - F_{ST}) = -0.021 + 0.0073 \ln(d)$

parameters α and β in the regression $F_{ST}/(1 - F_{ST}) = -0.021 + 0.0073 \ln(d)$ (Fig. 2). The genetic structures of the five sheep populations did not satisfy the criteria of the distance isolation model, and the population genetic divergence was not linearly related to the geographic distance, as indicated by a Mantel test ($P = 0.7936$).

Discussion

The average genetic diversity among the populations observed in our study was consistent with that reported in previous studies (Li et al. 2004; Diez-Tascon et al. 2000). The high PIC values obtained for most of the markers suggest their usefulness in the evaluation of the biodiversity of native Chinese sheep breeds. The Hu and Tong populations clustered first in group 1, and the Han, Wadi, and Tan populations were clustered in group 2, indicating that the phylogenetic relationships among the five sheep populations were not closely linearly correlated with their geographic distribution. This was in agreement with the findings of previous studies, which showed that these populations were originally different but shared the same genetic material because of natural and artificial selection and their ecological habitat (Lei 1999; Li 1993; Ran and Li 1998; Shan 1983). Populations with similar frequencies at microsatellite loci may still have adaptively important differences maintained by natural selection (Hedrick 1999). On the other hand, populations with differing frequencies of microsatellite loci may share adaptively important traits. Differentiation at microsatellite loci should reflect the potential for adaptive differences among populations (Gutiérrez-Espeleta et al. 2000).

In the present study, in order to test isolation by distance, F_{ST} values were plotted against the natural logarithm of pairwise geographic distances among the populations. In order to determine whether the degree of scatter shown in the scatter plots increases with geographic distance and whether the populations are in drift-gene flow equilibrium, the residuals from the linear regression F_{ST} were correlated with geographic distance. No correlation was found, which indicates that there is no relationship between the scatter of pairwise genetic and geographic distances. From a conservation genetics perspective, Hedrick and Miller (1992) indicated that populations should be managed so that adequate genetic variability is retained to provide for future adaptation and successful expansion of native and reintroduced free-ranging populations. Because we cannot directly evaluate the biological significance of the genetic differences between locations and because genetic differences are roughly proportional to geographic distances, the most conservative method of selecting stock for translocation would be to choose the closest available population to preserve local variation and/or potential adaptation (Gutiérrez-Espeleta et al. 2000). Finally, this study showed significant levels of genetic divergence between the five populations, and these findings indicate that the genetic process was associated with historical ecogeographic barriers.

Acknowledgments 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 special funded project (200902154), China Postdoctoral Science Foundation funded project (No. 20080430470), Support Foundation of China during the Eleventh Five-Year Plan Period (No. 2008BADB2B04), Natural Science Foundation of Jiangsu Province of China (BK2007556), Basic Natural Science Foundation for Colleges and Universities Jiangsu Province (NK051039), Jiangsu Government Scholarship for Overseas Studies Project, Qing Lan Project of Colleges and Universities Jiangsu Province, and the New Century Talent Project of Yangzhou University in China.

References

Arranz J, Bayon Y, San Primitivo F (1998) Genetic relationships among Spanish sheep using microsatellites. Anim Genet 29:435–440

Avise JC (2000) The history and formation of species. Harvard University Press, Cambridge, MA, p 439

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:3314–3331

Crawford AM, Dodds KG, Ede AJ, Pierson CA, Montgomery GW, Garmonsway HG, Beattie AE, Davies K, Maddox JF, Kappes SW, Stone RT, Nguyen TC, Penty JM, Lord EA, Broom JE, Buitkamp J, Schwaiger W, Epplen JT, Matthew P, Matthews ME, Hulme DJ, Beh KJ, McGraw RA, Beattie CW (1995) An autosomal genetic linkage map of the sheep genome. Genetics 140:703–724

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate change. Science 292:673–679

Diez-Tascon C, Littlejohn RP, Almeida PAR (2000) Genetic variation within the Merino sheep breeds: analysis of closely related populations using microsatellites. Anim Genet 31:243–251

Goudet J (2002) Fstat: a program to estimate and test gene diversities and fixation indices. Version 2.9.3.2. Available at http://www.unil.ch/izea/softwares/fstat.html

Gutiérrez-Espeleta GA, Kalinowski ST, Boyce WM, Hedrick PW (2000) Genetic variation and population structure in desert bighorn sheep: implications for conservation. Conserv Genet 1:3–15

Hall SJG, Bradley DG (1995) Conserving livestock breed diversity. Trends Ecol Evol 10:267–270

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution 53:313–318

Hedrick PW, Miller PS (1992) Conservation genetics: techniques and fundamentals. Ecol Appl 2:30–46

Hochberg Y (1988) A sharper Bonferroni procedure for multiple test of significance. Biometrika 75:800–802

Kantanen J, Olsaker I, Holm LE, Lien S, Vilkki J, Brusgaard K (2000) Genetic diversity and population structure of 20 North European cattle breeds. J Hered 91:446–457

Kimura M, Crow JF (1974) The number of alleles that can be maintained in a finite population. Genetics 49:725–738

Lei ZQ (1999) Tong sheep in Shaanxi. J Anim Sci Vet Med 18:35–36

Li XL, Gong YF, Zhang JW, Liu ZZ, Allessio V (2004) Study on polymorphisms of microsatellites DNA of six Chinese indigenous sheep breeds. Acta Genet Sin 31:1203–1210

Li ZN (1993) Sheep science in China. Agricultural publishing House, Beijing, pp 44–55

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

Ran RJ, Li GL (1998) Discuss on the breeding history for the Wadi sheep. China Sheep Sci 3:13–14

Raymond M, Rousset F (1995) Genepop: population genetics software for exact test and ecumenicism. J Hered 86:248–249

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetics distance. Genetics 105:767–769

Rohlf JF (2000) NTSYSpc: numerical taxonomy and multivariate analysis system. Version 2.1, Users Guide. Exeter Software, Setauket, New York

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219–1228

Ruane J (1999) A critical review of the value of genetic distance studies in conservation of animal genetic resources. J Anim Breed Genet 116:317–323

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory, NY, p 1659

Shan NS (1983) Primary discuss of history for Han sheep. China Semifine Wool 3:15–17

Slatkin M, Barton NH (1989) A comparison of three indirect methods of estimating average levels of gene flow. Evolution 43:1349–1368

Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, Ajmone-Marsan P (2008) Are cattle, sheep, and goats endangered species. Mol Ecol 17:275–284

Weir BS, Cockerham CC (1984) Estimation F-statistics for the analysis of population structure. Evolution 38:1358–1370

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

Wright S (1978) Evolution and the genetics of populations: variability within and among natural populations, vol 4. University of Chicago press, Chicago, IL, USA

Yeh F, Yang RC, Boyle T (1997) Popgene, the user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Center, University of Albert, Edmonton