Journal of Heredity 2005:96(5):494–501
doi:10.1093/jhered/esi100
Advance Access publication August 31, 2005

© The American Genetic Association. 2005. All rights reserved.
For permissions, please email: journals.permissions@oupjournals.org.

Mitochondrial Sequence Reveals High Levels of Gene Flow Between Breeds of Domestic Sheep from Asia and Europe

J. R. S. Meadows, K. Li, J. Kantanen, M. Tapio, W. Sipos, V. Pardeshi, V. Gupta, J. H. Calvo, V. Whan, B. Norris, and J. W. Kijas

From CSIRO Livestock Industries, Level 5 Queensland Bioscience Precinct, 306 Carmody Rd., St. Lucia 4067, Australia (Meadows, Whan, Norris, and Kijas); Department of Gene and Cell Engineering, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 10094, P.R. China (Li); Animal Production Research, MTT Agrifood Research Finland, Jokioinen, FIN-31600, Finland (Kantanen, Tapio); Clinical Department for Farm Animals and Herd Management, University of Veterinary Medicine, Vienna, A-1210 Austria (Sipos); Biochemical Sciences Division, National Chemical Laboratory, Pune, 411 008, India (Pardeshi, Gupta); and Unidad de Tecnologia en Produccion Animal, CITA, Zaragoza, 50059, Spain (Calvo).

Address correspondence to Dr. James W. Kijas at the address above, or e-mail: James.Kijas@csiro.au.

Abstract

Sequence variation present within the mitochondrial genome was used to investigate genetic diversity within sheep breeds from Asia and Europe. Comparison of 2027 bp of sequence from 121 animals revealed 44 phylogenetically informative nucleotide positions and a single insertion/deletion. A total of 57 haplotypes were observed which formed two distinct clades. Type A haplotypes were found in breeds from Asia (India, Indonesia, Mongolia, and Tibet), while type B haplotypes were observed at the highest frequency in breeds sourced from Europe (nine breeds from Austria, Åland, Finland, Spain, and northwestern Russia). The distribution of haplotypes indicates sheep appear to have the weakest population structure and the highest rate of intercontinental dispersal of any domestic animal reported to date. Only 2.7% of the sequence variation observed was partitioned between continents, which is lower than both goat (approximately 10%) and cattle (approximately 50%). Diagnostic restriction fragment length polymorphism polymerase chain reaction (RFLP-PCR) tests which distinguish type A and B haplotypes were used to test an additional 223 animals from 17 breeds of European and Asian origin. A mixture of the two lineages was found in every breed except Suffolk and the Indian Garole, indicating introgression has played a major part during breed development and subsequent selection.

Sheep are a highly adaptable and versatile domestic species, which has made them a critically important resource in human societies around the world. Following Mesolithic mans' domestication of the sheep approximately 8000–9000 years ago (Ryder 1984), selection has proceeded on traits such as coat color, environmental tolerance, wool characteristics, and meat and milk production. The result is a spectrum of phenotypic differences between breeds. Molecular genetics has proven highly informative for investigating the relationships between animal populations as well as for documenting the levels of genetic variation resident within breeds. Variation within autosomal microsatellites has been used successfully to make inferences about population history (Forbes et al. 1995; Walling et al. 2004) and to examine the relationship between sheep breeds from Europe (Arranz et al. 1998, 2001; Diez-Tascon et al. 2000; Tapio et al. 2003) and Asia (Chu et al. 2003). In most cases, analysis of the observed variation returned results broadly consistent with both historical and geographic knowledge for the breeds investigated.

A separate approach has focused on the maternal origins of both domestic sheep and their candidate wild progenitors via analysis of the mitochondrial genome. The major finding within domestic sheep is a biphyletic pattern where mitochondrial haplotypes form two distinct clades (Wood and Phua 1996). These are termed clades A and B (Hiendleder et al. 1998b, 2002; Wood and Phua 1996). Clade A has been found in two breeds from central Asia (Karakul/Kazakstan and Gizarr/Tadjikistan) as well as three breeds sampled in New Zealand (Romney, Coopworth, and Merino). Clade B haplotypes have been observed in a range of breeds from Europe, the Near East, and New Zealand, and includes sequences derived from European mouflon (*Ovis musimon*;

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024

Table 1. Summary statistics derived from mtDNA sequence for 10 populations of sheep

	European								Asian	
Sheep breeds[a]	ALD	FIN	VEP	VEN	TMS	CS	FS	TSS	JTT	TIB
Summary statistics[b]										
n	5	5	5	6	18	18	12	13	18	6
Psites	15	10	12	6	19	9	16	20	15	10
Hn	4	5	4	2	9	8	7	9	7	4
Hd	0.90	1	0.90	0.33	0.89	0.75	0.83	0.94	0.80	0.80
D	7.80	4.60	5.60	2.00	6.44	3.18	5.26	5.51	4.70	3.53
π	3.65	1.97	2.76	0.99	3.18	1.43	2.45	2.47	2.11	1.74

[a] Breed names are abbreviated as Åland (ALD), Finnsheep (FIN), Vepsia (VEP), Viena (VEN), Tyrolean mountain sheep (TMS), Carynthian sheep (CS), forest sheep (FS), Tyrolean stone sheep (TSS), Javenese thin tail (JTT), and Tibetan (TIB).

[b] *n* is number of individuals sampled; Psites is the number of polymorphic SNPs; Hn is the number of haplotypes observed including indels; Hd is the haplotype diversity including indels, alternatively described as the probability that two randomly selected haplotypes are different within the population; *D* is the average number of differences (including indels) between haplotypes in the population; π is the nucleotide diversity excluding indels $\times 10^{-3}$.

Hiendleder et al. 2002). Together this dual clustered pattern strongly suggests that multiple domestication events must have occurred where the mouflon is a possible progenitor of clade B haplotypes, while the maternal ancestor of sheep carrying type A haplotypes remains unknown.

The aim of this study was to investigate genetic diversity in breeds with different population histories and from distinct geographic regions. Individuals from a number of Asiatic breeds were collected from within their country of origin. This included animals from Indonesia (Javanese Thin Tail), India (Garole), Tibet, and Mongolia, where in each case breed development has occurred locally in response to specific climatic and production demands. This also applies to a set of European breeds where individuals were sourced from Austria (Carynthian sheep, Tyrolean mountain sheep, Tyrolean stone sheep, and forest sheep), Spain (latxa), and Finland (Finnsheep). Three populations of unimproved sheep were collected within Europe where animals are kept without registration or implemented breeding schemes. The Åland sheep originated from the Åland islands located between Sweden and Finland, and the Viena and Vepsia sheep from Russian Karelia. A separate set of animals was investigated where development or selection originally occurred in one location (usually Europe) before their importation and continued selection took place in other parts of the globe. Examples include Australian populations of Coopworth, Poll Dorset, Polwarth, Suffolk, and merino. The aim of this work was to search for the presence of additional major mitochondrial lineages and quantify the population structure and levels of introgression present within sheep populations from Asia and Europe.

Materials and Methods

Animals

Samples from 121 animals were used for sequence analysis of the mitochondrial genome (Table 1). Sample collection, DNA extraction, and breed characteristics for the Finnsheep (FIN, $n = 5$), Åland (ALD, $n = 5$), Russian Vepsia (VEP, $n = 5$), Russian Viena (VEN, $n = 6$), and Oxford Down (OXD, $n = 2$) are described elsewhere (Tapio et al. 2003). Similarly the Tyrolean mountain sheep (TMS, $n = 18$), Carynthian sheep (CS, $n = 18$), forest sheep (FS, $n = 12$), and Tyrolean stone sheep (TSS, $n = 13$) have been described previously (Sipos et al. 2002). The Mongolian sample (MON, $n = 1$) was collected from the Buryatia region near Lake Baikal which borders northern Mongolia, while the samples from the Tibetan breed (TIB, $n = 6$) were sourced from the northern and eastern parts of the country. Latxa (LAX, $n = 2$) and wild mouflon (MFL, $n = 4$) were collected from the Basque and Cuenca regions of Spain, respectively (Calvo et al. 2004) and the Javanese thin tail (JTT, $n = 18$) from Bogor, West Java. The Karakul (KAK, $n = 1$), a breed originating in central Asia, Romney (ROM, $n = 1$), merino (MER, $n = 1$), and Suffolk (SUF, $n = 3$) were all obtained from within Australia. A separate set of 223 animals was used for restriction fragment length polymorphism polymerase chain reaction (RFLP-PCR) analysis of the mitochondrial genome (Table 2). This included Garoles, which were originally from West Bengal (GAR, $n = 8$) and have been described previously (Davis et al. 2002). The remaining 215 animals represent 16 breeds collected within Australia. The largest sample set (merino, $n = 94$) was sourced from 20 separate breeders in different locations from across Australia. Unless otherwise stated, genomic DNA was extracted from blood using the QIAamp DNA Mini Kit (Qiagen, Australia) following the manufacturers instructions.

Sequencing and Genotyping of the Ovine Mitochondria

Two regions of the mitochondrial DNA (mtDNA) genome were sequenced using primers designed from the complete ovine mtDNA (AF010406; Hiendleder et al. 1998a). Primers CytB-F 5′-GTCATCATCATTCTCACATGGAATC-3′ and CytB-R 5′-CTCCTTCTCTGGTTTACAAGACCAG-3′ were used to amplify a 1272 bp region of the mitochondrial cytochrome *b* gene (AF010406 positions 14078 to 15349). Primers mtCR-F2 5′-AACTGCTTGACCGTACATAGTA-3′ and mtCR-R1 5′-AGAAGGGTATAAAGCACCGCC-3′ were used to amplify a 1246 bp fragment spanning part of the control region, the tRNA-Phe and 12S rRNA coding

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024

Table 2. Frequency of SNP haplotypes within 17 sheep populations

Breed	Animals tested	Frequency of SNP combination[a]		
		T291/A16454 type A	C291/G16454 type B	C291/A16454 type B
Australian sheep populations				
Black Suffolk	12	—	1.00	—
Border Leicester	3	0.33	0.33	0.33
Coopworth	10	0.50	0.30	0.20
Drysdale	1	1.00	—	—
English Leicester	11	0.36	0.64	—
Lincoln	13	0.31	0.54	0.15
Merino	94	0.47	0.34	0.19
Perendale	4	1.00	—	—
Poll Dorset	10	0.20	0.60	0.22
Polwarth	12	0.75	0.17	0.08
Shropshire	1	1.00	—	—
Southdown	10	0.20	0.50	0.30
Texel	6	0.17	0.50	0.33
Romney	16	0.81	0.06	0.13
Whiltshire Horn	3	0.67	0.33	—
White Suffolk	9	—	0.89	0.11
Indian population				
Garole	8	1.00	—	—
Total	223	0.45	0.40	0.15

[a] SNP combinations are composed of the genotype at position C291T and A16454G. Each are diagnostic of a subset of sequence haplotypes as follows: T291/A16454 (type A H51–57), C291/G16454 (type B H1, H3–H12, H14, H16–H22, H27–H50), and C291/A16454 (type B H2, H13, H15, H23–H26).

RNA genes (AF010406 positions 15983 to 592). The two primer sets were used for PCR amplification as described previously (Meadows et al. 2004), before products were directly sequenced using ABI Prism Big Dye Terminator 3.1 chemistry and an ABI 377 Prism DNA autosequencer (Applied Biosystems, Australia).

A rapid method was developed to distinguish the two major mtDNA lineages using single nucleotide polymorphism (SNP) genotyping. Polymorphisms at positions 16454 (control region) and 291 (12S rRNA gene) were assayed following enzymatic digestion of PCR products amplified using primer pair mtCR-F2/mtCR-R1. *Rsa*I (GT/AC) digestion resulted in 965 bp and 281 bp fragments where A is present at position 16454 or 516 bp, 449 bp, and 281 bp fragments where G is present. *Bst*NI (CC/WGG) digestion of the same PCR product resulted in 634 bp and 612 bp fragments where T is present at position 291 or 612 bp, 313 bp, and 321 bp fragments where C is present. Digested fragments were separated on 1.5% agarose gels and visually scored to determine the genotype at positions C291T and A16454G.

Analytical Methodologies

Sequence data from each PCR fragment was imported into Sequencher 4.2 (Gene Codes Corp., Ann Arbor, MI). Traces were trimmed to retain the maximum length sequence with reads in both the forward and reverse direction for all animals. The resulting data were aligned (CytB-F/R 967 bp; mtCR-F2/R1 1060 bp) and inspected using MEGA 2.0 (Kumar et al. 2001) to discriminate between singletons (polymorphic positions that appeared in only one animal) and sites that appeared at least twice. Singletons were removed by replacement with the consensus base, as they were not considered phylogenetically informative within the dataset and may represent sequencing artifacts. Sequences were deposited into GenBank (accession numbers AY879343–AY879584) prior to removal of singletons to allow an unbiased comparison with other datasets. Indices of sequence variation and haplotype structure were calculated using DnaSP 4.0 (Rozas et al. 2003). This included estimation of nucleotide diversity (π), the likelihood that a nucleotide position will be polymorphic when compared between individuals sampled at random from the population. The mismatch distribution is the number of differences that exist between every pairwise combination of haplotypes plotted as a function of frequency. The calculation of genetic distance using Kimura's two-parameter method, construction of the neighbor-joining haplotype tree and bootstrap analysis were performed with MEGA 2.0.

To estimate sequence diversity between the two major mtDNA lineages, the average number of nucleotide differences (D) and the average number of nucleotide substitutions per site (K) were estimated. Animals were first grouped as having either a type A (H51–57, 24 animals) or a type B haplotype (H1–50, 97 animals). Sequence diversity (D and K) between the populations was then calculated using DnaSP 4.0. The distribution of sequence variation was calculated using an analysis of molecular variance (AMOVA) (Excoffier et al. 1992) as implemented in Arlequin (http://anthropologie.unige.ch/arlequin/). To investigate the distribution of variation between geographic regions, two hierarchical groupings

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024

		POPULATION																		
		European Breeds									Asian Breeds			Imported Breeds					Wild	
		ALD	FIN	VEP	VEN	LAX	TMS	CS	FS	TSS	JTT	TIB	MON	KAK	OXD	ROM	MER	SUF	MFL	SUM
H1	CAACCCCATTCGTGGCATCTG-TGCCCCCATCTTAGGGCCTATTG			2			3	1		2										8
H2	....T.....A-.....G.....	1																		1
H3	...T.....C...-.A.....A	2																		2
H4	.....C...T.....A....C....							2												2
H5	.....-.....G.....C....							1												1
H6	.....C.....-..T.....A....C....							1												1
H7	.....-.....A....T...							9												9
H8	.....T.....A....T...							1												1
H9	.....C...-.....A....C....							2												2
H10	.....C...-.....A..T.C....							1												1
H11	.....-.....C..								5											5
H12	.....A.....-.....C.....								1											1
H13	..G.....A-.....								1											1
H14	.....G.....C...T.....								1											1
H15	..G.....AT.....								1											1
H16	.....G....C.....-.....						2		1											3
H17	T.....T.....C.										4									4
H18	T.....-.....C.										7									7
H19	T.....-...T.....C.										1									1
H20	.....T.....										1									1
H21	T.....G.-.....C.										1									1
H22	.....TG....-.....C....				5															5
H23	.....A-.....G.....C....			1																1
H24	.....T.....A-.....G.....C....			1																1
H25	....T.....AT.....G.....	1																1		2
H26	.....A-.....																	1		1
H27	.....A.....-.....		1																	1
H28	.....T.....T.....C....		1																	1
H29	.....TG....-.....C.....C....		1																	1
H30	.....T.A.....		1																	1
H31	...T..T.....-.....C.....				1															1
H32	.....AA.....-.....T.....		1																	1
H33	.....G.....T..-.....													1						1
H34	.....C.....-.....						1								1					2
H35	.....T..-.....C.....														1					1
H36	T.....-..T.....						1													1
H37	.....T.....-..T.....						3													3
H38	.....A.....-.....C....						1													1
H39	.....G.-...T.....						1													1
H40	.....GC.....-.....A.....									2										2
H41	.....GC.....T.....A.....									1										1
H42	.G.....-.....C.....T.....									3										3
H43	.G.....T.....C.....T.....									1										1
H44	T.....-.....									1										1
H45	.....A.....T.....T.....									1										1
H46	.G.....T.....C....TT.....									1										1
H47	.....-.....T.....											1								1
H48	.....C...-.....GG.....C....					1														1
H49	.....-.....G.....					1													2	3
H50	.....C.....-.....G.....																		2	2
H51	TG...TT.....C..A-..T.T....T...A.....	1		1					2		3	3	1			1	1	1		14
H52	TG.T.TT.....C....C..A-..T.T....T...A.....										1									1
H53	TG...TT...T.....C..A-..T.T....T...A.....						5													5
H54	TG...TT..C.....C..A-C.T.T....T...A.....						1													1
H55	TG...TT..C.....C..ATC.T.T....T...A.....									1										1
H56	TG...TT.....T.C..A-..T.T....T...A.....											1								1
H57	TG...TT.....A-..T.T....T...A.....											1								1
		5	5	5	6	2	18	18	12	13	18	6	1	1	2	1	1	3	4	**121**

Figure 1. Haplotype sequence and distribution in sheep populations with European and Asian origins. Vertical numbers indicate the SNP position relative to the reference sequence AF010406. A dash (−) indicates the position of the indel. Breed names are abbreviated as Åland (ALD), Finnsheep (FIN), Vepsia (VEP), Viena (VEN), Latxa (LAX), Tyrolean mountain sheep (TMS), Carynthian sheep (CS), forest sheep (FS), Tyrolean stone sheep (TSS), Javenese thin tail (JTT), Tibetan (TIB), Mongolian (MON), Karakul (KAK), Oxford Down (OXD), Romney (ROM), merino (MER), Suffolk (SUF), and Mouflon (MFL).

were imposed on sequence data derived from 113 animals (the 8 imported animals as defined in Figure 1 were excluded). Breeds were grouped as either European or Asian (defined in Figure 1). The mouflon was included as a breed within the European group. To investigate the component of variation attributable to introgression between the geographic regions, a second hierarchical grouping ignored breed membership. The 113 individuals were first sorted into four populations using mitochondrial haplotype (type A or B) and geography (Asian or European origin). The four populations were then grouped according to haplotype prior to calculation of the distribution of variance.

Results

mtDNA Sequence Variation and Genetic Diversity Between Populations

A total of 2027 bp of mitochondrial sequence was collected from each of 121 animals. The region sequenced was composed of the cytochrome *b* gene (*CytB*; 967 bp), the control region (CR; 525 bp), and the tRNA-Phe and 12S rRNA coding regions (RNA; 535 bp). Alignment of the 2027 positions revealed a total of 81 putative SNPs and a single insertion/deletion (indel). Inspection of the SNPs revealed that 37 were observed in only one animal (singletons) and these were

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024

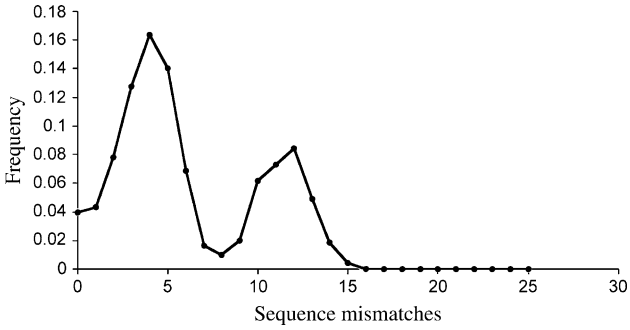


Figure 2. Frequency distribution of the number of sequence mismatches between pairwise combinations of 57 ovine mtDNA haplotypes. The distinct peaks indicate two groups of haplotypes exist.

subtracted from the dataset. The remaining 44 phylogenetically informative SNP positions were used to calculate nucleotide diversity from the complete 2027 bp ($\pi = 3.03 \pm 0.16 \times 10^{-3}$). Comparison between the three mitochondrial segments showed the control region contained both the highest number of SNPs and the highest nucleotide diversity (*CytB* SNP = 16, $\pi = 2.05 \pm 0.12 \times 10^{-3}$; CR SNP = 24, $\pi = 7.02 \pm 0.50 \times 10^{-3}$; RNA SNP = 4, $\pi = 0.90 \pm 0.11 \times 10^{-3}$).

The 44 SNPs and the single indel defined 57 unique haplotypes within the 121 individuals sequenced (Figure 1). This included animals from nine European breeds ($n = 84$), three Asian breeds ($n = 25$), five breeds imported into Australia ($n = 8$), and the European mouflon ($n = 4$). The genetic variation resident within each population (where $n \geq 5$) is presented in Table 1. In every breed where multiple animals were sampled, more than one haplotype was observed. Nucleotide diversity varied between breeds with the highest observed within Åland sheep ($\pi = 3.65 \times 10^{-3}$) and the lowest within the Viena breed of northwestern Russia ($\pi = 0.99 \times 10^{-3}$, Table 1).

Haplotype Structure and Distribution

Inspection of the distribution of haplotypes across breeds revealed the majority were confined to a single breed (51 of 57), while six haplotypes were present across multiple breeds (H1, H16, H25, H34, H49, and H51; Figure 1). To interrogate haplotype structure, the number of differences between each pairwise combination of the 57 sequences was calculated. The distribution of these differences shows two distinct peaks (Figure 2), which indicates the presence of divergent groups of haplotypes. This was supported following construction of a haplotype tree that contains two separate clusters (Figure 3). Haplotypes 1–50 form one large group that contains an example of a type B major ovine lineage, as described by others (Hiendleder et al. 2002; Wood and Phua 1996). This is separate from the remaining haplotypes (H51–57), which form a separate group likely to correspond with the type A lineage. Type A sequences were previously found in single animals from both Kazakhstan and Tadjikistan (Hiendleder et al. 2002). Consistent with the "Asian" origin of type A haplotypes, animals from Indonesia, Tibet, and Mongolia had closely related type A sequences (H51, H52, H56, and H57; Figures 1 and 3). Clade B sequences have previously been found within European breeds and the mouflon (Hiendleder et al. 2002; Wood and Phua 1996). The majority of European animals in this study contained a type B haplotype (73 of 84 animals from nine breeds; Figures 1 and 3). In addition, each of the four mouflon contained mtDNA sequences (H49 and H50) that clustered with type B haplotypes (Figure 3).

The sequence diversity separating the two major mtDNA lineages was calculated from the total dataset (2027 bp) and within each of the three component regions (CR, 525 bp; *CytB*, 967 bp; RNA, 535 bp). As expected, the average number of differences (D) and sequence diversity (K) between lineages was much higher within the control region compared with the other mtDNA components (CR $D = 7.52$, $K = 1.44\%$; *CytB* $D = 2.78$, $K = 0.29\%$; RNA $D = 1.10$, $K = 0.21\%$). The average across the three components ($D = 11.41$, $K = 0.56\%$) is lower than previous estimates derived from the entire control region of both domestic and wild species (3.3% to 5.6%; Hiendleder et al. 1998b, 2002).

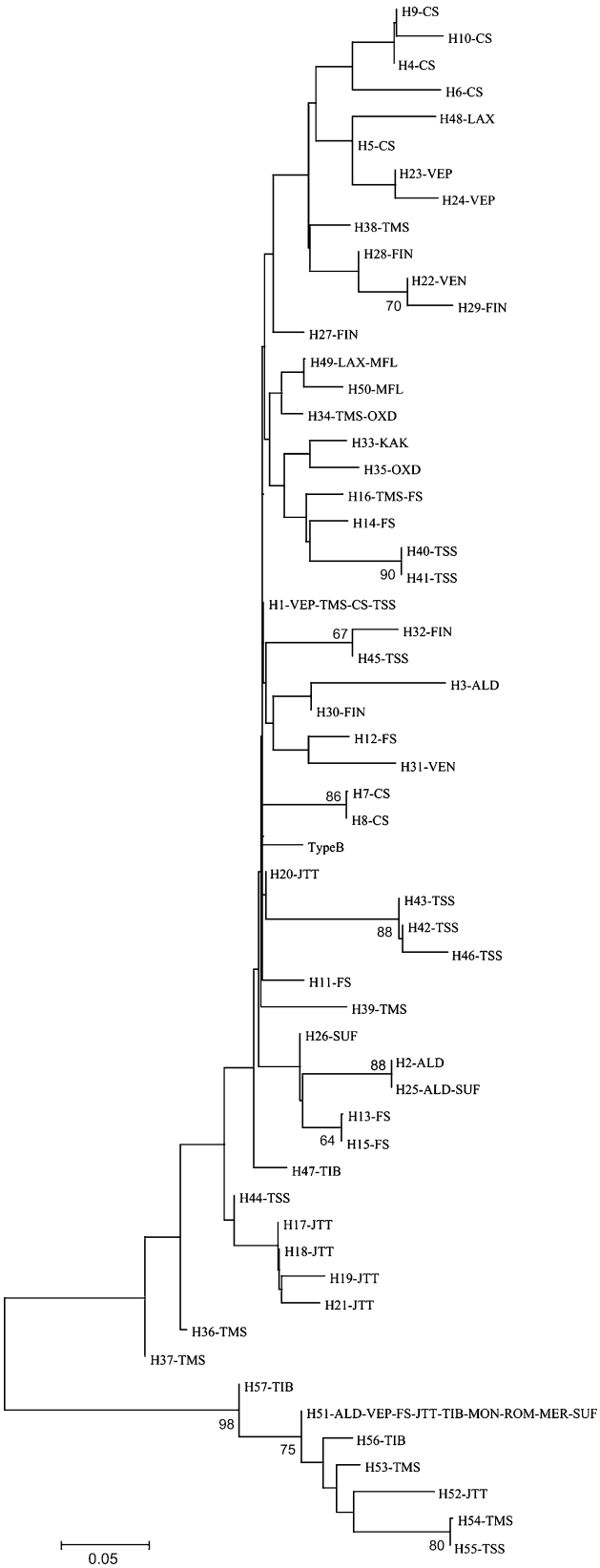
Molecular Evidence of Introgression and Population Structure

The correlation between sequence type and the geographic origin of breeds (type A in Asia and type B in Europe) is far from complete. Inspection of European animals revealed the presence of four type A sequences within animals collected from Åland (H51), northwestern Russia (Vepsia H51), and Austria (Tyrolean mountain sheep H53 and H54; Carynthian sheep H51; and Forest sheep H55; Figure 1). Conversely, sequencing of Asian animals revealed the presence of six type B haplotypes within the Javanese thin tail (H17–H21) and Tibetan breeds (H47; Figure 1). The degree of population structure was assessed by calculation of the distribution of sequence variation following imposition of two hierarchical groupings. When breeds were grouped according to geographic location (either Asian or European), the vast majority of diversity was present within breeds (81.5%), while only a fraction was diagnostic of the two geographic regions (2.7%). The remaining variation was present between breeds within each of the two regions (15.8%). To estimate the component of variation attributable to introgression occurring between the regions, a second hierarchical grouping removed information relating to breed membership. This resulted in an elevation in the component of variation existing between groups (7.9%).

SNP-Based Analysis of mtDNA Lineages

To investigate further the frequency of the two major mtDNA lineages, two mtDNA nucleotide positions were used for more rapid SNP-based genotyping in a second set of sheep. Selection of the two sites was based on the sequence dataset from 121 animals (Figure 1). Inspection of the

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024



sequence revealed all animals that carry the type A haplotype (n = 24) contain T at position 291 (T291) and A at position 16454 (A16454). All animals that carry the type B haplotype (n = 97) contain C291 and are further defined by either A (n = 8) or G (n = 89) at position 16454 (Figure 1). Enzymatic digestion assays were developed and used to genotype 223 animals from 17 breeds (Table 2). This included 16 breeds originally developed in Europe and subsequently imported and sampled from Australian flocks, as well as the Indian Garole. Examination of the SNP genotypes revealed three combinations were present (C291/G16454, C291/A16454, and T291/A16454). The fourth theoretical combination (T291/G16454) was not observed, consistent with the sequence dataset (Figure 1). The SNP combination diagnostic of type A mtDNA sequences (T291/A16454) had the highest frequency within the 223 animals and was widely distributed, being present in all but 2 of the 17 breeds. It appeared fixed within the Indian Garole and was absent only from the Black and White Suffolk. The type A haplotype occurred at moderate frequency (0.3–0.6) within Coopworth, English Leicester, merino, and Lincoln, and at high frequency (>0.6) within Polwarth and Romney.

Discussion

Sequence variation within the mitochondrial genome has proven highly informative for investigation of the origin of domestic animal species. Studies of cattle (Bradley et al. 1996), goat (Luikart et al. 2001), pigs (Giuffra et al. 2000), and sheep (Hiendleder et al. 2002) have revealed each has multiple maternal origins as evidenced by the presence of divergent groups of mtDNA sequence. Domestication events are thought to have occurred in Asia (buffalo and pigs), the Near East (cattle, sheep, and goats), and the Americas (alpaca and llama) (reviewed by Bruford et al. 2003). This study supports previous findings indicating sheep have two distinct lineages, each of which likely arose from geographically separate domestication events. The alternate interpretation where domestication occurred only once from an ancestral population carrying two divergent mtDNA types cannot be formally excluded, however, the level of divergence separating the major lineages significantly predates the depth of domestic history both for sheep (Hiendleder et al. 2002) and other domestic species that display strong population structure (MacHugh and Bradley 2001).

Sheep breeds in this work were sourced from Tibet, Mongolia, and Indonesia for mtDNA sequencing, and in each

←

Figure 3. Neighbor-joining tree of 57 mtDNA haplotypes (H1–H57) found within 117 domestic sheep and the European mouflon. Branch topography supported by bootstrap values in excess of 60% are indicated. The breed names are abbreviated as defined in Figure 1. The major mtDNA lineage B is represented (denoted as type B) by inclusion of sequence AF010406 (Hiendleder et al. 1998b).

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024

case haplotypes were identified that clustered together (type A) and separate from those of European origin. Breeds developed in Austria, northwestern Russia, Spain, and Finland predominantly contained animals with a second lineage (type B). Sequence data were collected from 17 different domestic ovine populations and these two lineages were the only ones identified (Figure 3). Additional breeds representing China and central Asia need to be examined, however, this study indicates additional major lineages may not be present. This would differentiate the sheep from goats, which have at least three major mtDNA clades within breeds from Asia and Europe (Luikart et al. 2001).

The level of sequence divergence that has accumulated between the two ovine lineages was calculated from three different segments of the mtDNA genome (CR K = 1.44%; *CytB* K = 0.29%; RNA K = 0.21%). Each region contained lower divergence than previous estimates (3.3% to 5.6%; Hiendleder et al. 1998a, 2002). Two factors likely account for this observation. First, previous estimates were derived from the complete control region, which includes regions known to accumulate substitutions at a higher rate compared to the mitochondrial components assayed in this study (Pesole et al. 1999). In addition, the inclusion of both domestic and wild species in previous estimates is expected to result in a higher sequence divergence.

This study provides molecular evidence for high levels of transportation of sheep between major geographic regions. Evidence for this was clear within the sequence dataset, where a number of European breeds contained animals carrying type A haplotypes (Figures 1 and 3). This strongly indicates the occurrence of introgression, however, the source could not be determined. The contributor to European animals of haplotype H51 remained unclear, as it is present within all three of the Asian breeds tested. In addition, the maternal contributor of haplotypes H53–H55 could not be identified, as they were not present in any of the Asian animals sampled. A high level of introgression was also revealed by investigation of 17 breeds sampled within Australia (16) and India (1) using SNP diagnostic for the two lineages A and B. The ancestors of 16 of the breeds were originally developed in Europe before populations were imported to Australia. The two SNP combinations diagnostic of type B haplotypes had a combined frequency of 0.57 (122 of 215 animals; Table 2). The remaining 43% of sheep carried a type A haplotype, suggesting widespread and repeated introduction of Asian animals during the development of every breed except Black and White Suffolk. Testing a cross section of Australian merinos showed approximately half had an Asian maternal origin.

The presence of type A haplotypes may be explained by reported introductions into merino lines of both the Garole from India (reviewed by Davis et al. 2002) and Chinese breeds carrying prolificacy genes. This is supported by the preliminary finding in this study that all of the Garole samples tested contained the SNP combination diagnostic of the type A lineage (T291/A16454). Whatever the source of this introgression into merinos, the observation of such a high frequency of type A haplotypes (0.47; Table 2) indicates the population can be considered a deep intercross with animals likely to contain genomic regions that arose from different domestication events. Interestingly, evidence for introgression in the opposite direction was also observed, as Asian animals were detected carrying type B sequences. The Javanese thin tail contained five different type B sequences within the 18 animals tested, suggesting introduction from different sources. Together this indicates that mixture of the two lineages is likely to be the case within every ovine breed in Asia and Europe. Such high levels of gene flow account for the weak population structure observed following calculation of the distribution of variation. Only 2.7% of the total variation existed between continental regions. This is the lowest of any domestic animal when compared with goats and cattle, where 10.6% and more than 50% of variation was present between continents, respectively (Bradley et al. 1996; Luikart et al. 2001).

The analysis of mtDNA sequences was insightful for assessing introgression, but less informative for establishing detailed relationships between breeds. Inspection of the haplotype tree (Figure 3) shows that definition within each major branch is poor, with few nodes supported by meaningful bootstrap values. This likely results from the small number of sequence differences that distinguish most haplotypes. In addition, the majority of haplotypes were observed only once, however, six sequences were observed across multiple populations (Figure 1). Haplotype 1 is present within the Vepsia breed of northwestern Russia and three of the four Austrian breeds, however, no introduction of Russian animals into Austrian populations is known to the authors. Although geographic neighbors, the four Austrian breeds appeared fairly genetically distinct, which is not surprising, as each are closed populations resulting from strong selection in an alpine environment. The Tyrolean mountain sheep and forest sheep both contained haplotype 16, suggesting shared ancestry. This is consistent with their breed history, as both are thought to be derived from the extinct Zaupel sheep (Sambraus 1987).

Interestingly, all three Asian breeds shared a haplotype (H51), which likely indicates the movement of animals has occurred during breed development. Analysis of the single Karakul individual, a breed developed in the central Asian region of present day Turkestan, revealed its haplotype (H33) clustered with the type B lineage (Figure 3). This groups it more closely with breeds from the Near East and distinct from breeds such as the Gizarr of geographic neighbors Kazakhstan and Tadjikistan (Hiendleder et al. 2002). An obvious omission from the sample set is representation of African and South American breeds. Future experiments are required to determine if a third clade is present within animals from these continents or if they are mixtures of the two clades A and B described here.

Acknowledgments

The authors thank Blair Harrison (CSIRO), Jill Maddox (University of Melbourne), Alexander Graphodatsky (Russian Academy of Sciences), Anthony Dufty (Rare Breeds Trust of Australia), and the Nimbkar Agricultural Research Institute, India, for contributions of ovine samples. Funding support

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024

for J. Kantanen and M. Tapio was received from the Academy of Finland, Ministry of Agriculture and Forestry and its program Sustainable Use of Natural Resources.

References

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

Arranz JJ, Bayon Y, and San Primitivo F, 2001. Differentiation among Spanish sheep breeds using microsatellites. Genet Sel Evol 33:529–542.

Bradley DG, MacHugh DE, Cunningham P, and Loftus RT, 1996. Mitochondrial diversity and the origins of African and European cattle. Proc Natl Acad Sci USA 93:5131–5135.

Bruford MW, Bradley DG, and Luikart G, 2003. DNA markers reveal the complexity of livestock domestication. Nat Rev Genet 4:900–910.

Calvo JH, Marcos S, Beattie AE, Gonzalez C, Jurado JJ, and Serrano M, 2004. Ovine alpha-amylase genes: isolation, linkage mapping and association analysis with milk traits in Manchega breed sheep. Anim Genet 35:239–332.

Chu MX, Wang JZ, Wang AG, Li N, and Fu JL, 2003. Association analysis between five microsatellite loci and litter size in small tail han sheep. Asian-Aust J Anim Sci 16:1555–1559.

Davis GH, Galloway SM, Ross IK, Gregan SM, Ward J, Nimbkar BV, Ghalsasi PM, Nimbkar C, Gray GD, Subandriyo, Inounu I, Tiesnamurti B, Martyniuk E, Eythorsdottir E, Mulsant P, Lecerf F, Hanrahan JP, Bradford GE, and Wilson T, 2002. DNA tests in prolific sheep from eight countries provide new evidence on origin of the Booroola (FecB) mutation. Biol Reprod 66:1869–1874.

Diez-Tascon C, Littlejohn RP, Almeida PA, and Crawford AM, 2000. Genetic variation within the Merino sheep breed: analysis of closely related populations using microsatellites. Anim Genet 31:243–251.

Excoffier L, Smouse PE, and Quattro JM, 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Forbes SH, Hogg JT, Buchanan FC, Crawford AM, and Allendorf FW, 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. Mol Biol Evol 12:1106–1113.

Giuffra E, Kijas JM, Amarger V, Carlborg O, Jeon JT, and Andersson L, 2000. The origin of the domestic pig: independent domestication and subsequent introgression. Genetics 154:1785–1791.

Hiendleder S, Kaupe B, Wassmuth R, and 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 Biol Sci 269:893–904.

Hiendleder S, Lewalski H, Wassmuth R, and Janke A, 1998a. The complete mitochondrial DNA sequence of the domestic sheep (*Ovis aries*) and comparison with the other major ovine haplotype. J Mol Evol 47: 441–448.

Hiendleder S, Mainz K, Plante Y, and Lewalski H, 1998b. 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.

Kumar S, Tamura K, Jakobsen IB, and Nei M, 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245.

Luikart G, Gielly L, Excoffier L, Vigne JD, Bouvet J, and Taberlet P, 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. Proc Natl Acad Sci USA 98:5927–5932.

MacHugh D and Bradley D, 2001. Livestock genetic origins: goats buck the trend. Proc Natl Acad Sci USA 98:5382–5384.

Meadows JRS, Hawken RJ, and Kijas JW, 2004. Nucleotide diversity on the ovine Y chromosome. Anim Genet 35:379–385.

Pesole G, Gissi C, Chirico AD, and Saccone C, 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. J Mol Evol 48:427–434.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, and Rozas R, 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497.

Ryder ML, 1984. Sheep. In: Evolution of domesticated animals (Mason IL, ed). New York: Longman; 63–84.

Sambraus HH, 1987. Atlas der Nutztierrassen. Stuttgart, Germany: Verlag Eugen Ulmer.

Sipos W, Kraus M, Schmoll F, Achmann R, and Baumgartner W, 2002. PrP genotyping of Austrian sheep breeds. J Vet Med A Physiol Pathol Clin Med 49:415–418.

Tapio M, Miceikiene I, Vilkki J, and Kantanen J, 2003. Comparison of microsatellite and blood protein diversity in sheep: inconsistencies in fragmented breeds. Mol Ecol 12:2045–2056.

Walling GA, Wilson AD, Mcteir BL, and Bishop SC, 2004. Increased heterozygosity and allele variants are seen in Texel compared to Suffolk sheep. Heredity 92:102–109.

Wood NJ and Phua SH, 1996. Variation in the control region sequence of the sheep mitochondrial genome. Anim Genet 27:25–33.

Received January 25, 2005
Accepted June 23, 2005

Corresponding Editor: Leif Andersson

Downloaded from https://academic.oup.com/jhered/article/96/5/494/2187700 by guest on 25 April 2024