Haplotype and Genetic Diversity of mtDNA in Indigenous Iranian Sheep and an Insight into the History of Sheep Domestication

M. H. Moradi¹*, S. H. Phua², N. Hedayat³, M. Khodaei-Motlagh¹, and M. Razmkabir⁴

ABSTRACT

The archaeological evidence suggests that Iran has been one of the first origins of sheep domestication in the world. This study aims to investigate the genetic diversity of indigenous Iranian sheep breeds using mitochondrial DNA (mtDNA) and to explore the evolutionary history of sheep domestication in Iran. Single Nucleotide Polymorphism (SNP) markers in the control region of mtDNA were used to genotype the unrelated sheep samples of Zel and Lori-Bakhtiari breeds which were collected from or near the center of the sheep domestication, using the Sequenom MassARRAY platform. Phylogenetic analysis of the mitochondrial SNPs classified all animals into either of two haplogroups A and B. The population differentiation (Fst) and gene flow (Nm) statistics were 0.054 and 4.715 respectively, indicating a low mitochondrial genetic differentiation and high gene flow between two sheep breeds. The Analysis of Variation (AMOVA) showed that around 97% of the total genetic diversity is distributed within the two breeds. Further analysis using SNP haplotyping identified nine different haplotypes within the animals; eight haplotypes were present in the Zel, while only four were seen in the Lori-Bakhtiari breed. Two haplotypes, designated H1 and H3, were present at higher frequencies in both breeds. Haplotypes H5, H6, H7, H8 and H9 were found as population-specific in the Zel, and haplotype H10 only occurred in Lori-Bakhtiari breed. The existence of two common Haplotypes (H1 and H3) in the animals suggest that the two Iranian breeds from strikingly different geographical regions, may share a common ancestry, and these haplotypes could be the origin haplotypes while the population specific haplotypes developed later.

Keywords: Genetic diversity, Haplogroups, Mitochondrial genome, Phylogenetic study, Sheep evolution.

INTRODUCTION

Domestication of livestock represents a crucial step in human history. The rise of civilisations could not happen without the domestication of plants and animals. Sheep (Ovis aries) is the first grazing animal known to be domesticated (Chessa et al., 2009). The archaeozoological evidence suggests that the domestication of sheep occurred during the Neolithic revolution approximately 9000 years ago (Kijas et al., 2009), in a region in northern Iraq and nearby regions close to Zagros...
mountains in Iran (Zeder, 1999). This means that the indigenous Iranian sheep breeds were among the first to be domesticated in the world. The wide geographical distribution patterns of retrovirus insertion site polymorphisms suggested that there were two major waves of sheep migrations: firstly, the spread of early domestic sheep presumably from the center of domestication, followed by the second wave of sheep (with improved traits) migration from Southwest Asia around 5,000 years BP (Chessa et al., 2009). Therefore, it is of essential interest to study the major native sheep breeds of Iran, Lori-Bakhtiari and Zel, using mitochondrial DNA (mtDNA) approach in an attempt to understand the history of domestication and migrations.

The genetic history of sheep has been investigated based on three major sources of genomic variations, namely the autosomes, Y chromosome and mitochondrial genome (Agaviezor et al., 2012). Presently, mtDNA has received much attention because its markers provide important phylogenetic information in relation to genetic diversities. The rate of mtDNA mutations is about 5 to 10 times faster than nuclear DNA, and its genes do not usually recombine (Upholt et al., 1977). Mammalian mtDNA is deemed to strictly follow maternal inheritance and is highly variable within a species, so mtDNA provides an important resource for phylogenetic studies. Based on mtDNA sequence, four distinct haplogroups (namely A, B, C and D) have been found in the domestic sheep, indicating four maternal lineages. The main haplogroups A and B are both found in Asia, while B dominates in Europe (Guo et al., 2005; Cinkulov et al., 2008). Haplogroup C has been found in Portugal, Turkey, the Caucasus and China (Tapio et al., 2006; Chen et al., 2006; Pardeshi, 2007; Zhao et al., 2011). Haplogroup D, present in Rumanian Karachai and Caucasian animals, is possibly related to the haplogroup A (Pariset et al., 2011). In contrast to cattle, the sheep haplogroups hardly correlate with geography.

Sheep production is the most important component of the Iranian livestock industries, with a population of 50 million animals which are well adapted to the harsh arid of sub-tropical areas (Mokhtari et al., 2015). Two main Iranian indigenous sheep breeds were investigated in the present study. The Lori-Bakhtiari sheep is one of the most common breeds, with a population of more than 1.7 million animals. It is well adapted to the hilly and mountainous Bakhtiari region which stretches out to the southern Zagros Mountains (Figure 1). The animals are kept mostly in villages under semi-intensive systems. Relative to other Iranian fat-tail breeds, Lori-Bakhtiari is a large animal, having the largest fat-tail by girth and weight (Vatankhah and Talebi, 2008). Zel is the only thin-tail Iranian breed; it is present largely on the northern slopes of the Elburz mountain range near the Caspian Sea (Moradi et al., 2012), and constitutes around 3% of the Iranian sheep population. This breed is also known as the Aryan breed since the historical evidence shows that the Aryans, who were living in these areas, attempted to domesticate these animals. It is a general belief that the first domesticated sheep were thin-tail and the fat-tail animals were developed later (Moradi et al., 2012). Hence, Zel breed is also known as a contour of wild and domesticated strains in Iran.

Very little work has been done on the genetic background and phylogenetic history of Iranian sheep breeds using mtDNA. The aim of the present study was to examine and describe the genetic diversity and phylogeny of indigenous Lori-Bakhtiari and Zel sheep breeds in Iran. With the animal samples collected from or near the center of domestication, we hope to shed some light on the
domestication event and on the history of Iranian sheep.

MATERIALS AND METHODS

Animal Samples

Blood samples of 150 Zel (134 females and 16 males) and 107 Lori-Bakhtiari (89 females and 18 males) animals were collected from farms distributed across the traditional rearing areas of each breed. Blood were collected in EDTA vacutainers by jugular venepuncture. The animals were from groups which were in the registration and recording system of the Animal Breeding Centre of Iran. Two criteria were used when selecting the experimental animals, namely the animals must be unrelated and from diverse geographical regions. While sampling pedigreed animals we ensured that they had no common antecedents, and in non-pedigreed animals we typically selected 4-5 animals of different ages from each group. The Zel breed animals were collected from the northern province of Mazandaran located in the south of the Caspian Sea. The Lori-Bakhtiari animals were from the Chaharmahal and Bakhtiari Province, which are located in the south western part of Iran close to the Zagros Mountain ranges.
SNP-Multiplex Design, DNA Extraction and Genotyping

In the 1,180-bp sequence of the mtDNA control region, there are 39 SNP markers which differentiate between mtDNA haplotypes A and B (Wood and Phua, 1996). These markers were used in the construction of a SNP-multiplex utilizing the MassARRAY Assay Design 4.0 software (Sequenom Inc., San Diego, CA, USA). The mtDNA sequence reported in Wood and Phua (1996) was used for designing the primers. We obtained a multiplex consisting of ten SNP markers (Table 1).

Total DNA, consisting of genomic and mitochondrial DNA, was extracted from whole blood using a modified salting out protocol (Moradi et al., 2012). DNA samples were diluted to 50 ng µl⁻¹ concentration. Genotyping with the SNP-multiplex was performed following the Sequenom’s recommended protocol, based on the iPLEX™ genotyping assay (Oeth et al., 2005). Reactions were carried out in a 384-well plate, using 25 ng of DNA per animal. The reaction products were transferred onto chips and analysed in a MassARRAY Compact 96 mass spectrometer.

Data Analysis

Nucleotide variations in the ten SNP sites of the 257 Lori-Bakhtiari and Zel animals were analysed and aligned using MEGA v.5.1 software, with a Kimura 2-parameter (with transitions and transversions) model and a bootstrap (1,000 replications) test (Tamura et al., 2011). DnaSP 5.1 software (Rozas et al., 2009) was used to calculate haplotype, indices of nucleotide variation and haplotype structure including nucleotide diversity (π), number of haplotypes (nh), Haplotype diversity (Hd) and average number of nucleotide substitutions (Dxy) per site between breeds. To evaluate the fit of major models of nucleotide substitutions based on obtained haplotypes, MEGA v.5.1 software
was applied and phylogenetic relationships among mtDNA haplotypes were obtained by Maximum Likelihood (ML) as implanted in this software. An Analysis of Molecular Variance (AMOVA) was computed to test significant differences in mtDNA diversity between the breeds using ARLEQUIN 3.01 (Excoffier et al., 2005). The Median Joining (MJ) networks were plotted using the Network 4.6.1.2 program (http://www.fluxus-engineering.com) to reveal the possible relationships between haplotypes.

RESULTS AND DISCUSSION

In this study, 24 widely-used models were investigated for phylogenetic analyses. Models with least Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC) and Maximum Likelihood (ML) information indices were used for description of the best substitution pattern. In the end, Kimura 2 (K2) model was selected as the best model for phylogenetic studies.

From the 257 analysed animals, nine haplotypes were found in Zel and Lori-Bakhtiari breed samples which were designated H1 to H9 (Table 2). In contrast to the Lori-Bakhtiari breed which has only four haplotypes (H1, H2, H3 and H4), Zel breed has higher mtDNA diversity with eight haplotypes (all haplotypes except H2). Although, most of the selection pressures do not act on the mtDNA, unless the selection occurs on energy metabolism, it is an observation that the low diversity of haplotypes in Lori-Bakhtiari breed may be due to the higher selection of pressure Imposed on this breed (Moradi et al., 2012). The higher haplotypes number seen in Zel breed is also likely to be due to the higher number of animals analysed: 150 Zel versus 107 Lori-Bakhtiari. Furthermore, the number of haplotypes present in the females was greater than that in the males (Table 3). Though most of the breeding programs in sheep are focused on rams, the latters do not pass on their mtDNA to the progeny. This again is possibly due to the higher number of female animals analysed (223 females versus 34 males).

The network analysis of haplotypes shows big differences between low-frequency haplotypes in Zel and Lori-Bakhtiari mtDNA diversity and history of sheep domestication  

Table 2. The two mtDNA haplogroups A and B, and the nine haplotypes present in the 257 Zel and Lori-Bakhtiari sheep. Haplogroups A and B are represented by H1 and H3 haplotypes, respectively. The other haplotypes are variants of the two ancestral haplogroups.

<table>
<thead>
<tr>
<th>Haplogroup</th>
<th>Haplotype</th>
<th>Breed</th>
<th>SNP variants&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>H1</td>
<td>Lori-Bakhtiari and Zel</td>
<td>TCCGAGGGTTG</td>
</tr>
<tr>
<td>A</td>
<td>H2</td>
<td>Lori-Bakhtiari</td>
<td>TCCGAGGGTA</td>
</tr>
<tr>
<td>B</td>
<td>H3</td>
<td>Lori-Bakhtiari and Zel</td>
<td>CTTAGAACCA</td>
</tr>
<tr>
<td>B</td>
<td>H4</td>
<td>Lori-Bakhtiari and Zel</td>
<td>CTTAGAACC</td>
</tr>
<tr>
<td>B</td>
<td>H5</td>
<td>Zel</td>
<td>CTCAGAACCA</td>
</tr>
<tr>
<td>A</td>
<td>H6</td>
<td>Zel</td>
<td>TCCGAGGTTG</td>
</tr>
<tr>
<td>A</td>
<td>H7</td>
<td>Zel</td>
<td>CTCAGGATT</td>
</tr>
<tr>
<td>B</td>
<td>H8</td>
<td>Zel</td>
<td>TTTAGAACCA</td>
</tr>
<tr>
<td>A</td>
<td>H9</td>
<td>Zel</td>
<td>TCTTAGGTTG</td>
</tr>
</tbody>
</table>

<sup>a</sup> Haplotype with the highest frequency is haplotype H1 which is used here as the reference sequence. The nucleotide variants which are different to the reference bases are bold.
Table 3. Population statistics and haplotype diversity in Lori-Bakhtiari and Zel breeds based on mtDNA SNP markers.

<table>
<thead>
<tr>
<th>Population</th>
<th>No. haplotypes</th>
<th>Haplotype diversity</th>
<th>No. different nucleotides (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>9</td>
<td>0.454</td>
<td>7.473</td>
</tr>
<tr>
<td>SubLori1</td>
<td>4</td>
<td>0.428</td>
<td>7.011</td>
</tr>
<tr>
<td>SubLori2</td>
<td>2</td>
<td>0.303</td>
<td>6.061</td>
</tr>
<tr>
<td>SubLori3</td>
<td>2</td>
<td>0.476</td>
<td>9.542</td>
</tr>
<tr>
<td>SubZel1</td>
<td>5</td>
<td>0.708</td>
<td>10.401</td>
</tr>
<tr>
<td>SubZel2</td>
<td>4</td>
<td>0.692</td>
<td>9.077</td>
</tr>
<tr>
<td>SubZel3</td>
<td>2</td>
<td>0.533</td>
<td>10.660</td>
</tr>
<tr>
<td>SubZel4</td>
<td>2</td>
<td>0.556</td>
<td>11.112</td>
</tr>
<tr>
<td>SubZel5</td>
<td>4</td>
<td>0.600</td>
<td>10.152</td>
</tr>
<tr>
<td>SubZel6</td>
<td>3</td>
<td>0.522</td>
<td>8.206</td>
</tr>
<tr>
<td>SubZel7</td>
<td>2</td>
<td>0.209</td>
<td>4.183</td>
</tr>
<tr>
<td>SubZel8</td>
<td>2</td>
<td>0.143</td>
<td>2.857</td>
</tr>
<tr>
<td>SubZel9</td>
<td>2</td>
<td>0.286</td>
<td>5.714</td>
</tr>
<tr>
<td>Lori</td>
<td>4</td>
<td>0.421</td>
<td>6.904</td>
</tr>
<tr>
<td>Zel</td>
<td>8</td>
<td>0.478</td>
<td>7.898</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>0.452</td>
<td>7.512</td>
</tr>
<tr>
<td>Male</td>
<td>4</td>
<td>0.480</td>
<td>7.471</td>
</tr>
</tbody>
</table>

Figure 2. Network analysis of observed haplotypes in Zel and Lori-Bakhtiari sheep. Black is Lori-Bakhtiari breed and yellow is Zel breed.

population (Figure 2). The existence of two common haplogroups (i.e. haplotypes H1 and H3) in the indigenous Iranian sheep breeds, which came from strikingly different geographical regions, suggests that these haplogroups would constitute the early haplotypes of the two breeds, while the population specific haplotypes evolved later. Since most of the genes on the mitochondrial genome are associated with energy metabolism, and are related with adaptation to different climatic conditions (Upholt et al., 1977), it seems feasible that the population specific haplotypes can be developed later due to the evolution and adaptation to different environments. As shown in Figure 2, there are four more potential haplotypes (designated mv1 to
mtDNA diversity and history of sheep domestication

...mv4); these theoretical potential haplotypes were not observed in the studied samples, but they may exist in a wider population.

Based on the statistical results, the Lori-Bakhtiari and Zel populations could be grouped into three and nine subpopulations, respectively (Table 3). The haplotype diversity and the number of different nucleotides for all subpopulations, breeds and sexes have been shown in Table 3, however, these statistics for all animal samples were 0.45 and 7.47, respectively. Mohammadhashemi et al. (2012) analysed ten mitochondrial regions in Moghani sheep breed and reported a haplotype diversity of 0.3, whereas Wang et al. (2006) in a study on Helan Mountain sheep population in China reported that the haplotype diversity was 0.79. The results obtained in our present study are within the reported ranges. Oner et al. (2013) reported that the average number of different nucleotides was 7.456 in nine Turkish native breeds, which is in agreement with the results of Iranian indigenous breeds. However, the haplotype diversity in Iranian breeds is low in comparison with sheep in China (Zhao et al., 2011), Asia, India, Europe (Pardeshi et al., 2007) and Africa (Agaviezor et al., 2012). This difference could be due to breed differences or the sampling method used in the studies.

Estimation of the genetic differentiation Fixation index ($F_{ST}$) and gene flow in all ten SNP sites between Zel and Lori-Bakhtiari breeds were 0.052±0.003 and 4.715±0.249, respectively. Lori-Bakhtiari is one of the most important Iranian fat-tailed breeds which have adapted to mountainous and valley regions in the southern part of Iran, along the Zagros Mountains. Zel is the only thin-tailed breed that is present largely on the northern part of Iran, near the Caspian Sea (Moradi et al., 2012). The findings of low $F_{ST}$ and high gene flow indicate that the different geographical areas have no barrier on the genetic structure of the populations. Different studies have been carried out to examine population differentiation and gene flow among different breeds of sheep all over the world. Agaviezor et al. (2012), in a study of genetic diversity in Nigerian sheep breeds using mtDNA analysis, reported a low $F_{ST}$ (between 0.0013- 0.0033) which is consistent with our results. Agaviezor et al. (2013), reported a presence of gene flow within four Nigerian sheep breeds. Similar findings were obtained by Moradi et al. (2012) based on SNP markers in the nuclear DNA. They reported low $F_{ST}$ with a mean of 0.024 (SD= 0.035) using the same Zel and Lori-Bakhtiari animal samples. Kijas et al. (2009) studied 23 domestic and two wild sheep breeds and showed that sheep in general exhibit low differentiation, consistent with the short evolutionary history of the species. The results of our study using mitochondrial genome are consistent with the reported findings.

Analysis of Molecular Variance (AMOVA) within and between populations was conducted for all ten mtDNA SNPs. The results revealed that for one marker, about 2% of the genetic variation is found between breeds, compared to 98% variation within breeds. For the remaining nine SNP markers, the variation within and between breeds were 97 and 3%, respectively. AMOVA revealed that mtDNA diversity is mainly distributed within breeds, mainly due to the maternal inheritance of the organelles. This is in agreement with the result of low $F_{ST}$ and high level of gene flow between populations. As it was shown earlier, both breeds are formed from two main haplogroups (haplotypes H1 and H3) with a relatively high frequency and these results correspond to low genetic diversity among breeds. Meadows et al. (2005) studied the mtDNA diversity in different Asian and European sheep breeds, and reported that sheep has the largest gene flow and weakest population structure in domestic animals, so that only 2.7 percent of the total genetic diversity is intercontinental diversity, whereas it is 10 and 50 percent for goat and cow, respectively.

Phylogenetic analysis of the Iranian animals based on mtDNA haplotypes is
shown in Figure 3. The phylogenetic tree was obtained using the maximum likelihood and Bootstrap method as the best model. The numbers on the phylogenetic tree are Bootstrap percentages in the formation of branches. The phylogenetic analysis classified all animals into two mtDNA haplogroups A and B. The two haplogroups were found in both breeds. Different variants of haplogroup A gave rise to haplotypes H1, H2, H6, H7 and H9; similarly, variants of haplogroup B generate H3, H4, H5 and H8 haplotypes. H1 and H3 haplotypes have the highest frequencies in both breeds, indicating that they are the ancestral types. Haplogroup A in terms of number of haplotypes and haplogroup B in terms of number of individuals has more frequency.

At least four mtDNA haplogroups have been reported in the literature (Wood and Phua, 1996; Tapio et al., 2006; Pariset et al., 2011). Wood and Phua (1996), in a study on the control region of mtDNA of 50 animals from five New Zealand sheep breeds, reported that all samples can be categorized into either of two haplogroups, A or B. Hiendleder et al. (1999) using RFLP markers analysed 239 samples from five European, three Central Asian and two African sheep breeds, and reported the same results. Lasagna et al. (2013) studied genetic diversity among three Italian breeds based on mtDNA, and detected two haplogroups A and B; there was one haplotype in haplogroup A and 82 haplotypes in haplogroup B. Mariotti et al. (2013) analysed 138 sheep samples belonging to seven Italian breeds; based on 68 SNPs in the mtDNA control region, they detected three haplogroups in their animals. In a similar study on 406 unrelated animals of 48 sheep breeds from European, Caucasian and Central Asian regions, Tapio et al. (2006) identified a total of four haplogroups, indicating four maternal lineages in the modern sheep breeds. Based on the different haplogroups reported in sheep (Wang et al., 2007; Zhao et al., 2011), it is likely that domestic sheep are formed from at least four different lineages of wild type Mouflon (Tapio et al., 2006). The presence of two mtDNA haplogroups in Iranian sheep breeds, may suggest that Iranian sheep originated from two sources of Mouflon, or from a large population containing both derivative haplogroups. However, further researches and complete mitochondrial sequencing are needed to establish the origin of the breeds.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Animal Breeding Center of Iran (ABCI) for access to the animals of the indigenous
sheep breeds. We also wish to thank the Animal Genomics Group at AgResearch Invermay for providing the facilities to conduct the experimental work. Thanks to the staff from Tehran University and AgResearch, especially Dr. A. Nejati-Javaremi, Dr. M. Moradi-Shahrbabak and Ms. Rayna Anderson who supported and helped with this research.

REFERENCES


تنوع های پلوتیوی و زننیکی ذنوم میتوکندریایی در گوسفندان بومی ایران و کاوش تاریخچه اهلی سازی گوسفند

م. ح. مرادی، س. م. فوازن. هدایتی، م. خدادی-مطلق و. م. رزم-کیار

چکیده

یافته‌های باستان‌شناسی نشان می‌دهد که ایران یکی از مهد‌های اصلی اهلی‌سازی گوسفند در جهان محسوب می‌شود. هدف از اجرای این تحقیق بررسی تنوع زننیکی نژاد‌های گوسفند بومی ایران با استفاده از اطلاعات ذنوم میتوکندریایی و کاوش تاریخچه تکاملی در این نژاد‌ها بود. به این منظور DNA نشان‌گرهای SNP در ناحیه کنترل mtDNA میتوکندریایی در نمونه‌های غیرخيراشته درندها از مادرین و لری-بختیاری (که در داخل یا خارجی مراکز اهلی‌سازی گوسفند قرار دارند) با استفاده از تیم‌های زننیکی مشاهده شدند. آنالیزهای فیلوزنیکی نشان داد که جمعیت‌های گوسفند بومی در دو هایپدرگوپ و A و B دارای می‌باشد. آمارهایهای G54 و F57 به ترتیب 54 تا 0/47/15 به عنوان دهنده میزان زننیکی پايين و جریان زنی بیابین این دو نژاد است. نتایج آنالیز واریانس تنوع زننیکی (AMOVA) نیز نشان داد که حداکثر 97% از کل تنوع زننیکی مشاهده شده در داخل نژاد‌ها پراکنده شده است. نتایج حاصل از آنالیزهای هایپولوپی مجموعاً به‌عث شناسایی 9 هایپلود مختالف در بین 4 جمعیت‌های گوسفند ایرانی شد که 8 مورد از آنها در نژاد‌های 2 و 1 درای شناسایی شدند. بدنه 4 هایپلود در نژاد‌های اهلی‌سازی مشاهده شدند. دو هایپلود H1 و H2 هایپلود محتلف در بین 4 جمعیت‌های گوسفند ایرانی یافت‌شدند. در این تحقیق هایپلود H2 مختص نژاد لری-بختیاری و هایپلود H1 و H3 و H4، H5، H6 و H7 در نژاد‌های اهلی‌سازی مشاهده شدند. دو هایپلود H1 و H2 در نژاد‌های اهلی‌سازی مشاهده شدند. (که در دو منطقه زننیکی کامل مشابه قرار دارند) دارای جد مشترکی بوده و این هایپلود های می‌توانند به عنوان هایپلود های اصلی متأثر از نژاد‌های اهلی‌سازی سطح بالا، در حالیکه هایپلودهای اصلی اختصاصی بعضاً در جهت انتقال و یا شرایط محیطی به وجود آمدند.