Genetic distances and phylogenetic trees of different Awassi sheep populations based on DNA sequencing

R.M. Al-Atiyat[1,2] and R.S. Aljumaah[1]

[1]Animal Production Department, King Saud University, Riyadh,
Kingdom of Saudi Arabia
[2]On Sabbatical Leave from Animal Production Department,
Mutah University, Jordan

Corresponding author: R.M. Al-Atiyat
E-mail: ralatiyat@ksu.edu.sa / raedatiyat@gmail.com

ABSTRACT. This study aimed to estimate evolutionary distances and to reconstruct phylogeny trees between different Awassi sheep populations. Thirty-two sheep individuals from three different geographical areas of Jordan and the Kingdom of Saudi Arabia (KSA) were randomly sampled. DNA was extracted from the tissue samples and sequenced using the T7 promoter universal primer. Different phylogenetic trees were reconstructed from 0.64-kb DNA sequences using the MEGA software with the best general time reverse distance model. Three methods of distance estimation were then used. The maximum composite likelihood test was considered for reconstructing maximum likelihood, neighbor-joining and UPGMA trees. The maximum likelihood tree indicated three major clusters separated by cytosine (C) and thymine (T). The greatest distance was shown between the South sheep and North sheep. On the other hand, the KSA sheep as an outgroup showed shorter evolutionary distance to the North sheep population than to the others. The neighbor-joining and UPGMA

trees showed quite reliable clusters of evolutionary differentiation of Jordan sheep populations from the Saudi population. The overall results support geographical information and ecological types of the sheep populations studied. Summing up, the resulting phylogeny trees may contribute to the limited information about the genetic relatedness and phylogeny of Awassi sheep in nearby Arab countries.

Key words: Phylogeny; Genetic Distance; Awassi sheep; DNA sequences

INTRODUCTION

The Awassi sheep is a dominant fat-tailed breed in most Arab countries (Galal et al., 2008). These sheep have great adaptability to tropical environmental conditions and are often used as a triple purpose sheep, better for milk production. The world's sheep were domesticated in the Mediterranean area in the Stone Age (Mason, 1984). At present, sheep breeds of Arab countries are expected to be most closely related to their ancestors of the Mediterranean area (Piper and Ruvinsky, 1997). However, there is not enough data published on Awassi sheep ancestry, phylogeny, biodiversity and/or its relatedness in Arab countries.

The use of advanced molecular DNA sequence technology has provided wide opportunities to analyze genetic relatedness at the DNA level in animal breeds. On one hand, DNA sequencing markers are widely used for functional genetics of specific sequences of an organism's genome. It has, on the other hand, a successful role in determining biodiversity and genetic relationships between and within sheep populations (Cho et al., 1998). Many conserved non-coding sequences are far from genes and cannot be assigned to defined functional classes but for evolutionary studies. Evolutionary studies of many prokaryotes and eukaryotes have been reported, covering evolutionary distance and phylogeny utilizing a DNA sequencing approach. The most common sequence technology is the bacteriophage T7 universal promoter (Dunn and Studier, 1983). The technology using the chain termination method is universally applicable to any piece of DNA and it has become the method of choice in most sequencing projects. The determination of genetic evolution based on DNA provides accurate information for genetic distance analysis allowing the ranking of populations according to the level of phylogenetic distinction (May, 1990). The establishment of the genetic phylogeny of populations and commercial strains is important for identifying common ancestors (Notter, 1999). DNA sequencing technology is widely used to study evolutionary distances and then to reconstruct phylogenies between and within different species or/and populations (Roderick and Holmes, 2009). This study aimed to reconstruct reliable molecular phylogenies for Awassi sheep populations from Jordan with Awassi sheep from the Kingdom of Saudi Arabia (KSA) using DNA sequencing technology.

MATERIAL AND METHODS

Sheep populations and sample collection

Thirty-two sheep individuals were randomly sampled from three different geographi-

cal areas of Jordan and KSA. Of the twenty-five Jordan Awassi sheep individuals, ten, nine, and six unrelated individuals represented sheep populations of the South, Middle, and North geographical parts of Jordan, respectively. In addition, seven unrelated individuals represented an outgroup population of Awassi sheep (Naemi) from a central region of KSA. A tissue sample of approximately 0.1 cm was taken from each individual and then stored in a deep freezer at -20°C until DNA extraction was performed.

DNA extraction and quantification

DNA extraction was performed using a commercially available kit/protocol of E.Z.N.A® MicroElute Genomic DNA extraction kit from OMEGA Bio-Tek; www.omega-biotek.com. Concentrations of DNA were estimated with a Nano-DNA spectrophotometer from AlphaSpec, www.alphainnotech.com, in which the quality of DNA was evaluated using the A260/A280 ratio. Each sample was diluted to give a DNA concentration of 10ng/µL.

DNA Sequencing

DNA sequencing service was provided by Macrogen® (www.macrogen.com) incorporating the universal primer T7promoter (5'-TAATACGACTCACTATAGGG-3') into a polymerase chain reaction (PCR) product using a high-throughput Applied Biosystems 3730XL sequencer.

Analysis of DNA sequences

Resulting DNA sequences were edited and assembled into contiguous sequences of 640 bp by using the MEGA program, version 5 (Tamura et al., 2011). The resulting DNA sequence information was analyzed for basic evolutionary analyses as detailed by Saitou and Nei (1987). Common analyses were evolutionary distances and phylogenetic reconstructions recommended by Kumar and Gadagkar (2001), Nei and Kumar (2000), and Tamura et al. (2011).

RESULTS AND DISCUSSION

Basic evolutionary analyses

Evolutionary statistics were performed to find the best model for evolutionary distances and phylogeny (Table 1). The best DNA model for estimating evolutionary distances is believed to be the nucleotide substitution model in which evolutionary distance between a pair of sequences is computed as the number of nucleotide substitutions occurring between them in a nucleotide-by-nucleotide comparison (Nei and Kumar, 2000). The results showed that the general time reversible (GTR) distance method was the best followed by the Hasegawa-Kishino-Yano (HKY) method (Table 1). The GTR model had the lowest Bayesian information criterion (BIC) score which was considered to be the best substitution pattern. Nei and Kumar (2000) stated that the best DNA nucleotide substitution model for estimating evolutionary distances must have the lowest BIC score. The Akaike information criterion corrected value (AICc) and estimated values of transition/transversion bias (R) are shown for each model in

Table 1. The GTR model was then considered for calculating the nucleotide frequency, substitutions, and further evolutionary analyses.

Table 1. Maximum likelihood fits of 24 different nucleotide substitution models.			
Model	BIC	AICc	R
GTR	34962	34445	0.78
HKY	34967	34480	0.78
GTR+G	34970	34446	0.8
GTR+I	34971	34447	0.78
HKY+G	34974	34479	0.8
TN93	34976	34482	0.78
HKY+I	34976	34482	0.78
TN93+G	34979	34478	0.81
GTR+G+I	34979	34448	0.8
HKY+G+I	34983	34481	0.8
TN93+I	34985	34484	0.78
TN93+G+I	34989	34480	0.81
T92	35035	34563	0.8
T92+G	35038	34559	0.82
T92+I	35045	34565	0.8
T92+G+I	35047	34561	0.82
K2	35082	34618	0.8
K2+G	35084	34612	0.82
K2+I	35091	34620	0.8
K2+G+I	35093	34614	0.82
JC	35152	34695	0.5
JC+G	35161	34697	0.5
JC+I	35161	34697	0.5
JC+G+I	35170	34699	0.5

GTR = general time reversible; HKY = Hasegawa-Kishino-Yano; TN93 = Tamura-Nei; T92 = Tamura 3-parameter; K2 = Kimura 2-parameter; JC = Jukes-Cantor; G = means using a discrete gamma distribution; I = means using a certain fraction of sites that are evolutionarily invariable. BIC = Bayesian information criterion; AICc = Akaike information criterion corrected value; R = assumed or estimated value of transition/transversion rate for each model.

In GTR distance model, the analysis involved 4 nucleotide sequences and codon positions, including 1st, 2nd, 3rd, and noncoding. The frequency of each nucleotide of adenine (A), cytosine (C), guanine (G), and thymine (T), was 0.28, 0.255, 0.2, and 0.265, respectively. Whereas the rate of nucleotide substitution ranged from 0.039 (GC) to 0.125 (GT) (Figure 1). As a consequence, the substitution pattern tests of homogeneity between sequences were performed to choose the best test to compute evolutionary distances between the sheep populations studied. The tests performed were pattern disparity index, maximum likelihood estimate of substitution matrix, and maximum composite likelihood method, as recommended by Kumar and Gadagkar (2001). Table 2 shows the values of the base composition bias disparity between and within sequence pairs in all populations studied. The disparity index per site is shown for all sequence pairs. Values greater than zero indicated the larger differences in base composition biases than expected on the basis of evolutionary divergence between sequences and by chance alone. The highest value was between KSA sheep group (outgroup) with the Jordan South sheep group, indicating the largest differences in base composition of the sequence pairs. However, the lowest value was found between the KSA Awassi sheep outgroup with Jordan Middle sheep population group. On the other hand, values of base composition bias disparity within sequences in all populations were the lowest in the Middle sheep popula-

tion and the largest in the North sheep population. This might be a result of high heterogeneous substitution patterns found within North sheep populations.
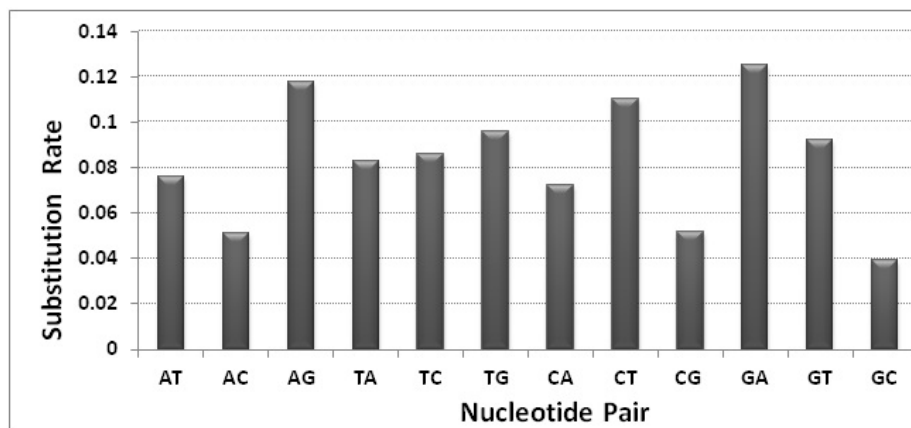

Figure 1. Nucleotide base substitutions rates for each nucleotide pair.

Table 2. Estimates of base composition bias disparity between and within sequences in all population groups.

	Middle JO sheep	North JO sheep	KSA sheep	South JO sheep
Middle JO sheep	**0.029**			
North JO sheep	1.311	**1.052**		
KSA sheep	0.102	1.631	**0.014**	
South JO sheep	1.621	0.528	1.845	**0.304**

JO = Jordan; KSA = Kingdom of Saudi Arabia. Values below diagonal represent the disparity index between populations, and diagonal bold values represent disparity index within each population.

The second test, maximum likelihood estimate of substitution method, was also performed in which the probability of substitution from one base to another was considered (Table 3). Rates of different transitional substitutions and those of transversion substitutions are shown in the Table 3. Each entry represents the probability of substitution from one base (row) to another base (column). Substitution pattern and rates were estimated under the GTR model using a discrete gamma distribution and evolutionarily in variable sites (GTR+G+I). Rates of different transitional substitutions are shown in bold, where the highest value was 12.64 for substitution A with G. Those of transversion substitutions are shown in italics with the highest value for substitution G with T (9.69) (Table 3).

Table 3. Maximum likelihood estimate of substitution matrix.

	A	T	C	G
A	-	*7.41*	*5.02*	**11.96**
T	*8.16*	-	**8.66**	*9.69*
C	*7.05*	**11.03**	-	*5.19*
G	**12.64**	*9.30*	*3.91*	-

(-) Each entry is the probability of substitution from one base (row) to another base (column). (-) Rates of different transition substitutions are shown in bold and those of transversion substitutions are shown in italics.

The last test, maximum composite likelihood method that is defined as a sum of related log-likelihoods (Tamura et al., 2004), was also computed. Since all pairwise distances in a distance matrix have correlations due to the phylogenetic relationships between the sequences, the sum of their log-likelihoods is a composite likelihood (Table 4). As in the previous tests, each entry in Table 4 shows the probability of substitution from one base to another base. Rates of different transition substitutions are shown in bold and those of transversion substitutions are shown in italics. In this test, the nucleotide frequencies were 26.83, 26.37, 26.78, and 20.01% for A, T, C, and G, respectively. The transition/transversion rate ratios were higher (k_1 = 1.542) for purines (A and G) than for pyrimidines (T and C) (k_2 = 1.46). The overall transition/transversion bias was R = 0.756, giving the probability of substitution from one base to another. Results from the selection test based on Tajima's neutrality test reported that the number of segregation sites were 416 with nucleotide diversity of 0.724. In general, recent ovine breeds showed signs of selection. In fact, the recent interrogation of 50,000 genetic loci provided unprecedented insight into the phylogeographic structure of sheep populations and the results of centuries of breeding practices of combining a global sheep sample (Kijas et al., 2012).

Table 4. Maximum composite likelihood estimate of nucleotide substitution.

	A	T	C	G
A	-	*7.53*	*5.71*	**11.79**
T	*7.66*	-	**8.34**	*7.64*
C	*7.66*	**10.99**	-	*7.64*
G	**11.81**	*7.53*	*5.71*	-

(-) Each entry shows the probability of substitution from one base (row) to another base (column). (-) Rates of different transition substitutions are shown in bold and those of transversion substitutions are shown in italics.

It seems that both methods assume nearly the same transition substitution rate for T and C nucleotides and different rates for A and G nucleotides. On the other hand, the transversion substitution rates were higher in the maximum likelihood estimate for all nucleotides except nucleotide A. Generally, the sums of nucleotide substitution pattern and substitution rates among sites in the two methods were different. In summing up, the maximum composite likelihood model was found to be more reliable for performing evolutionary analyses in this study. This summation is in agreement with Tamura et al. (2004), who reported that pairwise distances and related substitution parameters are accurately estimated by maximizing the composite likelihood than with other methods. Moreover, Tamura et al. (2004) found that, unlike cases of ordinary independent estimation of each pairwise distance, a complicated model had virtually no disadvantage in the composite likelihood method for phylogenetic analyses. Therefore, estimation of evolutionary distances and phylogeny reconstruction in the following section were conducted using the maximum composite likelihood model as implemented by Tamura et al. (2011).

Estimating evolutionary distances

Evolutionary distances are fundamental in molecular evolution study, phylogenetic reconstructions and the estimation of divergence times (Nei and Kumar, 2000). Most of widely used methods for distance between pair of sequences are usually measured by the number of

nucleotide substitutions occurring between them. The number of base substitutions per site between sequences is shown in Table 5. Analyses were conducted using the maximum composite likelihood model (Tamura et al., 2004). The number of base substitutions per site between sequences revealed reliable differences between the strains studied (Table 5). The results showed that the lowest number of base substitutions per site and the shortest genetic distance was between the North sheep population and KSA sheep population (outgroup) (0.545). This revealed that the two were much more closely related to each other than to the other groups. On the other hand, the North sheep showed closer evolutionary distance to the South sheep population (0.619) than to the Middle sheep population (0.643). Quite reliable differences were found between the KSA sheep outgroup population with the Middle sheep population (0.580), where the greatest distance was between the KSA sheep and South sheep (0.901). It could be assumed that the substitution rate of bases per site between the North sheep and KSA sheep was low for the DNA sequences studied. Furthermore, it should be noted that there was a possibility of closer ancestry and low selection sites in the sequence studied between those two populations in comparison with others. This is in agreement with the geographical information about these two populations and their common origin of the junction of north-east parts of Jordan and KSA (FAO, 1985). The possibility of common ancestry could go back centuries where the nomadic system of sheep production was dominant when nomads moved with their flocks all the year round looking for water and range for their animals in both parts of the countries in the Syrian desert. Currently, the transhumant system is still widely used across borders following certain routes of migration of residential tribes in region countries towards the east and the west sides (Valle Zerate et al., 2006). In addition, it is well know that Awassi lamb sheep from Jordan and Syria are exported to KSA for fattening and meat marketing, accounting for recent gene flow (Bourn, 2003). Gene flow was most commonly noted in sheep. In particular, Meadows et al. (2005) reported that mitochondrial sequences revealed high levels of gene flow between breeds of domestic sheep from Asia and Europe.

Table 5. Estimates of evolutionary divergence between sequences.

Population	North JO sheep	Middle JO sheep	KSA sheep
North JO sheep	-	-	-
Middle JO sheep	0.643	-	-
KSA sheep	0.545	0.580	-
South JO sheep	0.619	0.673	0.901

JO = Jordan; KSA = Kingdom of Saudi Arabia.

Phylogeny tree reconstruction

Reconstruction of the evolutionary history of populations is currently one of the most important subjects in molecular evolution. The reliable phylogenies shed light on the sequence of evolutionary events that generated the present day diversity and help to understand the mechanisms and history of evolution. Phylogenetic relationships of the sheep populations under investigation were described in a treelike form with a root utilizing three methods for constructing phylogenetic trees from molecular sequencing data. They are maximum likelihood (Li, 1997), neighbor-joining (Roderick and Holmes, 2009), and UPGMA (Nei and Kumar, 2000).

The phylogenetic trees in Figures 2, 3, and 4 revealed a considerable degree of differentiation of populations, supporting the results of evolutionary distances presented in Table 5.

The evolutionary history was inferred by using the maximum likelihood method based on the GTR model (Figure 2). The tree with the highest log likelihood (-17122.32) is shown in Figure 2. A discrete gamma distribution was used to model evolutionary rate differences between sites (+G, parameter = 36.95). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. Figure 2 shows the maximum likelihood tree construction of four major clusters (clades). The KSA sheep population considered in one cluster clearly separated from the North sheep population, which was in the second closer cluster. The KSA sheep population was clustered together with one root with the North sheep population, where the C nucleotide was a common nucleotide ancestor. Shorter topology from the root was seen within the KSA population with the most probable ancestors of T and C nucleotides. This maximum likelihood phylogenetic tree agreed with the previous results of evolutionary distances (Table 5). The results of maximum likelihood tree also indicated the most probable nucleotide ancestor for both populations with both Middle and South populations as indicated in Figure 2. A common rooted nucleotide was T, which separated the study population into two major groups, each of two clusters, whereas the G nucleotide was mostly within the Middle population. Finally, the South sheep population group was rooted with all individuals studied by the most probable A nucleotide ancestor.
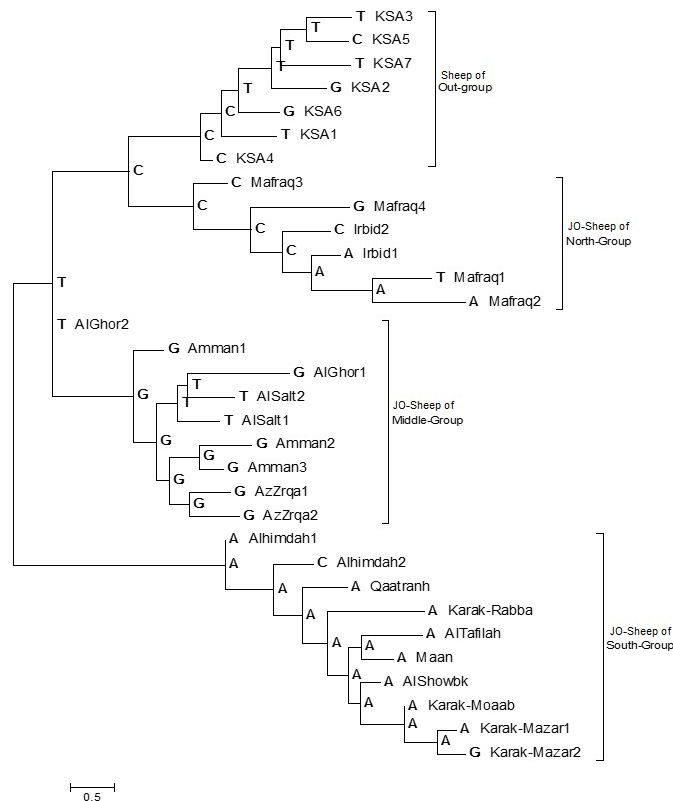

Figure 2. Molecular phylogenetic analysis by the maximum likelihood method with most probable ancestor of nucleotides (A, T, C, G) among the sheep populations studied.
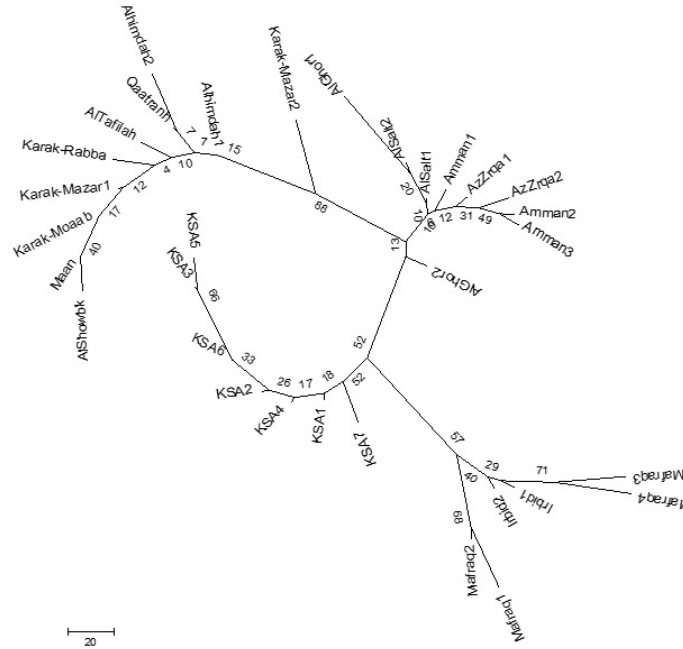
Figure 3. Evolutionary relationships of taxa using the neighbor-joining method.
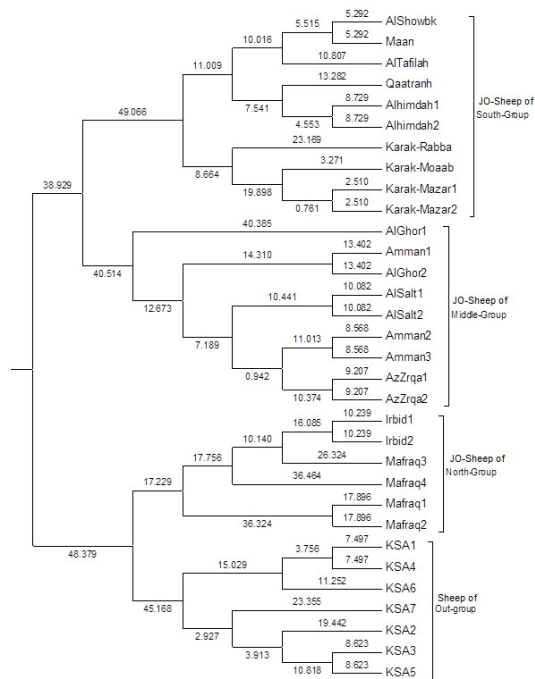

Figure 4. Evolutionary relationships of taxa using UPGMA tree among the sheep populations studied.

The evolutionary history was also inferred using the neighbor-joining method. The optimal tree with the sum of branch length (28.36) is shown in Figure 3. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Figure 3 shows the neighbor-joining tree construction of many branches representing closer distances of clusters considering separated branches from the South population that were divided into two branches. The first branch stated that there was more evolutionary relationship to each other than the other populations of Al-Karak, Altafilah, Maan, Al-Showbk and Alhimadeh, and Qaatranah that clustered in the second group. The Middle population was branched halfway to the other end of the tree where KSA sheep individuals clustered in a few branches. In general, the populations assorted into the same cluster have similar features. Such situation was likely due to sharing common sequences with less substitution rates from previous ancestors. For example, Al-Atiyat et al. (2012) reported that Jordan fat-tailed Awassi sheep populations of southern parts were genetically closer to each other according to the neighbor-joining method. Furthermore, it was shown that the neighbor-joining method was the best to infer evolutionary history of many species, such as clustering horseshoe crabs into seven clades (Ismail and Sarijan, 2011). Genetic subdivision of 57 European and Middle-Eastern sheep breeds revealed distinct groups of Middle-Eastern fat-tailed sheep, south-eastern European sheep and north-western/western European sheep (Peter et al., 2007).

The UPGMA method was also used to infer evolutionary history of the optimal tree with the sum of branch length around 28.35 (Figure 4). The tree was drawn to scale, with branch lengths (next to the branches) in the same units as those of the evolutionary distances used to infer the previous phylogenetic tree. Figure 4 shows the UPGMA tree for the sheep populations studied with three clusters and two roots. The tree clustered the populations as in the neighbor-joining tree regardless of topology length. The South sheep population considered in one cluster separated from the closer cluster. The third cluster consisted of the KSA population and North population indicating greater evolutionary relation to each other. These results support the results aforementioned of the maximum likelihood tree (Figure 2) and evolutionary distances (Table 5). In general, they were in contrast to the neighbor-joining tree and UPGMA, in which the KSA population and Middle population grouped in a closer cluster. This different result was most likely attributable to the different models used for estimation of evolutionary distances.

Overall, it could be concluded that on the basis of the results of the three phylogenetic methods, there is high evolutionary differentiation in Awassi sheep populations and between Jordan populations and the KSA outgroup population. Similar results were reported earlier in which the genetic phylogeny of Awassi sheep population was high when compared to different world sheep (Arranz et al., 2001). Furthermore, an example of a phylogenetic tree study among some geographical Awassi sheep ecotypes was also reported (Al-Atiyat et al., 2012). On the basis of microsatellite data, considerable genetic differentiation was recently reported in Saudi sheep (Musthafa et al., 2012).

The reconstruction of three major phylogentic trees indicated that the most interesting population was the KSA sheep as the sheep outgroup that shared close phylogeny with the North sheep (Figure 2) and Middle sheep (Figures 3 and 4). It was most probable that the KSA population had common ancestral breeds as well as gene flow with those two populations in the past. On the other hand, the North and Middle populations were in a close group of clusters compared to the others. Ozdemir et al. (2011) reported similar results of using various clustering methods for exhibiting two large clusters of different sheep populations in a phylogenetic

tree, where Awassi sheep was clearly separated from the other distinctive breeds. Furthermore, sequence comparisons showed that Awassi as one Asian sheep and European domestic sheep differed by 4.43% as separate lineages (Hiendleder et al., 1998). The Awassi sheep sequence diverged from Mouflon sheep by 4.52%, supporting the hypothesis that some modern domestic sheep and European Mouflon derive from a common ancestor, thus, providing evidence of an additional wild ancestor from Asia. Another study suggested that some mitochondrial lineage sheep arrived in northern Europe from the Near East fat-tailed sheep across Russia (Tapio et al., 2006).

CONCLUSION

Different phylogenetic trees reconstructed from 0.64-kb sequences provided reliable evidence for a close phylogenetic relatedness among Awassi sheep populations. The genetic relatedness between the three Awassi sheep populations from Jordan was evolutionarily separated from the KSA Awassi sheep that were considered as the outgroup. There was no distinctive phylogenetic branching of population clustering topology considering the different models applied. The maximum composite likelihood method was considered to be the best substitution pattern test of homogeneity between sequences to compute evolutionary distances between the populations studied. Quite reliable differences were found between the Jordan sheep population and Saudi population as a separate cluster. The evolutionary distances within individuals of the South sheep population showed the greatest distance with the North sheep population, revealing their closer relation compared with individuals of the other Jordanian sheep population. On the other hand, KSA sheep showed a closer relation and shorter evolutionary distance with the North sheep population compared with others. This different pattern was most likely attributable to the different models used for estimating distances and reconstructing phylogenies. The overall results support geographical information and ecological types of sheep populations from the countries of the study. It also shed light on the sequence of evolutionary events that generated the present day Awassi sheep populations studied. Summing up, the resulting phylogeny trees, on one hand, may help confirm historical and geographical information of common ancestors and genetic resource of Awassi sheep in the two countries. On the other hand, the overall results may contribute to the limited information of genetic diversity and relatedness of Awassi sheep populations in Arab countries.

ACKNOWLEDGMENTS

REFERENCES

Al-Atiyat RM, Tabbaa MJ, Salameh N and Tarawneh K (2012). Analysis of genetic variation of fat tailed-sheep in southern region of Jordan. *Asian J. Anim. Vet. Adv.* 7: 378-389.

Arranz JJ, Bayón Y and San PF (2001). Differentiation among Spanish sheep breeds using microsatellites. *Genet. Sel. Evol.* 33: 529-542.

Bourn D (2003). Livestock Dynamics in the Arabian Peninsula. Oxford: Consultancy Report by Environmental Research Group Oxford Limited to the Food and Agriculture Organization of the United Nations, Rome, 28-32.

Cho RJ, Fromont-Racine M, Wodicka L, Feierbach B, et al. (1998). Parallel analysis of genetic selections using whole genome oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A.* 95: 3752-3757.

Dunn JJ and Studier FW (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* 166: 477-535.

FAO (Food and Agriculture Organization) (1985). Small Ruminant Production in the Developing Countries, Rome.

Galal S, Gürsoy O and Shaat I (2008). Awassi sheep as a genetic resource and efforts for their genetic improvement - A review. *Small Rumin. Res.* 79: 99-108.

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

Ismail N and Sarijan S (2011). Phylogenetic inference from 18S rRNA gene sequences of horseshoe crabs, *Tachypleus gigas* between Tanjung Dawai, Kedah and Cherating, Pahang, Peninsular Malaysia. *Int. J. Agric. Biol. Sci.* 1: 95-98.

Kijas JW, Lenstra JA, Hayes B, Boitard S, et al. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* 10: e1001258.

Kumar S and Gadagkar SR (2001). Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. *Genetics* 158: 1321-1327.

Li WH (1997). Molecular Evolution. Sinauer Associates, Massachusetts.

Mason IL (1984). Evaluation of Domesticated Animals. Longman Group Ltd., Harlow.

May RM (1990). Taxonomy as destiny. *Nature* 347: 129-130.

Meadows JR, Li K, Kantanen J, Tapio M, et al. (2005). Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe. *J. Hered.* 96: 494-501.

Musthafa MM, Aljummah RS and Alshaik MA (2012). Genetic diversity of Najdi sheep based on microsatellite analysis. *Afr. J. Biotechnol.* 11: 14868-14876.

Nei M and Kumar S (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Notter DR (1999). The importance of genetic diversity in livestock populations of the future. *J. Anim. Sci.* 77: 61-69.

Ozdemir M, Bilgin OC and Esenbuga N (2011). Determination of genetic distance between Turkish breeds with various methods using some blood protein loci. *J. Anim. Plant Sci.* 21: 459-464.

Peter C, Bruford M, Perez T, Dalamitra S, et al. (2007). Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. *Anim. Genet.* 38: 37-44.

Piper L and Ruvinsky A (1997). The Genetics of Sheep. CAB International, New York.

Roderick DMP and Holmes EC (2009). Molecular Evolution: A Phylogenetic Approach, John Wiley and Sons. Blackwell Science Oxford.

Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.

Tamura K, Nei M and Kumar S (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U. S. A.* 101: 11030-11035.

Tamura K, Peterson D, Peterson N, Stecher G, et al. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731-2739.

Tapio M, Marzanov N, Ozerov M, Cinkulov M, et al. (2006). Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas. *Mol. Biol. Evol.* 23: 1776-1783.

Valle Zerate A, Musavaya K and Schaefer C (2006). Gene Flow in Animal Genetic Resources. A Study on Status, Impact and Trends. In: Institute of Animal Production in the Tropics and Sub-Tropics. University of Hohenheim. Germany.