The Eurasian Steppe is an important goat propagation route: A phylogeographic analysis using mitochondrial DNA and Y-chromosome sequences of Kazakhstani goats

Tabata, Risa ; Kawaguchi, Fuki ; Sasazaki, Shinji ; Yamamoto, Yoshio ; Bakhtin, Meirat ; Kazymbet, Polat ; Meldevekob, Alykhan ; Suleimenov,…

(Citation)
Animal Science Journal,90(3):317-322

(Issue Date)
2019-03

(Resource Type)
journal article

(Version)
Accepted Manuscript

The Eurasian Steppe is an important goat propagation route: a phylogeographic analysis using mitochondrial DNA and Y-chromosome sequences of Kazakhstani goats

Risa TABATA,[1] Fuki KAWAGUCHI,[1] Shinji SASAZAKI,[1] Yoshio YAMAMOTO,[2] Meirat BAKHTIN,[3] Polat KAZYMBET,[3] Alykhan MELDEVEKOB,[4] Maratbek Z. SULEIMENOV,[4] Masahide NISHIBORI,[2] and Hideyuki MANNEN[1,*]

[1]Laboratory of Animal Breeding and Genetics, Graduate School of Agricultural Science, Kobe University, Nada, Kobe 657-8501, Japan, [2]Graduate School of Biosphere Science, Hiroshima University, Higashi-Hiroshima 739-8528, Japan, [3]Astana Medical University, Radiobiology Scientific Center, Astana 010000, Kazakhstan, [4]Institute of Zoology, Almaty 050060, Kazakhstan

*Correspondence: Hideyuki Mannen, Graduate School of Agricultural Science, Kobe University, Nada, Kobe 657-8501, Japan. (E-mail: mannen@kobe-u.ac.jp)
TEL: +81-78-803-5803, FAX: +81-78-803-5803

Running title: Genetic diversity of Kazakhstan goat

Abstract

Goats (*Capra hircus*) were domesticated in the Fertile Crescent and propagated all over the world. The Silk Road through the Eurasian Steppe belt is a possible propagation route for domestic goats to Central Asia. Kazakhstan is in close geographical proximity to domestication centers and covers the majority of the Eurasian Steppe belt. In this study, we examined the genetic diversity and phylogeographic structure of Kazakhstani goats. The mtDNA sequences of 141 Kazakhstani goats were categorized into haplogroups A, C, and D, of which haplogroup A was predominant (97%), whereas haplogroups C and D were detected at low frequencies (1.4% each). The Kazakhstani haplotypes C were then categorized into Asian mtDNA type. Sequence analysis of the *SRY* gene on the Y-chromosome in 67 male Kazakhstani goats revealed two haplotypes: Y1A (64%) and Y2A (36%). Analysis of the distribution of mtDNA haplogroups and *SRY* haplotypes from Eurasia and Africa demonstrated genetic similarity among animals from Kazakhstan, Mongolia, and Northwest China located on the Eurasian Steppe belt. These phylogeographic results suggested that the Eurasian Steppe belt was an important propagation route for goats to Central Asia.

Key words

Capra hircus, Eurasian Steppe, propagation route, mtDNA D-loop, *SRY* gene

Introduction

Domestic goats (*Capra hircus*) have been raised for multiple purposes, including for their milk, meat, skin, and fiber. Archaeological evidence indicates that goat domestication occurred in the Fertile Crescent over 10,000 years ago (Zeder & Hesse, 2000; Luikart et al., 2006; Naderi et al., 2008). Because of their high tolerance of harsh environments, goats are one of the most important livestock in the world. Approximately 500–600 goat breeds have been reported worldwide (FAOSTAT, http://www.fap.org/faostat), and most of these are native breeds, which are well adapted to their environments but have not been sufficiently improved. Genetic information from these native breeds would shed light on the origins of domestication, their genetic relationships, environmental adaptation, and propagation history.

Previous molecular studies using mtDNA D-loop region were performed to investigate the origin and genetic relationship of goats (Luikart et al., 2001; Mannen, Nagata & Tsuji, 2001; Sultana, Mannen & Tsuji 2003; Joshi et al., 2004; Chen et al., 2005; Fernandez et al., 2006; Naderi et al., 2007; Lin et al., 2012). These studies revealed six highly divergent lineages of goat mtDNA (haplogroups A–G). Haplogroup A is distributed worldwide (Naderi et al., 2007), whereas most animals in haplogroup B are primarily found in South and Southeast Asia, with their frequency tending to increase southeastward (Lin et al., 2012). Haplogroup C is found in Asia and Europe with low frequencies, and the minor haplogroups

D, F, and G are observed in Asia, Sicily, and the Near East and northern Africa, respectively, at low frequencies (Naderi et al., 2007). These lineages are believed to have diverged prior to goat domestication (Luikart et al., 2001).

In addition, genetic variations on the Y-chromosome are useful for understanding paternal genetic diversity. Previous studies using *SRY* 3′UTR sequences (Canon et al*.,* 2006; Pereira et al., 2008; Pereira et al., 2009; Waki et al., 2015) identified four major *SRY* haplotypes worldwide (Y1A, Y1B, Y2A, and Y2B). The haplotypes Y1A and Y2A are found worldwide at high frequencies, whereas the haplotypes Y1B and Y2B are found specifically in Europe and Southeast Asia, respectively.

Goat domestication likely occurred at several sites in the Fertile Crescent (Zeder & Hesse, 2000; Luikart et al., 2006; Naderi et al*.,* 2008). From this center of domestication, many domesticated animals were propagated to Asia via the Silk Road, which was a series of ancient trade routes, including the Oasis Road, Sea Road, and Steppe Route (Eurasian Steppe belt). Nozawa (1991), Pereira & Amorim (2010), and Porter et al. (2016) also suggested that the Eurasian Steppe belt was a possible propagation route for goats, which are capable of adapting to the harsh environment of the Steppe. Kazakhstan and Mongolia occupy the majority of the Eurasian Steppe belt, and Kazakhstan in particular is in close geographical proximity to the goat domestication center and covers most of the Eurasian Steppe belt. Therefore, it is likely that Kazakhstani goats were

introduced to this region at an early stage of goat expansion, and their genetic information could provide important information toward understanding their propagation to Asia. Therefore, the objective of this study was to characterize the genetic diversity and the genetic relationships among countries and understanding the history of goat propagation in the Eurasian Steppe belt using mitochondrial and Y-chromosomal sequences of Kazakhstani goats.

Materials and Methods

Ethical Conditions

All procedures in the present study were performed according to the Research Guidelines for Kobe University.

Animals

We collected blood samples from 145 native goats from five regions in Kazakhstan: central (Astana city, n = 55), north (Aktobe and West Kazakhstan state, n = 31), southeast (Almaty state, n = 20), south (South Kazakhstan state, n = 20), and west (Mangystau state, n = 19). These geographic locations are shown in Figure 1. Genomic DNA was extracted from fresh blood using a standard phenol-chloroform method.

[Figure 1]

Sequencing

We sequenced the hypervariable segment I (HVI) of the mitochondrial D-loop from bp 15,707 to 16,187 (481bp) based on the goat mtDNA reference sequence (Accession No. AF533441.1) in 145 goats from Kazakhstan. Polymerase chain reaction was performed according to a previously described method (Sultana, Mannen & Tsuji, 2003). Standard double-strand DNA sequencing was performed using 20 ng of the amplified product using BigDye® Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, Tokyo, Japan), a primer with sequence 5′-TACCCACACAAACGCCAACACC-3′, and an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Tokyo, Japan).

We additionally sequenced the *SRY* 3′UTR region (478 bp) of 67 male Kazakhstani goats, covering nucleotide position bp 2,568 to 3,045 based on the goat *SRY* reference sequence (Accession No. D82963). This region was amplified and sequenced according to a previously published method (Waki et al., 2015).

Data analysis

Sequence alignment of HVI was performed using MEGA 7.0.14 (Kumar et al., 2016). To investigate the genetic relationships between mtDNA sequences, an unrooted neighbor-joining phylogenetic tree (Saitou & Nei, 1987) was constructed using the Tamura–Nei distance method (Tamura & Nei, 1993), including 22

reference haplotypes from six mtDNA haplogroups in goats selected by Naderi et al. (2007). The distance computation and phylogenetic tree construction were also conducted in MEGA 7.0.14, as was the estimation of the average number of nucleotide differences in each geographic region. Median-joining was used to construct a network (Bandelt, Forster & Röhl, 1999) in NETWORK 5.0.0.1, in which all mutations were equally weighted. To investigate the differences within/among populations, analysis of molecular variance (AMOVA) was conducted using ARLEQUIN 3.5.2.2 (Excoffier & Lischer, 2010).

To analyze the distribution of the mtDNA haplogroups in Eurasia and Africa, we used mtDNA data of 4,000 individuals from the DDBJ database (Table S1). In addition, to investigate regional differentiation within haplogroup A, we included 20 European, 20 Middle-Eastern, and 20 Eurasian Steppe goat sequences (Table S2). We also included sequences of 68 individuals with haplogroup C and 32 individuals with haplogroup D (Table S2). Furthermore, we analyzed the distribution of paternal haplotypes in goat populations in Eurasia and Africa using previously published data of 1,019 West and East Eurasian goats (Table S3).

Results

Maternal genetic structure of Kazakhstani goats

In total, we determined the mtDNA sequences of the 481 bp HV1 region from 141

Kazakhstani goats. The sequencing analysis identified 133 variable sites, including 109 transitions, 18 transversions, 5 transition/transversion events at the same locus, and 1 deletion. Based on these variations, 87 mtDNA haplotypes were identified. All the sequences were deposited in the DDBJ database (accession Nos. LC416627–LC416713). The neighbor-joining tree displayed the three highly divergent mtDNA lineages, A (n = 137), C (n = 2), and D (n = 2) in Kazakhstani goats (Fig. S1). The predominant haplogroup A was observed in all five regions, whereas haplogroup C was observed in the central (n = 1) and southeast (n = 1) regions; haplogroup D was present only in the southeast region (Figure 1). The median-joining network using haplogroup A sequences indicated that almost all the haplotypes were unique to each region and only three haplotypes were shared among two regions (Fig. S2).

The average number of pairwise differences was calculated for all individuals and separately for individuals with haplogroup A. When considering all the individuals, the values ranged from 0.0186 in the south to 0.0278 in the southeast region. When considering only haplogroup A, the values ranged from 0.0162 in the southeast to 0.0213 in the central region (Table S4). Genetic distances using haplogroup A ranged from 0.019 to 0.024 (Table S5), and no obvious differences were observed among the Kazakhstani regions. AMOVA revealed that approximately 97.33% and 97.22% ($p < 0.0000$) of the mtDNA diversity was distributed within the population (Table S6) when using all

haplogroups and only haplogroup A, respectively.

Geographic distribution of maternal haplogroups

Figure 2 shows the mtDNA haplotype distribution in Eurasia and Africa and indicates that haplogroup A was predominant in most countries. To investigate the diversity within haplogroup A, we constructed a tree using the neighbor-joining method for haplogroup A in Europe, West Asia, and Central Asia (Fig. S3). This tree illustrated the presence of highly divergent haplotypes but yielded no clear geographic differentiation.

[Figure 2]

Haplogroup C was found in Southern Europe, Central Asia, and Southern Asia at low frequencies. The neighbor-joining tree of 70 individuals with haplogroup C could clearly distinguish European and Asian haplotypes, and two Kazakhstani goats were categorized into the Asian cluster (Figure 3). Although haplogroup D was found in West and Central Asia at low frequencies, the phylogenetic tree for haplogroup D, constructed using 34 individuals, did not reveal a clear geographic structure (Fig. S4).

[Figure 3]

Paternal genetic structure and geographic distribution of Kazakhstani

goats

SRY sequences of Kazakhstani goats (n = 67) revealed two haplotypes, including the predominant haplotype Y1A (64.2%, n = 43) and haplotype Y2A (35.8%, n = 24). The Y1A haplotype was found in all five regions: central (n = 23/28), North (n = 8/25), southeast (n = 6/7), west (n = 5/5), and south (n = 1/2). The sequences obtained in this study were deposited in the DDBJ database (accession Nos. LC416985–LC416986).

Figure 4 illustrates the geographic distribution of goat *SRY* haplotypes in Eurasia and Africa. The haplotype Y1A was found in almost all regions. The haplotype Y2A was found in Europe, northeastern Asia, and northern Africa. Although the two major haplotypes Y1A and Y2A, were distributed worldwide, haplotypes Y1B and Y2B were found in limited areas of Europe and Southeast Asia, respectively.

[Figure 4]

Discussion

Maternal relationships in Kazakhstani goats

The mtDNA analysis of Kazakhstani goats identified three haplogroups: A, C, and D. A comparison of the genetic structure of the mtDNA haplogroups in Eurasia and Africa demonstrated that goat mtDNA in the Eurasian Steppe belt (Kazakhstan, Northwest China, and Mongolia) comprised predominantly haplogroup A and less haplogroups C and D (Figure 2).

Haplogroup A was predominant in Kazakhstani goats (n = 137; 97.2%), as well as in almost all the regions of the world. As many previous studies suggested, haplogroup A contains a high degree of genetic diversity, and there are small genetic differences among sampling regions (Fig. S3). Luikart et al. (2001) proposed that this weak phylogeographic structure likely resulted from frequent transportation for commercial trade or migratory and exploratory movements by humans. This weak phylogeographic structure was also supported by the domestication of multiple divergent wild goats (Daly et al., 2018).

Haplogroup C was also found in Kazakhstan (n = 2; 1.4%) and other European and Asian countries (Figure 2). The phylogenetic tree of haplogroup C revealed two highly divergent clusters of European and Asian types, as reported by Hughes et al. (2012) (Figure 3). Fernandez et al. (2006) suggested that gene introgression from wild goats and/or multiple domestication processes involving goats with haplogroup C occurred over time and in various geographic centers.

Therefore, the haplotypic data for haplogroup C should yield molecular markers capable of distinguishing European and Asian goat populations and identifying propagation routes. In this study, the two Kazakhstani haplotypes C were categorized into the Asian type, suggesting differentiation of the genetic roots of European and Asian goats and the derivation of goats in Kazakhstan from the Asian Steppe route.

Phylogenetic construction of haplogroup D (Fig. S4) revealed no clear differentiation among geographic regions, as well as haplogroup A. Therefore, it is also likely that introgression of haplogroup D into domestic goat populations occurred during the primary stage of goat domestication and propagation. A similar domestication and propagation scenario for haplogroup A would account for the lack of geographic differentiation between their haplotypes.

These data indicate that goat mtDNA in the Eurasian Steppe belt (Kazakhstan, Northwestern China, and Mongolia) has a similar phylogenetic structure in terms of both haplogroup distribution and haplotype clustering.

Paternal relationships in Kazakhstani goats

Kazakhstani goats contain two paternal haplotypes: Y1A and Y2A. These haplotypes are observed worldwide, whereas haplotypes Y1B and Y2B are specifically found in Europe and Southeastern Asia, respectively (Figure 4). Two hypotheses can be proposed for the origin of the phylogeographic haplotypes

Y1B and Y2B. First, they are novel haplotypes generated by a mutation during propagation. Second, they are the result of additional Y-chromosome introgression from wild male goats. Pereira et al. (2008) reported that the divergence times of Y1A and Y1B greatly predate the nearly 10,000 years of domestication, suggesting that Y1B is a result of introgression from wild goats into domestic populations. In addition, with respect to haplotype Y2B, the latter hypothesis is more likely because Southeast Asian goats specifically have maternal haplogroup B with high frequencies, which had a different origin (Luikart et al., 2001; Chen et al., 2005). Therefore, goats with Y2B and mtDNA haplogroup B may be derived from the same wild ancestor, as suggested previously (Waki et al*.,* 2015).

Because haplotypes Y1A and Y2A were detected in most parts of the world, these two haplotypes were probably the primary domesticated lineages and were transported throughout the world. However, the frequency of Y1A and Y2A tends to increase eastward and westward, respectively, in Eurasia. This differing distribution of the haplotypes may be influenced by founder and/or bottleneck effects on each propagation route. These haplotypes are found dominantly in Kazakhstan and Mongolia, which are located on the Eurasian Steppe belt. Thus, Y-chromosome sequences in the Eurasian Steppe belt reveal similar genetic characteristics based on haplotype structure and frequency, suggesting that goats in the Eurasian Steppe belt were propagated using the

same route.

In conclusion, analysis of mtDNA and Y-chromosome sequences revealed a similarity of genetic background among goats from the Eurasian Steppe belt. These phylogeographic results suggest that the Eurasian Steppe belt was probably an important goat propagation route to Central Asia. This contention is supported by the results of previous studies (Nozawa, 1991; Pereira & Amorim, 2010; Porter et al., 2016) and implies that goat transportation *via* the Eurasian Steppe was facilitated by this well adaptation of this species to harsh environments. This genetic information sheds light on the origin and propagation history of Central and East Asian domestic goats.

Acknowledgment

This work was supported in part by JSPS KAKENHI Grant Numbers 25304038 and 17H04643.

Reference

Bandelt, H.J., Forster, P. and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution, 16(1), 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Canon, J., Garcıa, D., Garcıa-Atance, M.A., Obexer-Ruff, G., Lenstra, J.A., Ajmone-Marsan, P., ⋯ The ECONOGENE Consortium (2006). Geographical partitioning of goat diversity in Europe and the Middle East. Animal Genetics, 37, 327-324. https://doi.org/10.1111/j.1365-2052.2006.01461.x

Chen S.Y., Su Y.H., Wu S.F, Sha T. and Zhang Y.P. (2005). Mitochondrial diversity and phylogeographic structure of Chinese domestic goats. Molecular Phylogenetics and Evolution, 37, 804 - 814. https://doi.org/10.1016/j.ympev.2005.06.014

Daly, K.G., Delser, P.M., Mullin, V.E., Scheu, A., Mattiangeli, V., Teasdale, M.D., ⋯ Bradley D.G. (2018). Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. Science, 361, 85–88. https://doi.org/10.1126/science.aas9411

Excoffier, L. and Lischer, H.E. (2010). Arlequin suite ver 3.5: a new series of

programs to perform population genetics analyses under Linux and Windows. Molecular Biology and Evolution, 10, 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fernandez, H., Hughes, S., Vigne, J.D., Helmer, D., Hodgins, G., Miquel, C., ⋯ Taberlet, P. (2006). Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas. PNAS, 103(42), 15375-15379. https://doi.org/10.1073/pnas.0602753103

Hughes, S., Fernandez, H., Cucchi, T., Duffraisse, M., Casabianca, F., Istria, D., ⋯ Taberlet P. (2012). A Dig into the Past Mitochondrial Diversity of Corsican Goats Reveals the Influence of Secular Herding Practices. PLoS ONE, 7(1), e30272. https://doi.org/10.1371/journal.pone.0030272

Joshi, M.B., Rout, P.K., Mandal, A.K., Smith, C.T., Singh, L. and Thangaraj, K. (2004). Phylogeography and Origin of Indian Domestic Goats. Molecular Biology and Evolution, 21, 454-462. https://doi.org/10.1093/molbev/msh038

Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution, 33(7), 1870-1874. https://doi.org/10.1093/molbev/msw054

Lin, B.Z., Odahara, S., Isida, M., Kato, T., Sansazaki, S., Nozawa, K. and Mannen, H. (2012). Molecular phylogeography and genetic diversity of East Asian goats. Animal Genetics 44, 79-85. https://doi.org/10.1111/j.1365-2052.2012.02358.x

Luikart, G., Fernandez, H., Mashkour, M., Englang, P.R. and Taberlet, P. (2006). Origins and Diffusion of Domestic Goats Inferred from DNA Markers Example Analyses of mtDNA, Y Chromosome, and Microsatellites. In Zeder & Decker-Walters & Bradley & Smith, *Documenting Domestication: New Genetic and Archaeological Paradigms* (pp. 294–306). University of California Press, Oakland, CA, USA.

Luikart, G., Gielly, L., Excoffier, L., Vigne, J.D., Bouvet, J. and Taberlet, P. (2001). Multiple maternal origins and weak phylogeographic structure in domestic goats. Proceeding of the National Academy of Science of United States of America, 98, 5927-5932. https://doi.org/10.1073/pnas.091591198

Mannen, H., Nagata, Y. and Tsuji, S. (2001). Mitochondrial DNA reveal that domestic goat (Capra hircus) are genetically affected by two subspecies of bezoar (Capra aegagurus). Biochemical Genetics, 39(5-6), 145-154. https://doi.org/10.1023/A:1010266207735

Naderi, S., Rezaei, H.R., Pompanon, F., Blum, M.G.B., Negrini, R., Naghash, H.R., ⋯ Taberlet, P. (2008). The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. PNAS, 105(46), 17659-17664. https://doi.org/10.1073/pnas.0804782105

Naderi, S., Rezaei, H.R., Taberrlet, P., Zundel, S., Rafat, S.A., Naghash, H.R., ⋯ Econogene Consortium (2007). Large-Scale Mitochondrial DNA Analysis of the Domestic Goat Reveals Six Haplogroups with High Diversity. PLoS ONE, 2(10), e1012. https://doi.org/10.1371/journal.pone.0001012

Nozawa, K. (1991). Domestication and history of goats. In K. Maijala, *Genetic Resources of Pig, Sheep and Goat* (pp. 391–403). Elsevier, Amsterdam, The Netherlands.

Pereira, F. and Amorim, A. (2010). Origin and spread of goat pastoralism. Encyclopedia of Life Sciences. Retrieved from https://doi.org/10.1002/9780470015902.a0022864.

Pereira, F., Carneiro, J., Soares, P., Nejmeddine, F., Lenstra, J.A ., Gusmao, L. and Amotim, A. (2008). A multiplex primer extension assay for the rapid

identification of paternal lineages in domestic goat (Capra hircus): Laying the foundations for a detailed caprine Y chromosome phylogeny. Molecular Phylogenetics and Evolution, 49, 663-668. https://doi.org/10.1016/j.ympev.2008.08.026

Pereira, F., Queiros, S., Gusmao, L., Nijman, I.J., Cuppen, E., Lenstra, J.A., ⋯ Amorim, A. (2009). Tracing the History of Goat Pastoralism: New Clues from Mitochondrial and Y Chromosome DNA in North Africa. Molecular Biology and Evolution, 26(12), 2765-2773. https://doi.org/10.1093/molbev/msp200

Porter, V., Alderson, L., Hall, S.J.G. and Sponenberg, D.P. (2016). Goats. In *Mason's World Encyclopedia of Livestock Breeds and Breeding* (volume 1) (pp. 350–2). CABI, Oxfordshire, UK.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, 4, 406-425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Sultana, S., Mannen, H. and Tsuji, S. (2003). Mitochondrial DNA diversity of Pakistani goats. Animal Genetics, 34, 417–421. https://doi.org/10.1046/j.0268-9146.2003.01040.x

Tamura, K. and Nei, M. (1993). Estimation of the nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution, 10, 512-526. https://doi.org/10.1093/oxfordjournals.molbev.a040023

Waki, A., Sasazaki, S., Kobayashi, E. and Mannen, H. (2015). Paternal phylogeography and genetic diversity of East Asian goats. Animal Genetics, 46, 337-339. https://doi.org/10.1111/age.12293

Zeder, M.A. and Hesse, B. (2000). The Initial Domestication of Goats (*Capra hircus*) in the Zagros Mountainlins 10,000 Years Ago. Science, 287(5461), 2254-2257. https://doi.org/10.1126/science.287.5461.2254

FIGURE LEGENDS

FIGURE 1 Sampling regions, sample size (N) and geographical distribution of mtDNA haplogroups in Kazakhstani goats. We sequenced 141 Kazakhstani goats from five regions; 55 Central (Astana city), 27 North (Altobe and West Kazakhstan state), 20 Southeast (Almaty state), 20 South (South Kazakhstan state), 19 West (Mangystau state).

FIGURE 2 Geographic distribution of caprine mtDNA haplogroups in the world. The size of each circle is proportional to the sample size and each specific haplogroup is represented by a different color. Data information used in this figure is summarized in Table S1.

FIGURE 3 Neighbor-joining tree of mtDNA haplogroup C from European, Near East and Asian goats. Scale bar indicates genetic distance.

FIGURE 4 Geographic distribution of caprine SRY haplotypes in the world. The size of each circle is proportional to the sample size and each specific haplotype is represented by a different color. Data information used in this figure is summarized in Table S3.

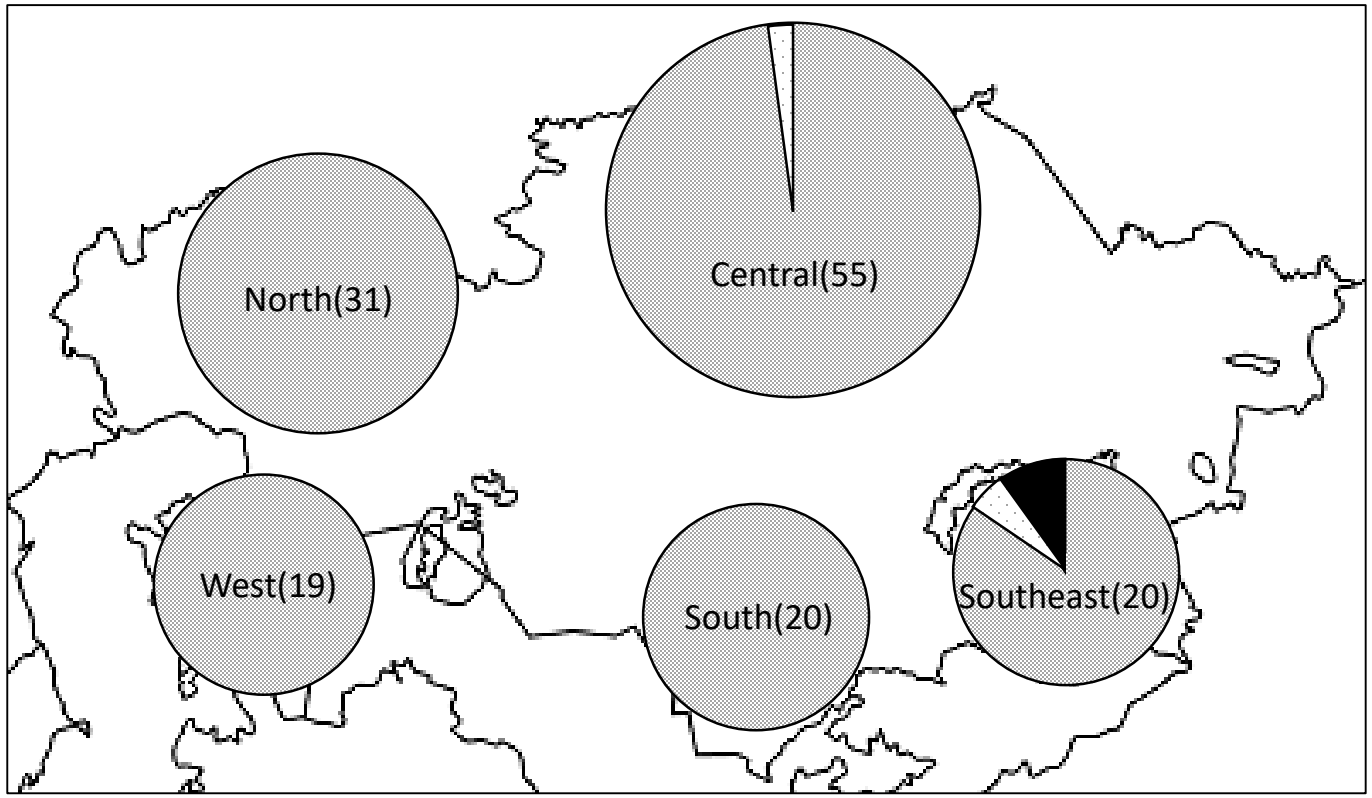

Figure 1.

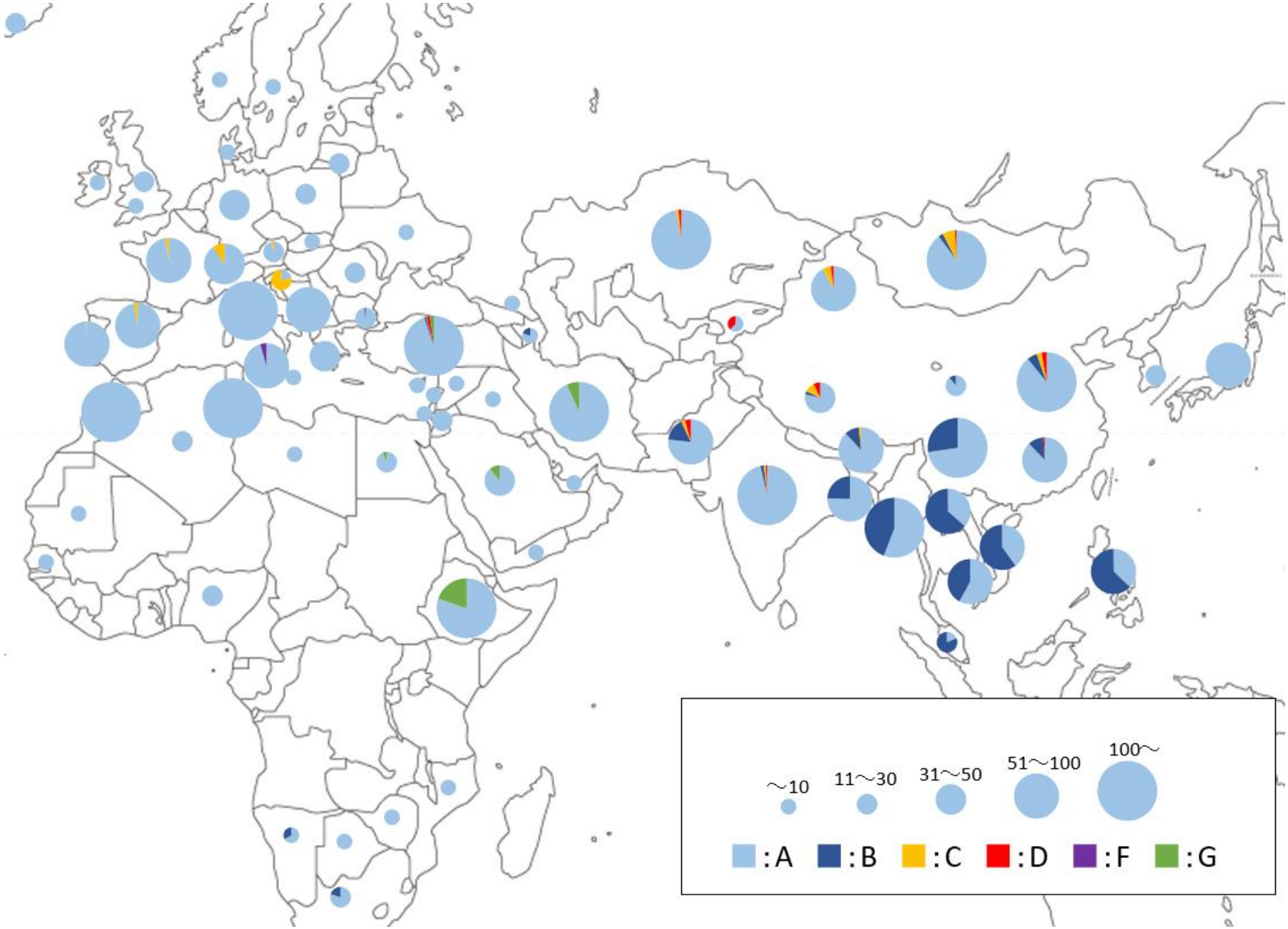

Figure 2

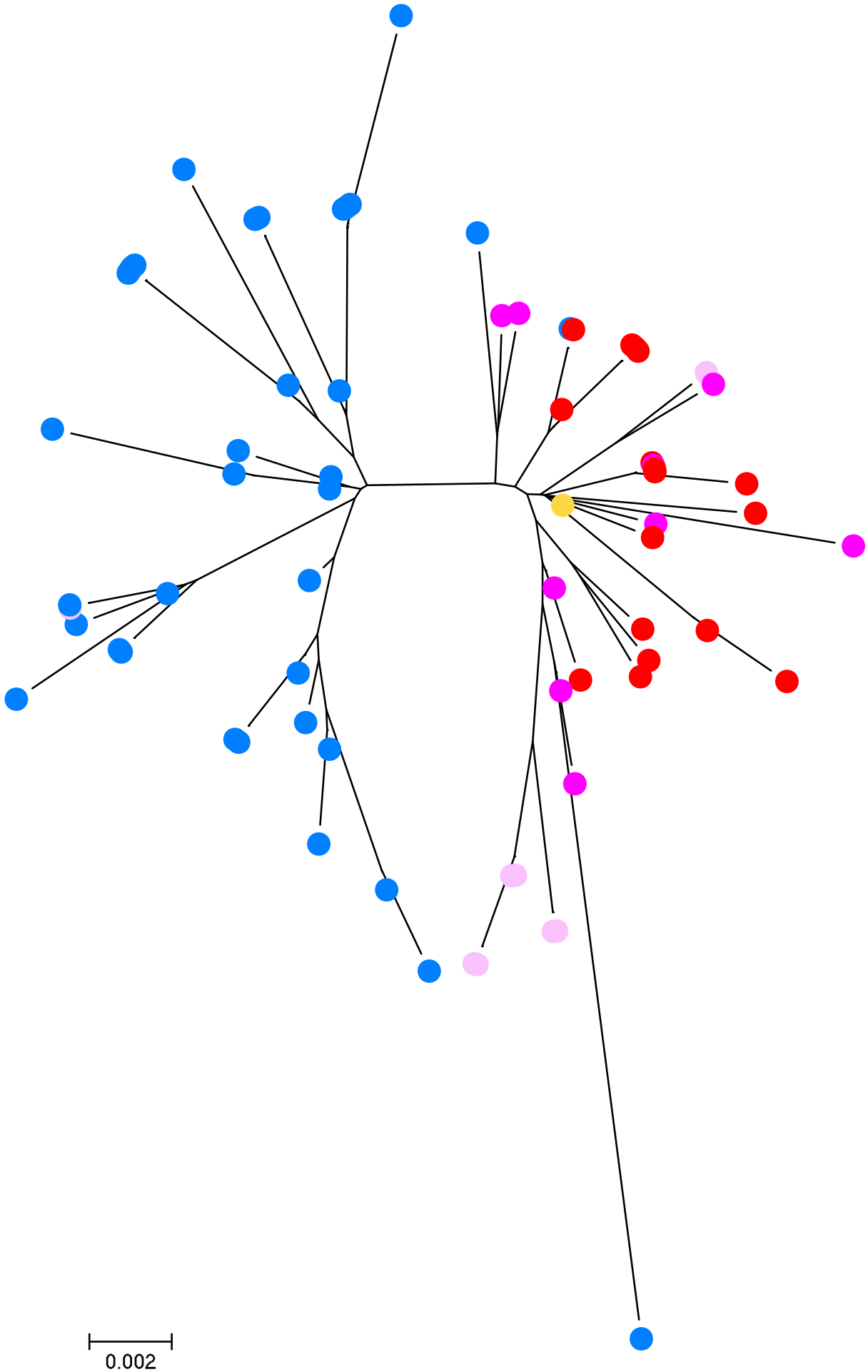

Figure 3.

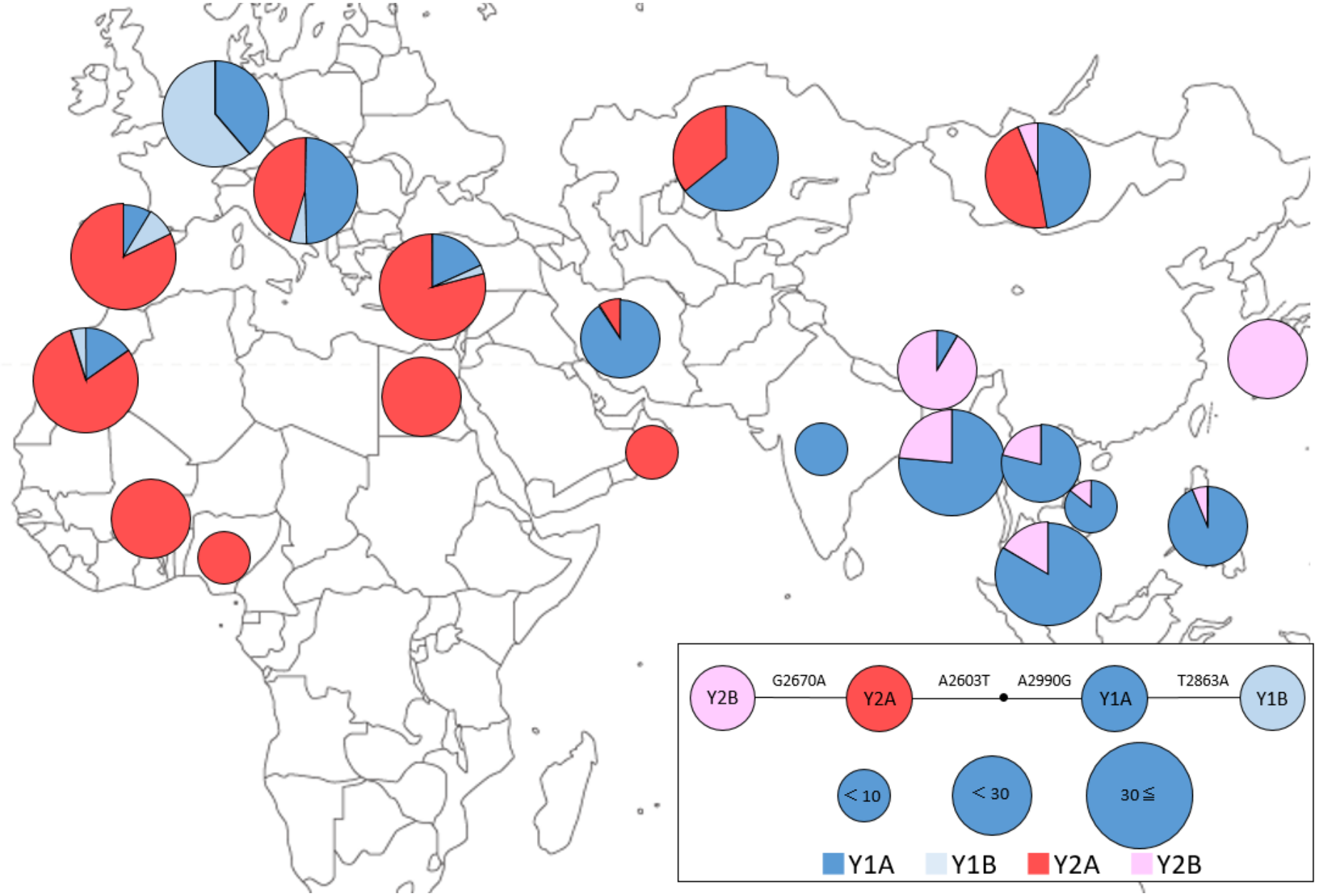
Y2B
G2670A
Y2A
A2603T
A2990G
Y1A
T2863A
Y1B
＜10
＜30
30≦
Y1A
Y1B
Y2A
Y2B

Figure 4.

Table 1. Average mean number of pairwise differencesof Kazakhstani five regions.

Regions	n	Average mean number of pairwise differences	
		Whole population	Only haplogroup A
Central	55	0.0236	0.0213
North	13	0.0197	0.0197
Southeast	20	0.0278	0.0162
South	20	0.0186	0.0186
West	19	0.0199	0.0199
All	127	0.0231	0.0203

Table 2. Genetic distance of Kazakhstani five regions (Tamura&Nei model). This distances were calcu

	Central	North	Southeast	South
North	0.023			
Southeast	0.021	0.019		
South	0.023	0.021	0.019	
West	0.024	0.021	0.019	0.021

ulated using haplogroup A animals.

Table 3. Hierarchical distribution of mtDNA (HV1) diversity for Kazakhstani goat population as computed

	Whole population	Only haplogroup A
among population	2.67	2.78
within population	97.33	97.22

| under AMOVA framework. (p< 0.00000)

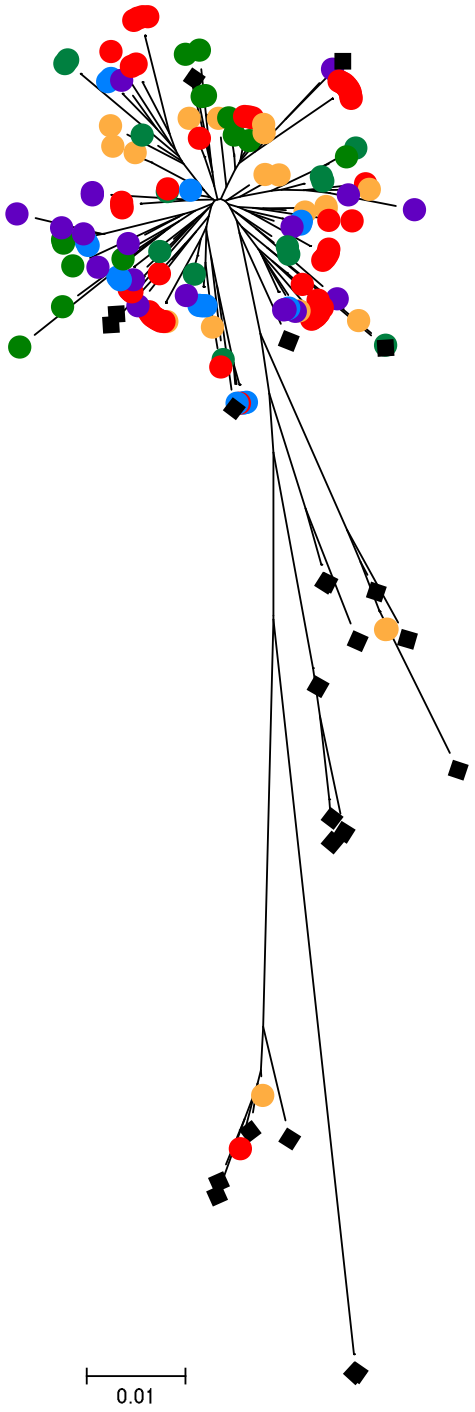

Figure S1.

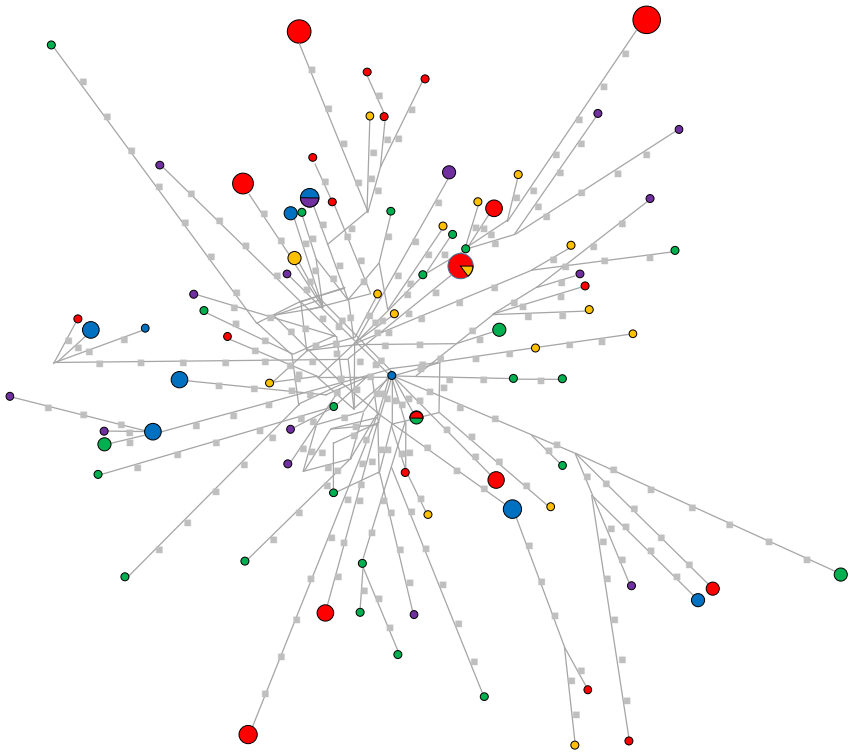

Figure S2

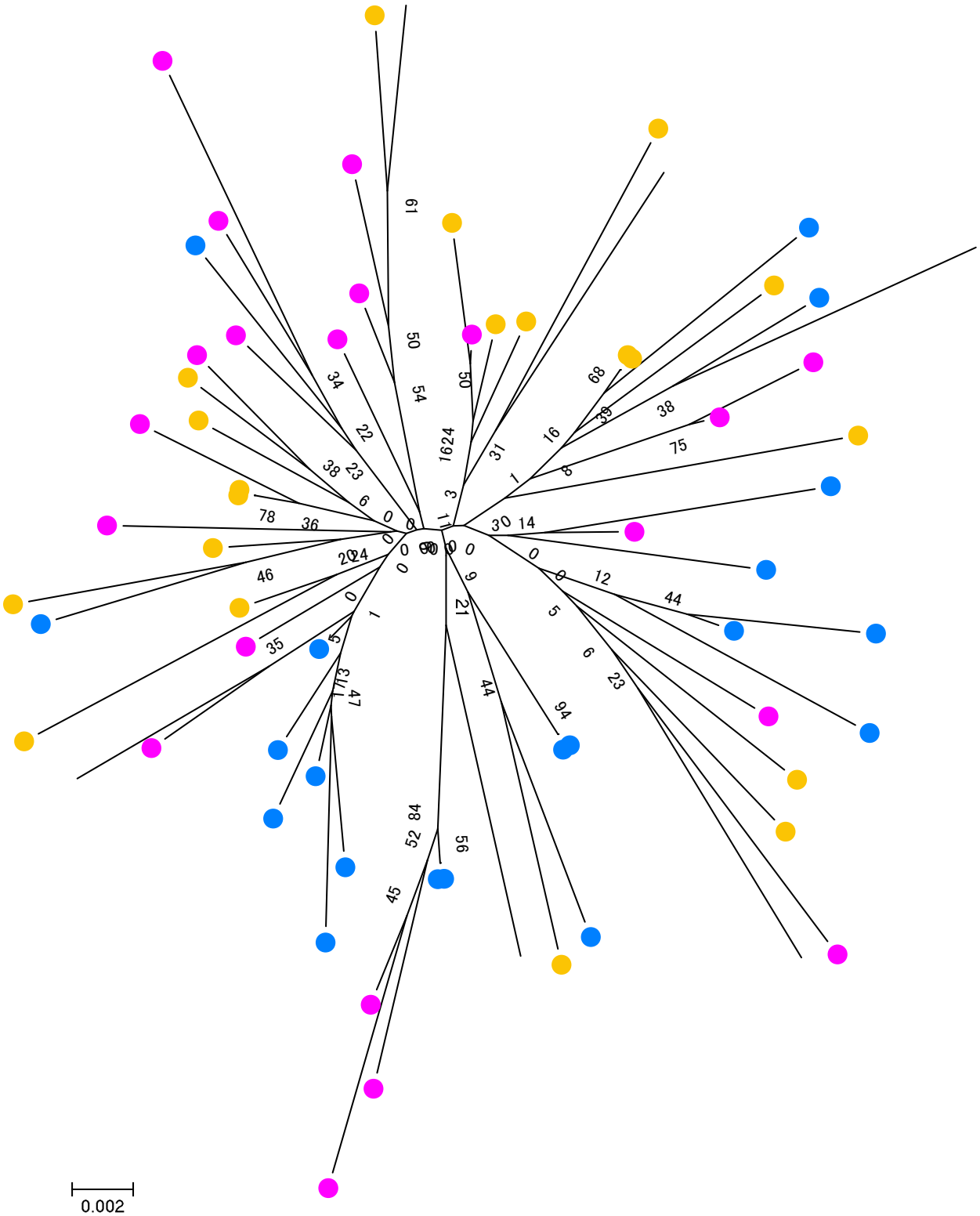

Figure S3.

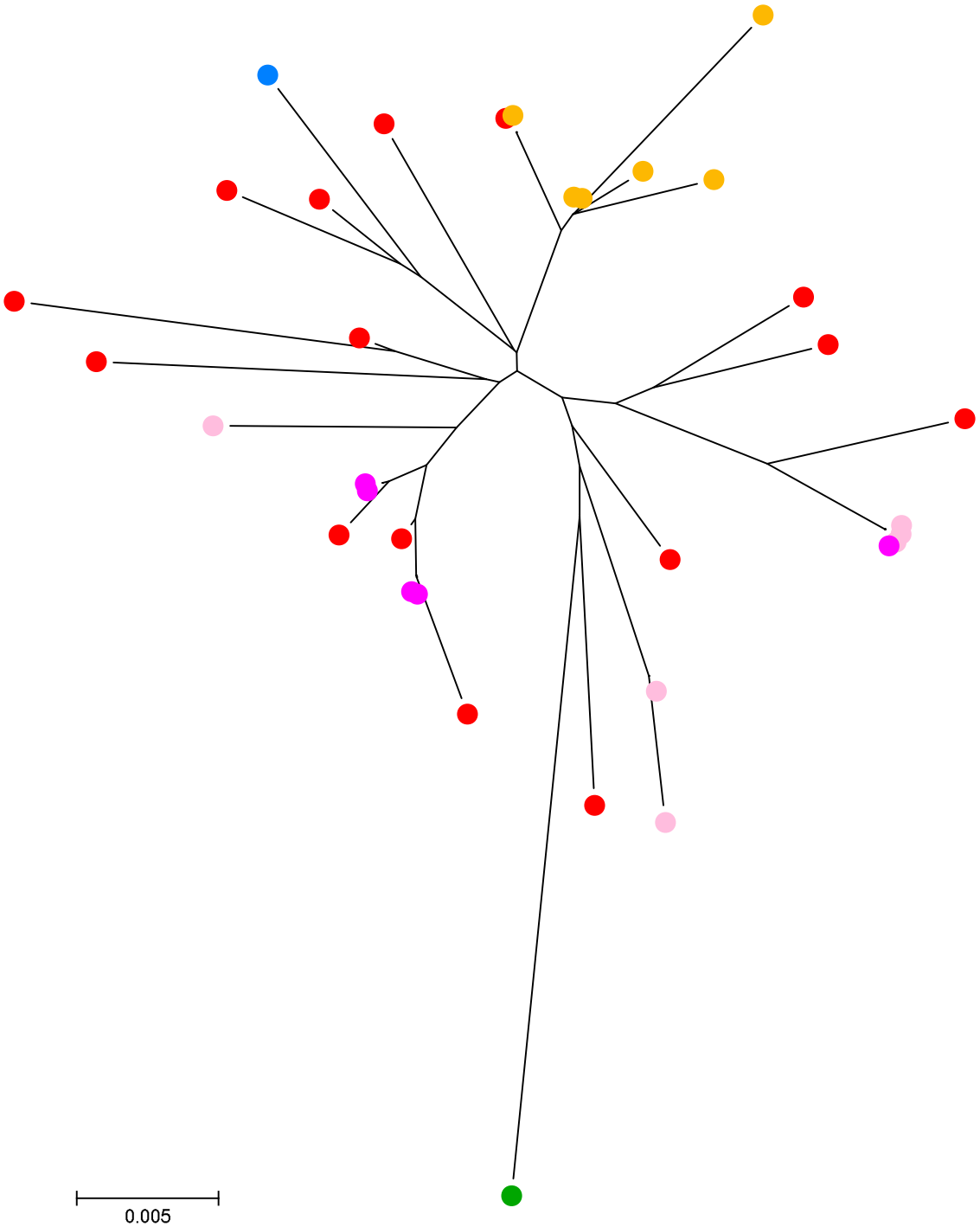

Figure S4.

Table S1. List of goat mtDNA haplogroups information from Genbank used in this study.

Region	Country	N	A	B	C	D	F
Middle East	Iran	222	207				
	Iraq	7	7				
	Jordan	19	19				
	Lebanon	8	8				
	Palestine	9	9				
	Pakistan	73	56	12	2	3	
	Saudi Arabia	40	35				
	Syria	9	9				
	Turkey	361	342	4	1	6	
	United Arab Emirates	7	7				
	Yemen	9	9				
Eastern Asia	China	667	532	118	10	7	
	Japan	56	56				
	Korea	19	19				
Central Asia	Kyrgyzstan	8	5			3	
	Mongolia	117	105	3	8	1	
Southeastern Asia	Bhutan	67	59	7	1		
	Cambodia	133	77	56			
	Laos	85	31	54			
	Malaysia	27	5	22			
	Myanmar	181	101	80			
	Philippine	30	8	22			
	Vietnam	65	26	39			
Western Asia	Azerbaijan	5	4	1			
	Dagestan	2	2				
South Asia	India	387	373	7	4	3	
	Bangladesh	53	40	13			
Northern Africa	Algeria	14	14				
	Egypt	29	27				
	Libya	7	7				
	Mauritania	8	8				
	Morocco	156	156				
	Nigeria	12	12				
	Senegal	3	3				
	Tunisia	22	22				
Sub Saharan Africa	Botswana	4	4				
	Mozambique	8	8				
	Namibia	4	2	2			
	South Africa	12	3				

	Zimbabwe	4	4				
Northern Europe	Austria	24	23			1	
	Denmark	2	2				
	England	3	3				
	France	79	77		2		
	Germany	32	32				
	Iceland	11	11				
	Ireland	8	8				
	Norway	3	3				
	Poland	27	27				
	Sweden	9	9				
	Switzerland	104	94		10		
	Slovakia	2	2				
	Slovenia	8	2		6		
	Wales	7	7				
Southern Europe	Albania	77	77				
	Cyprus	4	4				
	Greece	47	46	1			
	Italy	115	115				
	Malta	4	4				
	Portugal	321	320		1		
	Romania	26	26				
	Sicily	67	64				3
	Spain	73	71		2		
	Ukraine	6	6				
total		4000	3559	337	47	24	3

G	Reference
15	Naderi *et al.* (2007)
	Luikart *et al.* (2001)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Al-Araimi *et al.* (2017)
	Al-Araimi *et al.* (2017)
	Luikart *et al.* (2001); Sultana *et al.* (2003); Naderi *et al.* (2007)
5	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Al-Araimi *et al.* (2017)
8	Luikart *et al.* (2001); Naderi *et al.* (2007); Akis *et al.* (2014)
	Al-Araimi *et al.* (2017)
	Al-Araimi *et al.* (2017)
	Chen *et al.* (2005); Liu *et al.* (2006); Larkin *et al.* (2007); Li *et al.* (2002); Zhao *et al.* (2014)
	Lin *et al.* (2013)
	Odahara *et al.* (2006)
	Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007); Lin *et al.* (2013)
	Luikart *et al.* (2001); Lin *et al.* (2013)
	Lin *et al.* (2013)
	Mannen *et al.* (2001); Lin *et al.* (2013)
	Luikart *et al.* (2001); Naderi *et al.* (2007); Lin *et al.* (2013)
	Lin *et al.* (2013)
	Kato *et al.* (2013)
	Luikart *et al.* (2001); Naderi *et al.* (2007); Lin *et al.* (2013)
	Naderi *et al.* (2007)
	Naderi *et al.* (2007)
	Luikart *et al.* (2001); Joshi *et al.* (2004); Naderi *et al.* (2007)
	Kato *et al.* (2013)
	Luikart *et al.* (2001); Amills *et al.* (2004); Al-Araimi *et al.* (2017)
2	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Naderi *et al.* (2007); Al-Araimi *et al.* (2017)
	Al-Araimi *et al.* (2017)
	Luikart *et al.* (2001); Naderi *et al.* (2007); Benjelloun *et al.* (2011); Al-Araimi *et al.* (2017)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001)
	Luikart *et al.* (2001); Vacca *et al.* (2010); Al-Araimi *et al.* (2017)
	Luikart *et al.* (2001)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)

	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Pereira *et al.* (2005); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Sardina *et al.* (2006)
	Luikart *et al.* (2001); Naderi *et al.* (2007)
	Luikart *et al.* (2001); Naderi *et al.* (2007)

Table S2. List of goat mtDNA haplogroup A sequence used in this st

Region	country	N
Europe	Austria, Switzerland, Poland, Germany, France, Sweden, Slovenia, Wales, Ireland, Ukraine, Slovakia, Cyprus, Portugal, Greece, Italy, Albania, Romania, Sicily, Spain	20
Middle East	Iran	10
	Turkey	10
Eurasian Steppe belt	Kazakhstan	10
	Mongolia	5
	Northwestern China	5

udy.

Accession number
EF617678; EF618423; EF618264; EF617788; EF617730; EF618415; EF618346; EF618542; EF618085; EF917851; EF618540 (Naderi *et al.* 2007); AJ317653 (Luikart *et al.* 2001); EF617707; EF618287; EF617816; EF618087; EF617601; EF618296(Naderi *et al.* 2007); DQ241305 (Sardina *et al.* 2006); EF617357 (Naderi *et al.* 2007)
EF617863; EF617873; EF617883; EF617893; EF617903; EF617913; EF617923; EF617933; EF617943; EF617953 (Naderi *et al.* 2007)
EF618492; EF618497; EF618502; EF618507; EF618512; EF618517; EF618522; EF618527; EF618532; EF618533 (Naderi *et al.* 2007)
sampling No.1; 71 (Central); 28; 91 (North) ; 31; 41 (Southeast); 51; 61 (South); 121; 131 (West)
EF618234; EF618238; EF618236 (Naderi *et al.* 2007); AJ317534; AJ317545 (Luikart *et al.* 2001)
DQ089434; DQ089444; DQ089454 (Chen *et al.* 2005); EF103509; EF103518(Wang *et al.* 2008)

Tsble S3 List of goat mtDNA haplogroup C s

Region	country	N
Europe	France	2
	Sardinia	7
	Slovenia	6
	Spain	9
	Switzerland	12
Middle East	Turkey	1
Central Asia	Mongolia	7
Eastern Asia	China	16
Southeastern Asia	Bhutan	1
South Asia	India	4
	Pakistan	2

sequences used in this study.

Accession number
EF617786; EF617787 (Naderi *et al.* 2007)
JN085564; JN085593; JN085604; JN085668; JN085676; JN085695; JN085770 (Piras *et al.* 2012)
AJ317835; AJ317837 (Luikart *et al.* 2001); EF618347; EF618348; EF618349; EF618350 (Naderi *et al.* 2007)
AY424915 (unpublished data of Amills *et al.* 2003); AY918060 (Azor 2005); EF618413; EF618414 (Naderi *et al.* 2007); EU910305 (Royo *et al.* 2009); HQ713401; HQ713415 (Martinez *et al.* 2012); KM893320 (Ferrando *et al.* 2015); KR059225 (Colli *et al.* 2015)
AJ317836; AJ317838; AJ317839; AJ317840 (Luikart *et al.* 2001); EF618486; EF618487; EF618488; EF618489; EF618490; EF618491 (Naderi *et al.* 2007); KR059223; KR059224 (Colli *et al.* 2015)
KC574160 (Akis *et al.* 2014)
AB440769; AB440770; AB440771; AB440772; AB440773; AB440774 (Lin *et al.* 2013); AJ317834 (Luikart *et al.* 2001)
DQ089187; DQ089188; DQ089191; DQ089192; DQ089460; DQ089461 (Chen *et al.* 2005); DQ188886; DQ188890; DQ188892 (Liu *et al.* 2005); EF103513 (Wang *et al.* 2008); EF368288; EF368314 (Chen *et al.* 2007); EU035991; EU035998; EU036007 (Wi *et al.* 2009); EU130770 (Liu *et al.* 2009); KC190465 (Zhong *et al.* 2012)
AB440749 (Lin *et al.* 2013)
AY155708; AY155877; AY156000; AY156009 (Joshi *et al.* 2004)
AB110555; AB110559 (Sultana *et al.* 2003)

Table S4 List of goat mtDNA haplogroup D sequen

Region	Country	N
Europe	Austria	1
Africa	Nigeria	1
Middle East	Turkey	6
Central Asia	Kyrgyzstan	3
	Mongolia	1
East Asia	China	14
South Asia	India	3
	Pakistan	3

ces used in this study.

Accession number
EF617701 (Naderi *et al.* 2007)
KU292693 (Ajibike *et al.* 2016)
KC574241; KC574296; KC574323; KC574324; KC574343; KC574345 (Akis *et al.* 2014)
EU618217; EF618218; EF618219 (Naderi *et al.* 2007)
AB440766 (Lin *et al.* 2013)
DQ089350 (Chen *et al.* 2005); DQ188888; DQ1888893 (Liu *et al.* 2005); EF103496; EF103516; EF103529; EF103537 (Wang *et al.* 2008); EU036025; EU036029; EU036032; EU036038; EU036055; EU036061 (Wu *et al.* 2007); EU130714 (Liu *et al.* 2009)
AY155952; AY155964; AY155965 (Joshi *et al.* 2004)
AB110587; AB110588; AB110589 (Sultana *et al.* 2003)

Table S5. List of goat SRY haplotypes information from Genbank used in this study.

Region	country	N	Y1A	Y1B	Y2A	Y2B
Southeast Asia	Bhutan	24	2			22
	Cambodia	36	30			6
	Laos	14	11			3
	Myanmar	34	26			8
	Philippine	16	15			1
	Vietnam	7	6			1
South Asia	India	7	7			
East Asia	Japan	11				11
Central Asia	Mongolia	32	15		15	2
Near East	Iran	22	20		2	
	Oman	1			1	
Eastern Mediterranean	-	196	36	4	156	
Central Mediterranean	-	137	58	26	53	
Western Mediterranean	-	164	14	15	135	
Central and Northern Europe	-	219	85	134		
Africa	Morocco	66	10	53	3	
	Burkina Faso	14			14	
	Nigeria	4			4	
	Egypt	15			15	
total		1019	335	232	398	54

reference
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Waki *et al.* (2015)
Vidal *et al.* (2017)
Vidal *et al.* (2017)
Canon *et al.* (2006); Kul *et al.* (2015)
Canon *et al.* (2006); Vidal *et al.* (2017)
Canon *et al.* (2006); Pereira *et al.* (2009); Vidal *et al.* (2017)
Canon *et al.* (2006); Vidal *et al.* (2017)
Pereira *et al*. (2009)
Vidal *et al.* (2017)
Vidal *et al.* (2017)
Vidal *et al.* (2017)