Published December 4, 2014

Genetic structure and diversity among sheep breeds in the United States: Identification of the major gene pools[1,2]

H. D. Blackburn,*[3] S. R. Paiva,† S. Wildeus,‡ W. Getz,§ D. Waldron,# R. Stobart,‖ D. Bixby,¶ P. H. Purdy,* C. Welsh,* S. Spiller,* and M. Brown**

*National Animal Germplasm Program, National Center for Genetic Resources Preservation, ARS, USDA, Ft. Collins, CO 80525; †Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Genetic Resources and Biotechnology, Brasilia, Distrito Federal, Brazil 70770-917; ‡Agriculture Research Station, Virginia State University, Petersburg 23806; §College of Agriculture, Fort Valley State University, Fort Valley, GA 31030; #Animal Science Department, Texas A&M University, San Angelo 76901; ‖Animal Science Department, University of Wyoming, Laramie 82071; ¶American Livestock Breeds Conservancy, Pittsboro, NC 27312; and **Grazing Lands Research Laboratory, ARS, USDA, El Reno, OK 73036

ABSTRACT: Understanding existing levels of genetic diversity of sheep breeds facilitates in situ and ex situ conservation activities. A comprehensive evaluation of US sheep breeds has not been previously performed; therefore, we evaluated the genetic diversity among and within 28 US sheep breeds. Both major and minor breeds were included in the analysis and consisted of 666 animals from 222 producers located in 38 states. The level of within-breed genetic diversity was variable and not dependent upon status of a breed as a major or minor breed. Bayesian cluster analysis indicated the breeds were grouped more by physiological differences (meat vs. wool production) rather than geographic origin. Results suggest several actionable items to improve in situ and ex situ conservation. The results clearly identify breeds in need of increased in situ and ex situ management (e.g., Hog Island and Karakul) and allow several suggestions for in situ management of flocks. Conversely, several of the breeds appear genetically similar and therefore require less emphasis on collecting germplasm samples for the gene bank. Commercially important breeds (e.g., Rambouillet and Suffolk) were found to have substantial variation, which should enable breeders to proceed, unencumbered by genetic diversity concerns, with selection strategies that maximize profit.

Key words: animal genetic resource conservation, Bayesian cluster analysis, domestic sheep, genetic diversity

©2011 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2011. 89:2336–2348
doi:10.2527/jas.2010-3354

INTRODUCTION

Geographically and functionally diverse breeds of sheep were imported into the United States from the 16th century to the present (Dohner, 2001). Once imported to the United States, populations found a niche, disappeared, or were crossed with other breeds (Wood and Orel, 2001; Blackburn and Gollin, 2009). Traditionally, sheep breeders have divided breeds into categories based on wool- and meat-producing characteristics. Wool categories are typically fine, medium, or coarse. Meat breeds tend to produce wool of lesser quality and quantity, have greater growth rates, and usually have superior carcass characteristics compared with wool breeds. Although all types are used for meat and wool production, within-breed selection emphasis for these traits has differed.

Phylogenetic studies have differentiated sheep breeds and their origins. For example, the genetic distances between domestic (*Ovis aries*) and wild species of sheep suggest 2 to 5 mitochondrial lineages in the domestication process (Hiendleder et al., 2002; Pedrosa et al., 2005; Meadows et al., 2007). Microsatellites and SNP

[1]The authors thank the owners of the numerous flocks that provided tissue samples for this project. We also thank Y. Plante (Agriculture and Agri-Foods Canada, Saskatoon) and L. Kuehn (USDA/ARS, Clay Center, NE) for their constructive comments on the draft manuscript.

[2]Mention of a trade name or proprietary product does not constitute a guaranty or warranty by the USDA and does not imply approval to the exclusion of other products that may be suitable. The USDA, ARS, Northern Plains Area, is an equal opportunity/affirmative action employer. All agency services are available without discrimination.

[3]Corresponding author: Harvey.blackburn@ars.usda.gov

Received July 26, 2010.
Accepted February 22, 2011.

have been useful in quantifying the genetic distances among sheep breeds in various geographic regions (Diez-Tascón et al., 2000; Tapio et al., 2005; Baumung et al., 2006; Peter et al., 2007; Lawson Handley et al., 2007; Kijas et al., 2009). Such studies have shown that introgression has played an important role in breed formation and utilization in sheep over time (Meadows et al., 2005).

Genetic diversity among livestock breeds has contracted and the Interlaken Declaration and Global Plan of Action (**GPA**) for Animal Genetic Resources highlighted the need for better livestock population characterization (FAO, 2007a,b). Therefore, the objectives of this study were to quantify the levels of genetic diversity in and among US sheep breeds, provide information for breed conservation, and fulfill GPA obligations.

MATERIALS AND METHODS

Animal Care and Use Committee approval was not obtained for this study because the study did not use live animals.

Breeds, Sample Acquisition, and DNA Extraction

Twenty-eight sheep breeds were selected for study based on their economic importance or rarity. As background, Table 1 reviews the origin or use, or both, of the breeds evaluated in the study. Specific information for several breeds is presented for consideration in the interpretation of the results. The Black Welsh Mountain breed has had several importations during the last decade, and the samples from that breed represent both preexisting and recently imported genetics. The Navajo Churro and Gulf Coast Native are thought to be descendents from early Spanish Churro importations during the 16th century. Several importations of Dorper since the early 1990s have occurred. The first importation was from Australia, but subsequent importations occurred directly from South Africa. The Hog Island is thought to be representative of colonial sheep brought to the United States from England. For a substantial period of time the sheep were feral on Hog Island off the Virginia coast but were removed in 1974 (Dohner, 2001). It has been speculated that Santa Cruz Island breed, a population from Santa Cruz Island off the California coast, might contain some portion of Churro breeding from Spain. Several breeds in the study are composite breeds formed during the 20th century that utilized Rambouillet. For example, the Warhill, which has been maintained principally as a single population (ranging in size from 5,000 to 20,000 animals), was developed using Rambouillet and Rambouillet composite breeds. Populations are contracting, and organizations like the FAO or the American Livestock Breeds Conservancy have developed classification systems for designating the degree of rarity or endangerment. For this paper we will use the terms major, minor, or rare breeds to denote if a breed plays an important role or does not contribute in a substantial manner to the economic vitality of the US sheep industry.

From 28 breeds, blood or semen samples from 666 animals were collected from 222 producers in 38 states. Criteria for within-flock animal sampling included the acquisition of tissue samples from both sexes and no known genetic relationship. In addition, efforts were made to sample from flocks that had been relatively independent in their breeding programs. Blood samples were collected by the owner of the animal or a collaborator and shipped to this laboratory for processing. Upon receipt, blood samples were cryopreserved until ready for DNA extraction and analysis. The semen utilized in the study was acquired while developing breed collections for the gene bank. The blood and semen cryopreservation protocols used are found at the website http://www.ars.usda.gov/Main/docs.htm?docid=16979 (last accessed June 2009). The DNA was isolated from cryopreserved blood and semen samples using the BloodPrep chemistry protocol and the NucPrep chemistry protocols (Applied Biosytems, Foster City, CA), respectively, in conjunction with the ABI Prism 6100 Nucleic Acid PrepStation (Applied Biosystems).

Analysis of Microsatellite Data

For this analysis the Food and Agriculture Organization of the United Nations/International Society for Animal Genetics (**FAO/ISAG**) panel of 31 microsatellite markers were used (FAO, 2004) to maintain the option for combining this data set with other studies using the same complement of markers. The consensus panel markers are located across 21 chromosomes and had as a selection criterion that they should be known to be unlinked (FAO, 2004). A commercial company (GeneSeek, Lincoln, NE) constructed the multiplex system, amplified the DNA samples by PCR, and made the allele calls. Appendix Table A1 lists the markers (chromosome number), number of alleles, and the percentage of missing data from each microsatellite. Markers MCM140, OarFCB20, and BM1329 did not amplify well and were deleted from the analysis.

The GENALEX 6 program (Peakall and Smouse, 2006) was used to compute the average and effective number of alleles, allele frequency per locus, observed and expected heterozygosity, private alleles of a breed, principal component analysis, and the analysis of molecular variance (**AMOVA**). The AMOVA was performed using the codominant allelic distance matrix with 999 permutations. Inbreeding ($\mathbf{F_{is}}$) was calculated using FSTAT (Goudet, 2000). Genetic distance was determined by using the program MICROSAT (Minch

Table 1. Breeds, phenotypic descriptors, and number of animals, breeders, and states included in the study in addition to the country of origin of the breed

Breed	Conservation priority[1]	Phenotypic descriptors[2]	Animal numbers	Number of breeders	Number of states	Country of origin[2]
Barbados Blackbelly	Recovering	Hy, m, prolific	18	7	6	Barbados
Black Welsh Mountain	Recovering	Sw, m	30	3	3	United Kingdom
Columbia	—	Mw, m	21	7	5	United States[3,4]
Cotswold	Threatened	Lw, m	9	4	4	United Kingdom
Dorper	—	Hy, cw, m, fr	44	23	12	South Africa
Dorset	—	Sw, m	27	12	6	United Kingdom
Finnish Landrace	—	Prolific, w, m	20	3	3	Finland
Gulf Coast Native	Critical	Cw, mw	30	8	5	United States[3]
Hampshire	—	Sw, m	27	20	12	United Kingdom
Hog Island	Critical	Cw, mw, m	24	6	2	United States[4]
Jacob	Threatened	Mw, m	24	9	6	United Kingdom
Karakul	Threatened	Ft, d	19	4	4	Uzbekistan
Katahdin	Recovering	Hy, m, prolific	29	13	10	United States
Leicester Longwool	Critical	Lw, m	29	9	6	United Kingdom
Lincoln	Watch	Lw, m	19	12	9	United Kingdom
Navajo Churro	Threatened	Cw	31	16	7	Spain[3]
Polypay	—	Prolific, w, m	17	11	9	United States[3,4]
Romanov	—	Prolific, cw	24	4	3	Russia
Rambouillet	—	Fw, m	46	20	8	France
Romney	—	Lw, m	20	13	9	United Kingdom
Santa Cruz Island	Critical	Mw	21	1	1	United States[3]
St. Croix	Threatened	Hy	26	11	7	United Kingdom
Southdown	Recovering	M, sw	7	5	5	United Kingdom
Suffolk	—	M, sw	26	13	7	United Kingdom
Targhee	—	Mw, m	18	4	4	United States[3,4]
Texel	—	M, lw	20	7	7	the Netherlands
Tunis	Watch	M, ft	14	5	5	Tunisia
Warhill	—	Mw, m	26	1	1	United States[3,4]

[1]American Livestock Breeds Conservancy conservation priority ranking.

[2]Breed origin and phenotypic descriptors from Mason (1996): m = meat, fw = fine wool, sw = short wool, mw = medium wool, lw = English long wool, hy = hair, cw = coarse wool, w = wooled, ft = fat tail, fr = fat rump, d = dairy.

[3]Origin of foundation breeds: Spanish.

[4]Origin of foundation breeds: British.

et al., 1996), where 1,000 bootstraps were performed in computing Nei's estimate of genetic distance (Nei, 1972). Nei's genetic distances were used as input to the PHYLIP program version 3.67 (Felsenstein, 2007) after 1,000 bootstraps to construct a neighbor joining tree illustrating the association among the breeds.

The Bayesian clustering program, STRUCTURE (Pritchard et al., 2000), was used with an admixture model with the correlations between loci option. The program was run using a burn-in of 100,000 iterations followed by an additional 500,000 iterations, the results of which were used in the analysis. Within each specified cluster (**K**) ranging from 1 to 28, 3 replicates were run and averaged for use in the analysis. The averaged likelihood at each K was used to calculate ΔK (Evanno et al., 2005; McKay et al., 2008), which may be used as an ad hoc indicator of population number. The ΔK indicated several minor peaks in addition to an increased region that ranged from 20 to 23 populations, with ΔK reaching a maximum when K = 21 (Appendix Figure A1). Based on the ΔK, K values of 2, 3, 4, 5, 11, and 21 were evaluated. Graphical representation of cluster assignments was constructed using DISTRUCT (Rosenberg, 2004).

RESULTS AND DISCUSSION

Within-Breed Diversity

Alleles per locus ranged from 6 to 24, with an average of 14.1 alleles per locus; these values are within the ranges of previously reported values using the same markers (Lawson-Handley et al., 2007; Peter et al., 2007; Dalvit et al., 2009). Across breeds, the average number of alleles (Table 2) was 5.86 with a range of 3.75 (Black Welsh Mountain) to 8.18 (Rambouillet). Among European breeds, Peter et al. (2007) reported an average number of alleles per breed of 6.42 and a range of 5.0 to 7.52 alleles. Minor sheep breeds had less than the average number of alleles, with the Black Welsh Mountain, Cotswold, Southdown, and Hog Island ranking the lowest. In addition, Hog Island, Black Welsh Mountain, Cotswold, and Romanov each had one locus that was not polymorphic. Among rare breeds, the Gulf Coast Native and Navajo Churro ranked high in mean number of alleles (Table 2).

Fifty-three private alleles were found among the breeds. However, 43 of the private alleles were present in frequencies of less than 0.05 (Appendix Table A2).

Table 2. Measures of within-breed allelic richness and genetic diversity[1]

Breed	Average number of alleles	F_{is}	H_o	H_e	Loci not in HWE ($P < 0.05$)
Barbados Blackbelly	4.86	0.223***	0.513	0.636	9
Black Welsh Mountain	3.75	−0.001	0.467	0.456	5
Columbia	5.78	0.092***	0.597	0.639	6
Cotswold	3.86	0.099*	0.543	0.561	0
Dorper	7.57	0.141***	0.602	0.690	10
Dorset	7.32	0.105***	0.639	0.698	10
Finnsheep	6.57	0.133***	0.634	0.707	7
Gulf Coast Native	7.57	0.174***	0.606	0.718	12
Hampshire	6.71	0.073***	0.607	0.641	8
Hog Island	4.11	0.134***	0.370	0.417	5
Jacob	5.68	0.089***	0.572	0.613	2
Karakul	5.61	0.264***	0.473	0.620	12
Katahdin	6.71	0.123***	0.591	0.659	10
Leicester Longwool	5.14	0.112***	0.484	0.534	9
Lincoln	4.89	0.139***	0.521	0.585	6
Navajo Churro	7.71	0.151***	0.610	0.703	10
Polypay	6.39	0.086***	0.649	0.686	5
Rambouillet	8.18	0.147***	0.617	0.714	14
Romanov	5.18	0.069	0.571	0.598	6
Romney	5.89	0.124***	0.591	0.654	7
Santa Cruz Island	5.25	0.004	0.609	0.596	5
Southdown	4.14	0.100*	0.554	0.568	2
St. Croix	5.93	0.140***	0.582	0.661	10
Suffolk	6.07	0.139***	0.578	0.655	8
Targhee	6.32	0.122***	0.624	0.687	3
Texel	6.21	0.086***	0.611	0.647	6
Tunis	4.86	0.162***	0.537	0.611	6
Warhill	6.07	0.002	0.711	0.688	3

[1]H_o (observed heterozygosity), H_e (expected heterozygosity), and loci not in HWE computed by GENALEX; inbreeding (F_{is}) computed by FSTAT. HWE = Hardy-Weinberg equilibrium.

***Significant at $P < 0.001$. *Significant at $P < 0.05$.

Private alleles with high frequencies were found in the Tunis (0.62 and 0.15) and Warhill (0.30). The Black Welsh Mountain, Cotswold, and Santa Cruz Island had no private alleles.

Breed estimates of F_{is} were computed using FSTAT (Goudet, 2000) and ranged from −0.001 to 0.264 (Table 2). Breeds with the greatest F_{is} values were Barbados Blackbelly, Gulf Coast Native, Tunis, Navajo Churro, and Karakul. Near zero F_{is} were found for Black Welsh Mountain, Santa Cruz Island, and Warhill. Both Santa Cruz Island and Warhill were relatively large, but isolated, populations (with as many as 80,000 and 20,000 animals, respectively) until the last decade. When inbreeding values for Warhill were calculated (Falconer and Mackay, 1996) based upon typically used sex ratios and generation intervals for the range sheep industry, we found general concurrence between the calculations and the results reported in this study. New importations of Black Welsh Mountain from a broad array of United Kingdom flocks during the last decade is a likely explanation for the low inbreeding level found in this report.

Many of the breeds (major and minor) in this study had relatively large levels of observed heterozygosity (Table 2). The measured heterozygosities in this study are in agreement with Muigai et al. (2002) for the St. Croix, Barbados Blackbelly, Rambouillet, and Gulf Coast Native. Breeds showing the lowest levels of observed heterozygosity included the Black Welsh Mountain, Hog Island, and Karakul.

Averaged across breeds, 25% of the loci were not in Hardy-Weinberg equilibrium (**HWE**), which was greater than the 18% reported by Peter et al. (2007). The increased percentage of loci not in HWE is due to 8 breeds with 10 to 14 loci not in HWE (Table 2). Among the 8 breeds, the mean number of alleles was above the average for all breeds in the study.

Genetic Structure of Breeds

The AMOVA indicated that 13% of the total variation was present among breeds, and the majority of genetic variation was found within breed. Similar partitioning of variance has been reported in several studies (Lawson Handley et al., 2007; Peter et al., 2007).

Based in part on the ΔK analysis, evaluation of K at 2, 3, 4, 5, 11, and 21 is shown to illustrate the progression of breed structure (Figure 1). The initial partition (K = 2) separated the long wool breeds (Leicester Longwool, Lincoln, Cotswold, Karakul, and Romney) from a set of breeds that are dissimilar (Southdown,

Katahdin, Barbados Blackbelly, Rambouillet, Hampshire, Suffolk, and Finnsheep). In addition, 16 breeds were almost equally split between clusters 1 and 2. Setting K = 3 resulted in partitioning of breeds into long/coarse wool, medium/fine wool, and those typically considered meat-producing (e.g., Hampshire, Suffolk, Saint Croix, Barbados Blackbelly) breeds (Figure 1). At K = 4 the meat breeds were subdivided so that the Caribbean hair breeds were separated, and when K = 5 the Black Welsh Mountain and a large proportion of the Hog Island were placed in the new cluster. Tracking the assignments for Santa Cruz Island, Gulf Coast Native, Jacob, and Navajo Churro from 3 to 5 clusters indicated substantial proportions of admixture for these breeds. Throughout the 19th and 20th centuries, Dohner (2001) discusses a variety of situations where these breeds were known to be crossbred. In contrast, very distinct wool breeds like the Leicester Longwool and Rambouillet were consistently placed into their respective clusters with little admixture.

Delta K indicated that K = 11 was a significant partition of the 28 breeds (Figure 1). The initial clusters of long/coarse wool and fine/medium wool breeds remained grouped as single clusters. But at this point in the analysis the original meat breed cluster (from K = 3) was decomposed into 4 different clusters.

Delta K analysis indicated that partitioning of the 28 breeds into 21 clusters (K) encapsulated the US populations (Figure 1). This deviated considerably from work with cattle which typically show K = 2 due to the major genetic differentiation between *Bos taurus* and *Bos indicus* (McKay et al., 2008). Increasing the number of populations (**K**) to 21 and the greatest ΔK resulted in further decomposition of the physiological groups to specific breeds (12 of 28 breeds had exclusive clusters vs. 4 of 28 when K = 11). At K = 21, the Hampshire-Suffolk, Warhill-Rambouillet, Lincoln-Cotswold were consolidated, matching commonly held perceptions about how these breeds have been managed during the last 40 yr. Comparing K = 21 through 23 generally showed a stabilization of breed assignments. Albeit minor, the shifting of population proportions between various clusters when K ranged from 21 to 23 suggested the partitioning of these proportions was not particularly strong in relation to the distinctness of other breeds.

Among-Breed Diversity

Nei's genetic distance was computed and used (Figure 2). Distances ranged from 0.04 (Hampshire to Suffolk) to 0.54 (Barbados Blackbelly to Black Welsh Mountain). The similarity between the Hampshire and Suffolk was anticipated due to their similar function in the sheep industry, purported admixture, and a common ancestral breed. Genetic distances of <0.15 were found among the Rambouillet-based breeds (Columbia, Targhee, Polypay, and Warhill). The Navajo Churro and Dorset also had close association with the Rambouillet grouping. We believe the close association of the Dorset was in part due to the role it had in the formation of the Polypay, which is 25% Dorset and 43% Rambouillet. Of particular interest was the genetic distance (0.10) between the Gulf Coast Native and Navajo Churro. In the United States these 2 breeds are geographically distinct, but are thought to have descended from Spanish breed(s) imported in the 16th century. The closeness of these 2 breeds may suggest similar founding populations.

An unrooted tree was constructed using Nei's neighbor joining tree (Felsenstein, 1993) with 1,000 bootstraps performed (Appendix Figure A2). The joining of the nearest neighboring breeds generally matched known history of the breeds.

Principal component (**PC**) analysis was performed using the covariance matrix computed from the genetic distances (Figure 2). The first 3 PC explained 61.9% of the total variation. The first PC (24.7%) separated the breeds into 4 main groupings consisting of the Hog Island-Black Welsh Mountain, Hampshire-Suffolk-Romanov, Leicester Longwool-Lincoln-Cotswold-Columbia, and all other breeds (Figure 3). The second PC (20.4%) separated breeds into 4 groups consisting of Leicester Longwool-Lincoln-Cotwold-Karakul-Hog Island, Romanov-Southdown, Barbados Blackbelly-St. Croix, and the remaining breeds. The second PC groups appeared to be grouped by wool type: hair breeds vs. long coarse wool breeds vs. fine and medium wool breeds. The third PC (16.8%) grouped breeds throughout the range of values with Black Welsh Mountain-Tunis and Hog Island-Southdown at the extremes. The PC analysis indicated a relatively close alignment of the rare breeds originating from Spain (Gulf Coast Native-Santa Cruz Island-Navajo Churro) along the axes of the first and third PC.

Conservation Approaches

A rationale for performing this study was to better define actions for ex situ conservation of sheep breeds. In the United States a programmatic decision was made to develop germplasm collections for all breeds (Blackburn, 2009). To date the repository has collected 49,912 germplasm and tissue samples from 1,735 animals, representing 38 sheep breeds (http://www.ars.usda.gov/Main/docs.htm?docid=16979; last accessed December 2010). Although these collections have been initiated, they are not complete in terms of quantity of germplasm collected and within-breed genetic diversity represented.

Based upon number of alleles, heterozygosity, and genetic distance measures, it appears that a relatively broad range of ovine genetic resources is present in the United States. Several rare and major breeds had genetic diversity measures that were robust in terms of average number of alleles and observed heterozygosity

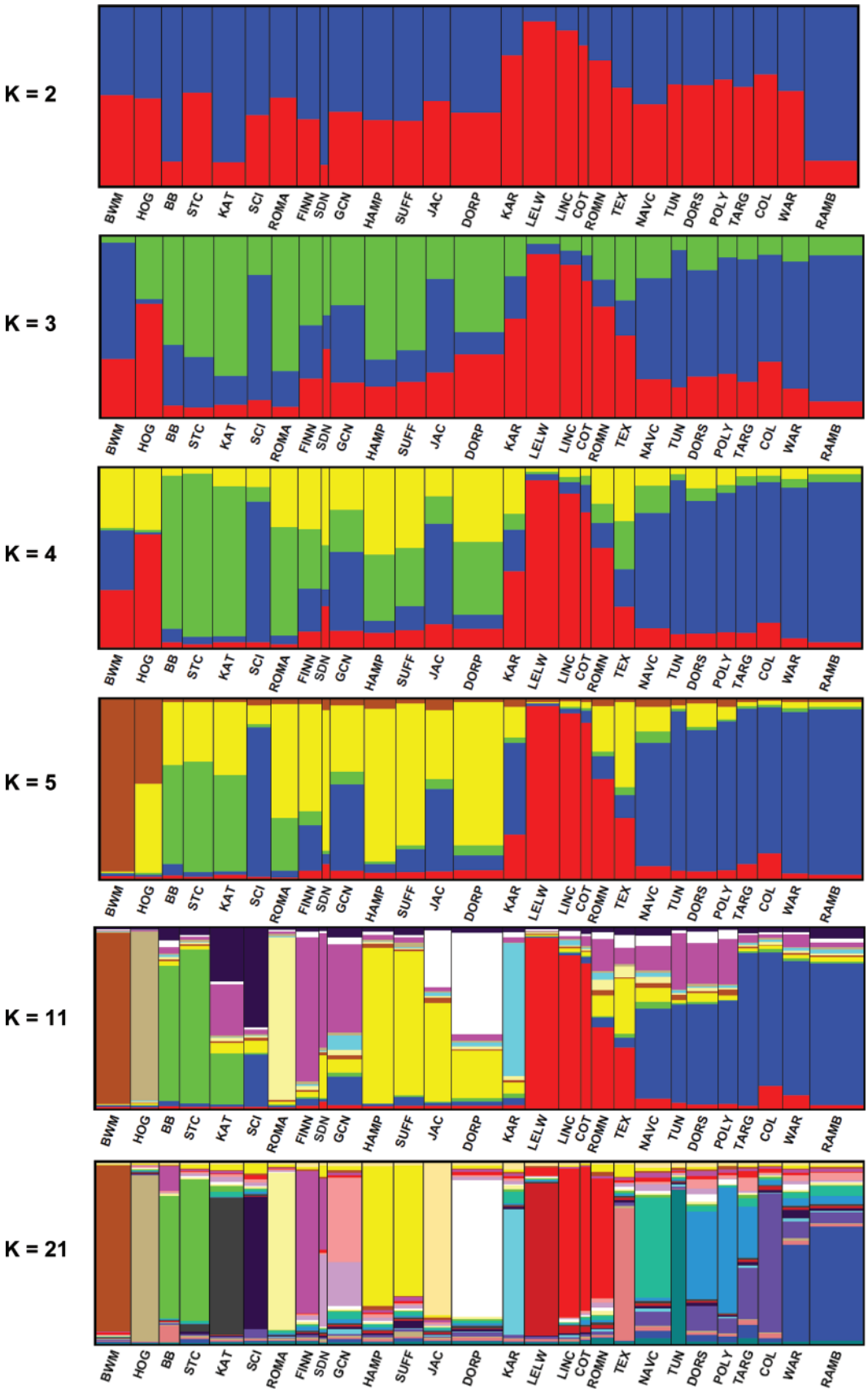

Figure 1. Breed assignment to color-coded clusters based on a Bayesian analysis (using STRUCTURE) when the number of clusters (K) was set to 2, 3, 4, 5, 11, and 21. Plots for each cluster were constructed using the program DISTRUCT; the width of each segment is based on breed sample size. BB = Barbados Blackbelly; BWM = Black Welsh Mountain; COL = Columbia; COT = Costwold; DORP = Dorper; DORS = Dorset; FINN = Finnsheep; GCN = Gulf Coast Native; HAMP = Hampshire; HOG = Hog Island; JAC = Jacob; KAR = Karakul; KAT = Katahdin; LELW = Leicester Longwool; LINC = Lincoln; NAVC = Navajo Churro; POLY = Polypay; RAMB = Rambouillet; ROMA = Romanov; ROMN = Romney; SCI = Santa Cruz Island; SDN = Southdown; STC = Saint Croix; SUFF = Suffolk; TARG = Targhee; TEX = Texel; TUN = Tunis; WAR = Warhill. Color version available in the online PDF.

	BB	BWM	COL	COT	DORP	DORS	FINN	GCN	HAMP	HOG	JAC	KAR	KAT	LELW	LINC	NAVC	POLY	RAMB	ROMA	ROMN	SCI	SDN	STC	SUFF	TARG	TUN	WAR
BB	-																										
BWM	0.54	-																									
COL	0.22	0.45	-																								
COT	0.37	0.43	0.21	-																							
DORP	0.32	0.42	0.21	0.22	-																						
DORS	0.26	0.31	0.10	0.17	0.16	-																					
FINN	0.31	0.36	0.26	0.27	0.19	0.15	-																				
GCN	0.30	0.32	0.19	0.23	0.18	0.11	0.16	-																			
HAMP	0.38	0.32	0.27	0.28	0.22	0.22	0.21	0.16	-																		
HOG	0.51	0.46	0.45	0.42	0.36	0.33	0.34	0.34	0.37	-																	
JAC	0.33	0.35	0.26	0.30	0.23	0.16	0.23	0.20	0.27	0.40	-																
KAR	0.45	0.47	0.32	0.32	0.25	0.28	0.30	0.28	0.33	0.40	0.36	-															
KAT	0.31	0.49	0.31	0.31	0.22	0.24	0.24	0.24	0.24	0.50	0.33	0.40	-														
LELW	0.37	0.51	0.20	0.19	0.28	0.25	0.29	0.30	0.37	0.48	0.32	0.37	0.39	-													
LINC	0.35	0.49	0.18	0.12	0.22	0.19	0.24	0.24	0.26	0.46	0.28	0.28	0.35	0.15	-												
NAVC	0.24	0.31	0.14	0.20	0.16	0.09	0.17	0.10	0.20	0.37	0.18	0.29	0.22	0.29	0.23	-											
POLY	0.31	0.34	0.14	0.19	0.19	0.07	0.16	0.17	0.23	0.37	0.24	0.28	0.28	0.26	0.22	0.15	-										
RAMB	0.28	0.37	0.13	0.27	0.26	0.11	0.21	0.15	0.28	0.45	0.29	0.33	0.26	0.30	0.28	0.10	0.14	-									
ROMA	0.36	0.44	0.37	0.41	0.26	0.26	0.24	0.26	0.28	0.44	0.32	0.36	0.32	0.49	0.39	0.25	0.28	0.32	-								
ROMN	0.37	0.33	0.22	0.19	0.21	0.18	0.20	0.20	0.16	0.41	0.27	0.26	0.26	0.24	0.18	0.21	0.19	0.25	0.31	-							
SCI	0.37	0.37	0.23	0.36	0.26	0.24	0.33	0.24	0.26	0.44	0.32	0.42	0.34	0.40	0.33	0.22	0.32	0.21	0.37	0.30	-						
SDN	0.38	0.41	0.30	0.36	0.29	0.22	0.21	0.28	0.23	0.37	0.28	0.44	0.33	0.33	0.33	0.28	0.24	0.33	0.38	0.27	0.34	-					
STC	0.21	0.48	0.30	0.42	0.29	0.29	0.32	0.25	0.33	0.49	0.36	0.45	0.23	0.44	0.43	0.22	0.34	0.27	0.43	0.36	0.33	0.41	-				
SUFF	0.35	0.35	0.23	0.28	0.19	0.18	0.21	0.15	0.04	0.36	0.24	0.32	0.22	0.36	0.28	0.16	0.21	0.23	0.28	0.19	0.20	0.21	0.32	-			
TARG	0.21	0.36	0.08	0.17	0.23	0.08	0.18	0.13	0.25	0.38	0.22	0.32	0.27	0.24	0.20	0.07	0.10	0.09	0.29	0.20	0.22	0.27	0.25	0.23	-		
TEX	0.30	0.42	0.21	0.22	0.22	0.17	0.20	0.22	0.22	0.45	0.26	0.38	0.30	0.24	0.19	0.16	0.22	0.24	0.26	0.25	0.28	0.35	0.34	0.24	0.18	-	
TUN	0.40	0.37	0.31	0.36	0.34	0.22	0.29	0.23	0.35	0.53	0.30	0.38	0.34	0.39	0.32	0.21	0.31	0.22	0.36	0.32	0.36	0.44	0.39	0.30	0.23	0.31	-
WAR	0.26	0.37	0.12	0.24	0.22	0.11	0.21	0.13	0.27	0.42	0.23	0.28	0.26	0.25	0.22	0.11	0.18	0.07	0.34	0.24	0.21	0.32	0.28	0.23	0.12	0.21	0.22

Figure 2. Nei's genetic distance calculated from MICROSAT using 1,000 bootstraps. BB = Barbados Blackbelly; BWM = Black Welsh Mountain; COL = Columbia; COT= Costwold; DORP = Dorper; DORS = Dorset; FINN = Finnsheep; GCN = Gulf Coast Native; HAMP = Hampshire; HOG = Hog Island; JAC = Jacob; KAR = Karakul; KAT = Katahdin; LELW = Leicester Longwool; LINC = Lincoln; NAVC = Navajo Churro; POLY = Polypay; RAMB = Rambouillet; ROMA = Romanov; ROMN = Romney; SCI = Santa Cruz Island; SDN = Southdown; STC = Saint Croix; SUFF = Suffolk; TARG = Targhee; TEX = Texel; TUN = Tunis; WAR = Warhill. Genetic distance with no shading >0.40; light gray 0.26 to 0.40; medium gray 0.11 to 0.25; and dark gray 0.0 to 0.10.

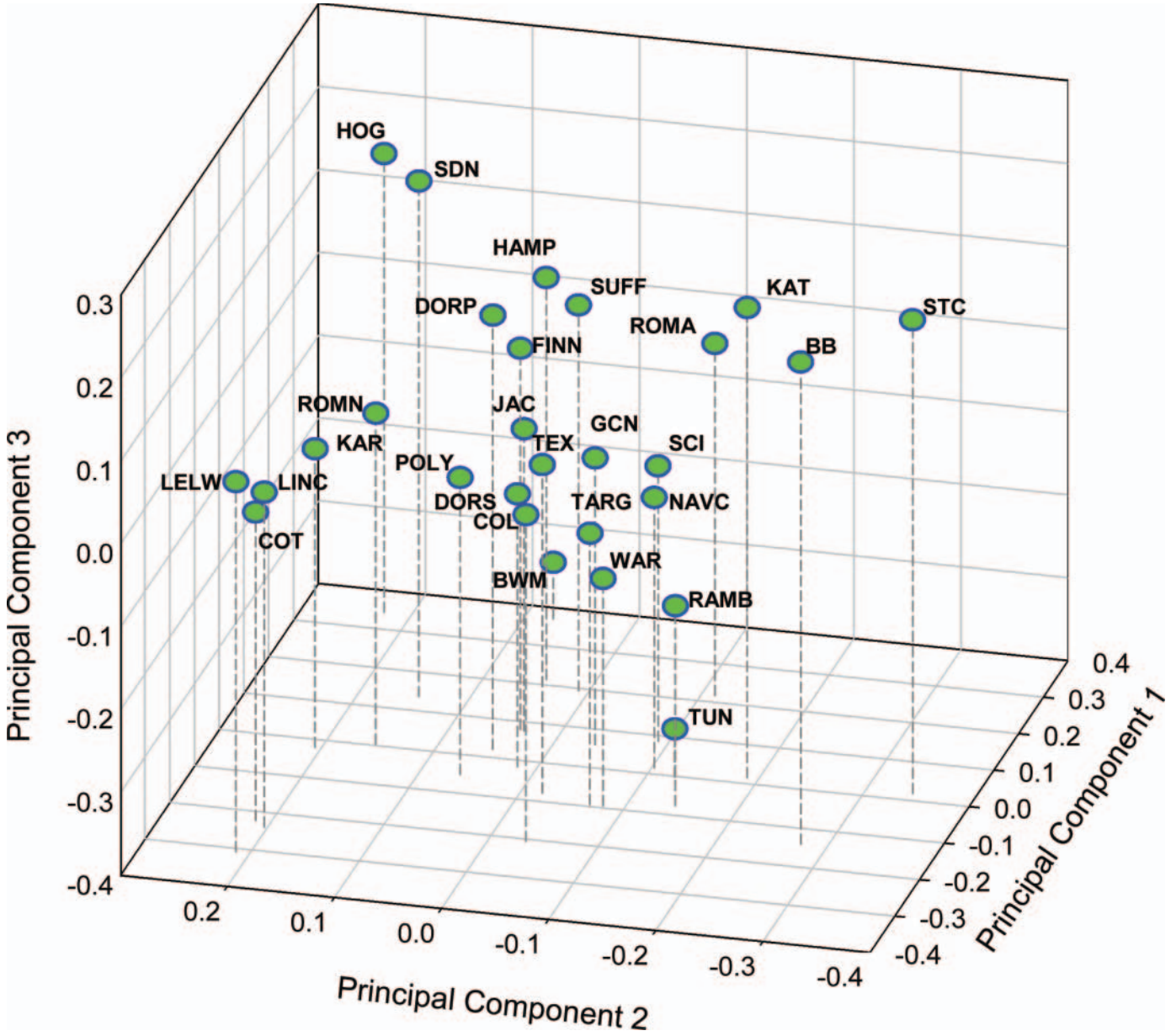

Figure 3. Breed placement by the primary principal components, where the variation explained by each principal component was as follows: principal component 1 = 24.7%, principal component 2 = 20.4%, and principal component 3 = 16.8%. BB = Barbados Blackbelly; BWM = Black Welsh Mountain; COL = Columbia; COT= Costwold; DORP = Dorper; DORS = Dorset; FINN = Finnsheep; GCN = Gulf Coast Native; HAMP = Hampshire; HOG = Hog Island; JAC = Jacob; KAR = Karakul; KAT = Katahdin; LELW = Leicester Longwool; LINC = Lincoln; NAVC = Navajo Churro; POLY = Polypay; RAMB = Rambouillet; ROMA = Romanov; ROMN = Romney; SCI = Santa Cruz Island; SDN = Southdown; STC = Saint Croix; SUFF = Suffolk; TARG = Targhee; TEX = Texel; TUN = Tunis; WAR = Warhill. Color version available in the online PDF.

(Dorset, Dorper, Gulf Coast Native, Navajo Churro, and Rambouillet). These relatively high levels would indicate that several US populations have adequate genetic diversity to continue adapting populations to meet breeder/industry goals.

The results suggest the Warhill and Targhee should be considered genetically similar to Rambouillet; such a conclusion could result in a scaling back of collection goals for these populations. Across the analyses, there is justification for limiting the acquisition of samples for both Suffolk and Hampshire. A possible alternative would be to treat the 2 breeds as 1 population. However, concurrence with the respective breed associations is desirable before collection development is altered.

Two hair breeds (Barbados Blackbelly and Saint Croix) and hair breed composites (Dorper and Katahdin) were analyzed. It has been hypothesized that the Barbados Blackbelly and Saint Croix are of African origin. Muigai et al. (2002) reported that Barbados Blackbelly and Saint Croix had a closer association with Iberian Peninsula wool breeds rather than African hair breeds. Our results placed the Barbados Blackbelly and Saint Croix in a group of breeds known to have originated from the Iberian Peninsula (Gulf Coast Native, Navajo Churro, Rambouillet). Figure 3 illustrates considerable genetic distance between the Barbados Blackbelly-Saint Croix and the Dorper, suggesting differences exist between the fat-rumped/fat-tailed

breeds of eastern and southern Africa and the other hair breeds.

The study quantifies the genetic diversity found in the various breeds useful for in situ conservation. Results indicate there are breeds in need of more intensive conservation management than others (e.g., Hog Island, Tunis, Karakul, Black Welsh Mountain, Cotswold, and Southdown). Of these breeds, F_{is} for Barbados Blackbelly and Karakul are moderately high (>20%), and therefore inbreeding depression may be of concern, which has been shown in other species (MacNeil et al., 1989). In addition, breeds with F_{is} greater than 15% warrant close monitoring and potentially assistance in developing in situ conservation management strategies.

Because of the perceived importance of the Hog Island as a breed closely resembling breeds from colonial populations, efforts to manage current diversity levels are needed. Our results indicate that the Hog Island is more limited in genetic variation, based on the average number of alleles and observed heterozygosity. Positively, inbreeding level is not high. But if the need arises to broaden the genetic base for this breed, the present results suggest the Southdown as the most appropriate breed to accomplish that goal, based on its proximity to the Hog Island (Figure 3; Appendix Figure A2).

Sheep populations in the United States represent a wide variety of breed types that were originally developed to fulfill varying product or production system needs (Ryder, 1964; Wood and Orel, 2001). Molecular characterization confirmed the differentiation between many of the breed types. Therefore, a relatively broad range of ovine genetic resources is present based on number of alleles, heterozygosity, and genetic distance measures. As a result, these populations may have sufficient genetic diversity for breeders to continue adapting populations to meet breeder and industry goals. This study provided insight and the basis for recommendations concerning in situ and ex situ conservation of US sheep breeds that will be employed.

LITERATURE CITED

Baumung, R., V. Cubric-Curik, K. Schwend, R. Achmann, and J. Solkner. 2006. Cenetic characterization and breed assignment in Austrian sheep breeds using microsatellite marker information. J. Anim. Breed. Genet. 123:265–271.

Blackburn, H. D. 2009. Genebank development for the conservation of livestock genetic resources in the United States of America. Livest. Sci. 120:196–203.

Blackburn, H. D., and D. Gollin. 2009. Animal genetic resource trade flows: The utilization of newly imported breeds and gene flow of imported animals in the United States of America. Livest. Sci. 120:240–247.

Dalvit, C., M. De Marchi, E. Zanetti, and M. Cassandro. 2009. Genetic variation and population structure of Italian native sheep breeds undergoing in situ conservation. J. Anim. Sci. 87:3837–3844.

Diez-Tascón, C., R. P. Littlejohn, P. Almeida, and A. Crawford. 2000. Genetic variation within the Merino sheep breed: Analysis of closely related populations using microsatellites. Anim. Genet. 31:243–251.

Dohner, J. V. 2001. Historic and Endangered Livestock and Poultry Breeds. Yale Univ. Press, New Haven, CT.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14:2611–2620.

Falconer, D. S., and T. F. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Prentice Hall, Harlow, UK.

FAO. 2004. Secondary Guidelines for Development of National Farm Animal Genetic Resource Management Plans: Measurement of Domestic Animal Diversity (MoDAD): Recommended Microsatellite Markers. FAO, Rome, Italy.

FAO. 2007a. Global Plan of Action for Animal Genetic Resources and the Interlaken Declaration. FAO, Rome, Italy.

FAO. 2007b. The State of the World's Animal Genetic Resources for Food and Agriculture. FAO Rome, Italy.

Felsenstein, J. 2007. PHYLIP (phylogeny inference package) version 3.67. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle. Accessed March 2007. http://evolution.genetics.washington.edu/phylip.html.

Goudet, J. 2000. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9). Accessed Dec. 2009. http://www.unil.ch/izea/softwares/fstat.html.

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

Kijas, J. W., D. Townley, B. P. Dalrymple, M. Heaton, J. Maddox, A. McGrath, P. Wilson, R. G. Ingersoll, R. McCulloch, S. McWilliam, D. Tang, J. McEwan, N. Cockett, V. H. Oddy, F. W. Nicholas, H. Raadsma, and International Sheep Genomics Consortium. 2009. A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS ONE 4:e4668 10.1371/journal.pone.0004668.

Lawson Handley, L. J., K. Byrne, F. Santucci, S. Townsend, M. Taylor, M. W. Bruford, and G. M. Hewitt. 2007. Genetic structure of European sheep breeds. Heredity 99:620–631.

MacNeil, M. D., D. Dearborn, L. Cundiff, C. Dinkel, and K. Gregory. 1989. Effects of inbreeding and heterosis in Hereford females on fertility, calf survival and preweaning growth. J. Anim. Sci. 67:895–901.

Mason, I. L. 1996. A World Dictionary of Livestock Breeds, Types and Varieties. 4th ed. CAB Int., Wallingford, UK.

McKay, S. D., R. D. Schnabel, B. Murdoch, L. Matukumalli, J. Aerts, W. Coppieters, D. Crews, E. Dias Neto, C. Gill, C. Gao, H. Mannen, Z. Wang, C. Van Tassell, J. L. Williams, J. F. Taylor, and S. S. Moore. 2008. An assessment of population structure in eight cattle breeds of cattle using a whole genome SNP panel. BMC Genet. 9:37 doi:10.1186/1471-2156-9-37.

Meadows, J. R., I. Cemal, O. Karaca, E. Gootwine, and J. Kijas. 2007. Five ovine mitochondrial lineages identified from sheep breeds of the near east. Genetics 175:1371–1379.

Meadows, J. R., K. Li, J. Kantanen, M. Tapio, W. Sipos, V. Pardeshi, V. Gupta, J. H. Calvo, V. Whan, B. Norris, and J. W. Kijas. 2005. Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe. J. Hered. 96:494–501.

Minch E., A. Ruiz-Linares, D. Goldstein, M. Feldman, and L. L. Cavali-Sforza 1996. Microsat (version 1.5): A computer program for calculating various statistics on microsatellite allele data. Stanford University Medical Center, Stanford, CA.

Muigai, A., J. Hirbo, S. Sharkey, E. Rege, H. Blackburn, and O. Hannotte. 2002. Genetic diversity and relationships of hair sheep breeds of the Americas: First results. 7th World Congress on Genetics Applied to Livestock Production (Montpellier, FR). 33:573–576.

Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283–292.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6:288–295.

Pedrosa, S., M. Uzun, J. Arranz, B. Gutierrez-Gil, F. San Primitivo, and Y. Bayon. 2005. Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. Proc. Biol. Sci. 272:2211–2217.

Peter, C., M. Buford, T. Perez, S. Dalamitra, G. Hewitt, G. Erhart, and the ECONGENE Consortium. 2007. Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. Anim. Genet. 38:37–44.

Pritchard, J. K., M. Stephans, and P. Donnelly. 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.

Ryder, M. L. 1964. The history of sheep breeds in Britain. Agric. Hist. Rev. 12:1–12.

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

Wood, R. J., and V. Orel. 2001. Genetic Prehistory in Selective Breeding: A prelude to Mendel. Oxford Univ. Press, New York, NY.

APPENDIX

Table A1. Number of alleles per locus and percentage of missing data per locus[1]

Locus (chromosome No.)	No. alleles	% Missing data	Locus (chromosome No.)	No. alleles	% Missing data
BM1824 (1)	10	4.1	OarAE129 (5)	9	2.9
BM8125 (17)	8	3.3	OarCP34 (3)	8	22.0
DYMS1 (20)	21	3.9	OarCP38 (10)	12	12.7
HUJ616 (13)	23	3.3	OarFCB128 (2)	14	2.7
ILSTS11 (9)	9	17.9	OarFCB193 (11)	17	2.1
ILSTS28 (3)	16	9.5	OarFCB304 (19)	22	6.4
ILSTS5 (7)	12	16.8	OarFCB226 (2)	18	1.8
INRA063 (14)	21	13.0	OarHH47 (18)	15	5.0
MAF209 (17)	12	3.8	OarJMP29 (24)	20	5.5
MAF214 (16)	10	5.0	OarJMP58 (26)	20	1.2
MAF33 (9)	14	1.8	OarVH72 (25)	10	9.8
MAF65 (15)	11	2.6	SRCRSP1 (CHI13)	8	1.1
MAF70 (4)	24	3.6	SRCRSP5 (18)	6	3.3
MCM527 (5)	12	4.7	SRCRSP9 (12)	13	4.2

[1]Markers not used: MCM140 (6), OarFCB20 (2), BM1329 (6).

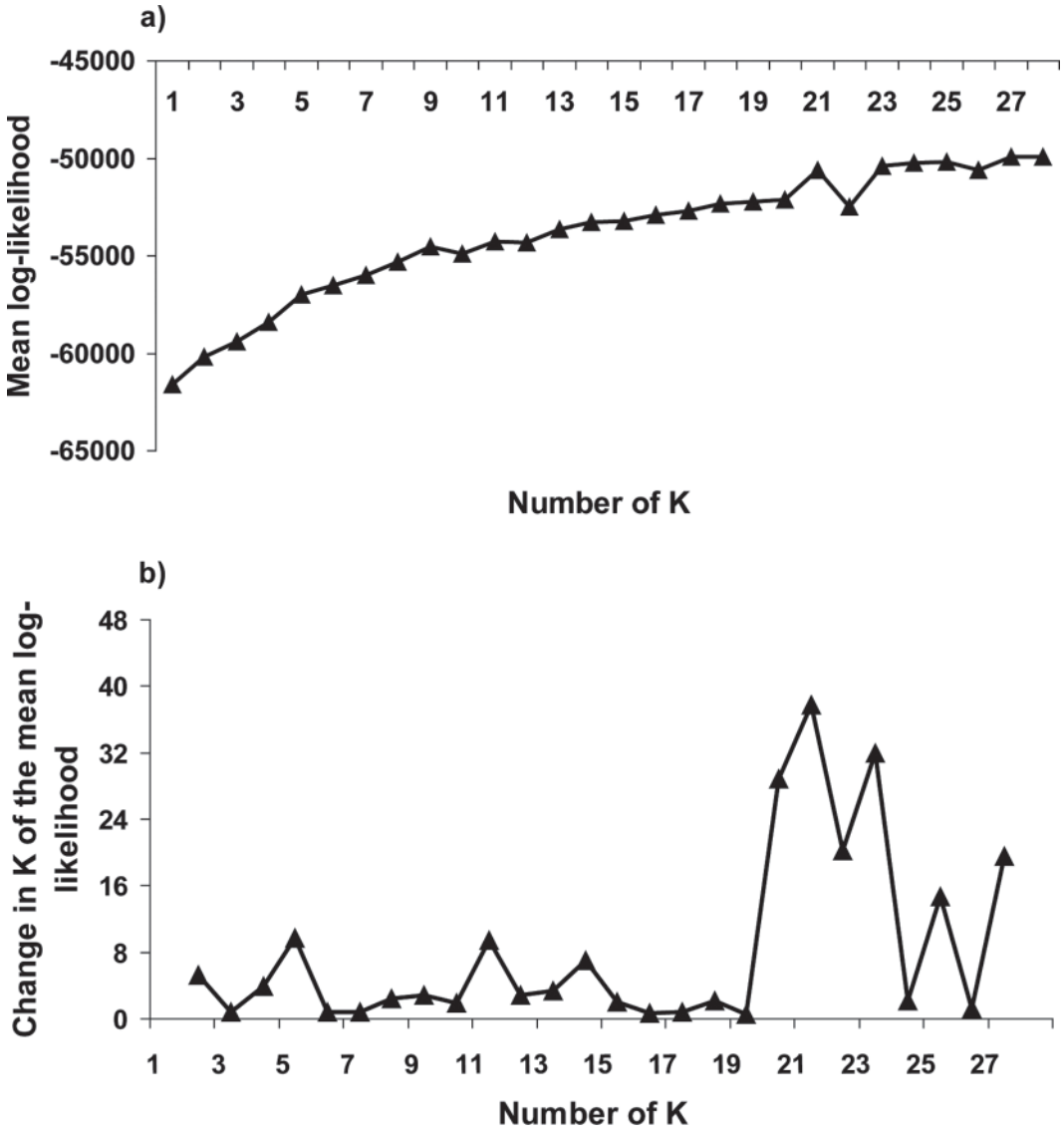

Figure A1. a) Mean estimated log-likelihood probabilities for 28 clusters obtained from microsatellite markers in 28 US sheep breeds and b) ΔK from the mean log-likelihood probabilities from STRUCTURE runs where inferred clusters (K) ranged from 1 to 28.

Table A2. Number of private alleles and identification of the alleles per loci unique to a breed

Breed	Private alleles	Locus	Allele	Frequency	Animals with private alleles, n (% of total sample)
Barbados Blackbelly	1	MAF214	240	0.042	1 (5.5)
Black Welsh Mountain	0	—	—	—	0 (0.0)
Columbia	2	MAF70	185	0.025	2 (9.5)
		OarJMP58	181	0.024	
Cotswold	0	—	—	—	0 (0.0)
Dorper	2	OarFCB128	134	0.045	4 (9.0)
			136	0.011	
Dorset	3	INRA063	198	0.019	6 (11.1)
		INRA063	212	0.038	
		OarJMP29	136	0.056	
Finnsheep	3	MAF70	141	0.079	6 (30.0)
		OarCP38	154	0.031	
		OarJMP58	197	0.026	
Gulf Coast Native	4	ILSTS28	178	0.019	6 (20.8)
		OarCP38	158	0.017	
		OarFCB226	177	0.018	
		OarJMP58	189	0.033	
Hampshire	0	—	—	—	0 (0.0)
Hog Island	3	OarFCB204	207	0.087	4 (12.5)
		SRCRSP9	133	0.025	
		SRCRSP9	151	0.025	
Jacob	1	SRCRSP9	149	0.022	1 (4.2)
Karakul	1	HUJ616	159	0.026	1 (5.3)
Katahdin	3	BM8125	126	0.017	2 (10.3)
		DYMS1	193	0.017	
		OarFCB226	179	0.017	
Leicester Longwool	3	ILSTS11	290	0.034	4 (11.4)
		OarHH47	136	0.019	
		OarJMP58	131	0.034	
Lincoln	1	OarAE129	161	0.029	1 (5.3)
Navajo Churro	4	MAF70	187	0.018	4 (12.9)
		OarAE129	183	0.016	
		OarFCB128	120	0.017	
		OarFCB193	123	0.017	
Polypay	1	DYMS1	223	0.063	2 (5.9)
Rambouillet	4	HUJ616	154	0.011	4 (8.7)
		ILSTS28	162	0.012	
		MAF65	134	0.011	
		MCM527	175	0.011	
Romanov	2	OarCP38	130	0.036	5 (8.3)
		OarFCB226	149	0.083	
Romney	1	INRA063	214	0.063	2 (5.0)
Santa Cruz Island	0				0 (0.0)
South Down	0				0 (0.0)
St. Croix	1	SRCRSP5	164	0.038	1 (3.8)
Suffolk	2	BM1824	198	0.021	2 (7.7)
		MCM527	182	0.020	
Targhee	0				0 (0.0)
Texel	5	DYMS1	179	0.028	4 (25.0)
		ILSTS28	176	0.031	
		ILSTS5	232	0.038	
		OarAE129	185	0.028	
		OarJMP29	150	0.056	
Tunis	3	MAF209	151	0.615	10 (21.4)
		MAF33	147	0.038	
		SRCRSP1	143	0.154	
Warhill	3	ILSTS28	144	0.022	4 (11.5)
		MAF70	145	0.022	
		OarJMP29	142	0.300	

Figure A2. Nei's neighbor joining tree (1,000 bootstraps).