

Original Research Article**Cytochrome b diversity and phylogeny of six Egyptian sheep breeds**

Running title: Cytochrome b diversity of Egyptian sheep

Abstract

Aim: Cytochrome b (*Cyt-b*) regions of mtDNA have received more attention due to their role in the genetic diversity and phylogenetic studies in different livestock. By using Cytochrome B sequencing, we aimed to clarify the genetic affinities and phylogeny of six Egyptian sheep breeds.

Methodology: The genomic DNA was extracted and the specific primers were used for conventional PCR amplification of the Cytochrome b region of mtDNA (1134-bp) in sheep. The alignment of sequences was done to identify the sequence variations and polymorphic sites in the amplified fragments.

Results: The results showed the presence of 39 polymorphic sites leading to the formation of 29 haplotypes (MG407500 - MG407528) with total haplotype diversity 0.814 and nucleotide diversity 0.00359. The lowest distance was observed between Rahmani and Saidi while the highest distance was observed between Ossimi and Sohagi. The sequences of 111 analyzed samples were aligned with references sequences of different haplogroups; A, B, C, D and E. The result showed that 86 out of 111 tested animals cluster with haplogroup B (77.48%), whereas 12 tested animals cluster with each of both haplogroups A and C (10.81%) and one animal belongs to haplogroup E (0.90%) with the absence of haplogroup D.

Conclusion: Cytochrome b regions of mtDNA have an important role in the genetic diversity and phylogenetic studies in farm animals as well as many other mammalian species.

Key words: *Cyt-b*, Diversity, Genetic distance, Phylogeny, Egyptian sheep

1. Introduction

Sheep (*Ovis aries*) is considered one from the early domesticated livestock in the world since 8-11 thousand years ago [1]. Parallel to human migration and trade, sheep was spread widely all over the world and constitute a principal part in different human societies [2,3] especially in desert and rural societies where it is considered an important source of

meat, milk, wool and other products. Many reports suggested that the early sheep domestication began in the Fertile Crescent from Asian mouflon (*O. orientalis*). Currently, there are more than 1400 sheep breeds distributed in all geographical regions with different environmental challenges [4]. Human activities including - long term natural and intense artificial selection - lead to the formulation of the modern breeds with highly production and at the same time, the progressively substitution of the less productive, locally adapted, native breeds.

The genetic resources conservation of sheep breeds through the identification of genetic characterization and variability among different breeds has more attention to avoid the disappearance of the local breeds [5]. The advances in paleontological and molecular genetics help in the identification of the origins and migration expansion of domestic sheep [6-8]. Mitochondrial DNA sequence analysis declared the presence of 5 maternal lineages; A and B [9], C [10], D [11] and E [6]. Most ovine mtDNA investigations focused mainly on the Cytochrome b and control regions [10]. Cytochrome b (*Cyt-b*) gene is an important portion of mtDNA genome and it is used for the classification of different species and the assessment of phylogenetic relations among diverse mammalian species where its nucleotide sequence is highly conserved [12-15].

Due to the location of Egypt in the East-North gate of Africa where the inherent pathways of the pastoral trips for domesticated sheep populations was done through Sinai to North Africa and also due to the presence of fat tailed sheep breeds - character quite common in Turkey and Syria- where genotypes that seem quite primitive, the affinities and phylogeny of Egyptian sheep breeds well shed some lights on the domestication history and migration pathways of this economically important species. Using *Cyt-b* sequencing, this work aimed to assess the genetic variations among six Egyptian sheep breeds, the affinities of these breeds to different haplogroups and their phylogeny with shedding a light on sheep domestication and migration pathways.

2. Materials and Methods

2.1. Blood samples and genomic DNA extraction:

Blood samples were collected from 111 sheep animals belonging to six breeds reared in Egypt; 19 (Barki), 19 (Rahmani), 20 (Ossimi), 14 (Saidi), 22 (Sohagi) and 17 (Fallahi). Genomic DNA was extracted from the whole blood according to the method described by Miller et al. [16] with minor modifications. Briefly, blood samples were mixed with cold 2x

sucrose-triton and centrifuged at 5000 rpm for 15 min at 4°C. The nuclear pellet was suspended in lysis buffer, sodium dodecyl sulfate and proteinase K and incubated overnight in a shaking water bath at 37°C. Nucleic acids were extracted with saturated NaCl solution. The DNA was picked up and washed in 70% ethanol. The DNA was dissolved in 1xTE buffer. The DNA concentration was determined, using Nano Drop1000 thermo scientific spectrophotometer and then diluted to the working concentration of 50 ng/μl.

2.2. Polymerase chain reaction:

The PCR primers used to amplify *Cyt-b* gene were designed according to Meadows et al. [6]. The PCR amplifications were conducted in a 50 μl volume containing 5 μl of 10x reaction buffer, 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.2 μM each primer, 1.5 U *Taq* DNA polymerase and approximately 100 ng of sheep genomic DNA. The amplification conditions were as follows: initial denaturation at 95°C for 5 min followed by 35 cycles of denaturation at 95°C for 1 min, annealing at 58°C for 1 min, and extension at 72°C for 1.5 min, then the final extension at 72°C for 10 min. The PCR products were electrophoresed on 2% agarose gel stained with ethidium bromide to test the amplification success. The amplified products were purified with a DNA purification kit (ExoSap-IT, USB Corporation) according to the manufacturer's instructions to remove residual primers and dNTPs. *Cyt-b* sequencing was performed in Macrogen Incorporation (Seoul, Korea).

Cyt-b F: GTC ATC ATC ATT CTC ACA TGG AAT C

Cyt-b R: CTG GTC TTG TAA ACC AGA GAA GGA G

2.3. Data analysis:

Cytochrome b sequences of tested sheep animals were aligned together using the BioEdit software [17] in order to identify and trace individual haplotype mutations. Haplotype structure, sequence variation, average number of nucleotide differences (D) and average number of nucleotide substitutions (Dxy) per site between populations were calculated using DnaSP 5.00 software [18]. Neighbour-joining (NJ) tree for tested sheep sequences and the phylogenetic tree of Egyptian sheep animals with references sequences of different haplogroups; A, B, C, D and E was constructed using Mega 5.0 software [19].

3. Results and Discussion

The domestication history of human-related species such as cattle, sheep and goat is clarified from the recent study of finer variations in the genomes. In sheep, search approaches

for polymorphisms in broadband [7], polymorphisms insertions stabilized retroviruses [3] and the study of mitochondrial genome [6] have not yet determined the exact number of domestication events that occurred from wild populations to the present species in the region between Turkey and the Caucasus.

The first event of domestication would have happened since 10 000 to 11 000 years in the Fertile Crescent [20]; the birthplace of agriculture, urbanization and trade. Its existence is supported by the presence of the oldest sheep archaeozoological evidence in present-day in Iran, Turkey and Cyprus. The population of wild sheep behind this first wave of domestication was highly heterogeneous. In addition, the "genetic bottle-neck" was not pronounced in the domestication of sheep as it was for other domesticated breeds [7], which results in greater genetic variability. At first, the sheep were domesticated for meat, and the specialization for wool and milk began 5000 years ago in Southwest Asia and 4000 years in Europe.

Currently, more than 1400 different breeds of sheep are known for more than 1.1 billion animals [4]. However, 181 breeds are now extinct and more than 12% of identified breeds are threatened. They suffer the decrease of genetic variability due to the increase of inbreeding resulting from the selection methods exclusively oriented toward high productivity [21]. So, the maintenance of large genetic variability between sheep breeds in countries - with old, not highly selected breeds - is of more interest for genetic resources conservation.

Studies based mainly on sequencing of the control region of mitochondrial DNA showed that there are five maternal lineages for domestic sheep (*Ovis aries*), called haplogroups A, B, C, D and E. The haplogroups A and B were highlighted by Hiendleder et al. [9], haplogroup C by Pedrosa et al. [10], haplogroup D by Tapio et al. [11] and recently, Meadows et al. [6] added the haplogroup E.

Because of the eastern location of Egypt in the Mediterranean basin and the presence of fat tailed sheep breeds- character quite common in Turkey and Syria- where genotypes that seem quite primitive, the phylogenetic studies of Egyptian sheep breeds is become particularly attractive. Work of Chessa et al. [3] showed that African sheep populations

(Ossimi and Barki breeds in Egypt) have a common retrotype R2 with the population in Southwest Asia, suggesting direct migration links between these regions.

The tested sheep animals in this work belong to six Egyptian sheep breeds; Barki, Ossimi, Rahmany, Saidi, Sohagi and Fallahi. The amplified fragments represented *Cyt-b* gene at 1134-bp were aligned together using BioEdit software. DnaSP 5.00 software was used to identify the sequence variation and polymorphic sites in the aligned sequences. The result declared the presence of 39 polymorphic sites result in the formation of 29 different haplotypes (**Fig. 1**). The nucleotide sequences of these haplotypes were submitted to GenBank with the accession Nos.: MG407500 - MG407528. Haplotype no. 2 is the most common one which appeared in 43 animals belonging to all tested breeds followed by haplotype no. 1 which appeared in 19 animals whereas most of the other haplotypes excited in one animal (**Table 1**).

Table 1: The animals belonging to each detected haplotype

Haplotype No.	Animals
1	Bar1, Bar7, Bar9, Bar10, Bar13, Bar16, Rah11, Oss2, Oss6, Oss7, Sai3, Sai5, Sai6, Soh8, Fal1, Fal5, Fal11, Fal14
2	Bar2, Bar3, Bar5, Bar12, Rah1, Rah3, Rah4, Rah5, Rah6, Rah9, Rah13, Rah14, Rah16, Rah17, Rah19, Oss4, Oss5, Oss9, Oss12, Oss13, Oss14, Oss16, Sai1, Sai7, Sai8, Sai9, Sai11, Sai12, Sai13, Soh1, Soh3, Soh9, Soh10, Soh14, Soh18, Soh20, Soh22, Fal3, Fal6, Fal12, Fal15, Fal16, Fal17
3	Bar4, Rah7, Sai4, Sai10, Soh6, Soh12, Soh13
4	Bar6, Soh19
5	Bar8
6	Bar11, Bar18, Rah2, Rah8, Rah12, Rah15, Oss19, Fal7, Fal8, Fal13
7	Bar14
8	Bar15
9	Bar17
10	Bar19, Soh21
11	Rah10
12	Rah18
13	Oss1
14	Oss3
15	Oss8, Oss10, Oss17, Oss20
16	Oss11, Oss18
17	Oss15
18	Sai2
19	Sai14
20	Soh2, Soh4
21	Soh5, Fal10
22	Soh7

23	Soh11
24	Soh15
25	Soh16
26	Soh17
27	Fal2
28	Fal4
29	Fal9

Bar: Barki, Rah: Rahmani, Oss: Ossimi, Sai: Saidi, Soh: Sohagi, Fal: Fallahi

	111
	1111222333333334444455566677778888999011
	122301502267994789957839912381112239201
	736579892468367645831333632590398590746
Hap_1	TATTTGGCACGATTCAGGATGGTTTCTCCTACTAGCGAC
Hap_2A.....T.....
Hap_3A.T.....A.....T.....
Hap_4	...CA.T...CC.G.A.C...CT.T..G...AT...
Hap_5TA.....
Hap_6A.....A...T.....
Hap_7A.....A...TG.....
Hap_8G
Hap_9TA.....G
Hap_10	...CA.T...CC.G.A.....CT.T..G...AT...
Hap_11	...C.A.....T.....
Hap_12A.....T.....T.....
Hap_13A.....G.....T.....T.....
Hap_14A..
Hap_15	...CA.T...CCTG.A.....CT.T..G...AT...
Hap_16	...CA.T...CC.G.A...C.CT.T..G...AT...
Hap_17A.....AA...T..T.....
Hap_18	.G...A.T.....A.....T.....
Hap_19	C....A.....T.....
Hap_20	...CA.T...CC.G.AGC...CT.T..G...AT...
Hap_21	..C..A.....T.....
Hap_22	...CA.T...C..G.A.....CT...G...A....
Hap_23A.T.....A.....T...C.....
Hap_24A.....T.....T.....
Hap_25A.....C.T.....
Hap_26G.....C.....
Hap_27A.G.....A.....T.....T.....
Hap_28A.T.....CA.....T.....
Hap_29AA.....A...T.....

Fig. 1: The nucleotide sequences of 29 haplotypes.
The polymorphic (variable) site nos. in red.

The haplotype diversity recorded in tested six sheep breeds ranged from 0.643 in Rahmani breed (with an average number of nucleotide differences K: 0.982) to 0.871 in Barki breed (K: 4.491) whereas the total haplotype diversity and average number of nucleotide differences in all tested breeds were 0.814 and 4.074 respectively. On the other hand, the nucleotide diversity ranged from 0.00087 in Rahmani to 0.00595 in Ossimi breed with the total nucleotide diversity 0.00359 in all breeds (**Table 2**). The genetic distances (D) and the average number of pairwise differences (Dxy) between sheep breeds were estimated (**Table 3**). The lowest distance was observed between Rahmani and Saidi (D: 1.436 and Dxy: 0.00127) whereas the highest distance was observed between Ossimi and Sohagi (D: 6.050 and Dxy: 0.00534) (**Fig. 2**).

Table 2: The genetic diversity data of the six sheep breeds

Breeds	No. of sequences (N)	No. of polymorphic sites (S)	No. of haplotypes (H)	Haplotype diversity (Hd)	Average number of pairwise differences (K)	Nucleotide diversity (π)
Barki	19	19	10	0.871	4.491	0.00396
Rahmani	19	7	6	0.643	0.982	0.00087
Ossimi	20	21	8	0.837	6.753	0.00595
Saidi	14	6	5	0.725	1.736	0.00153
Sohagi	22	21	12	0.861	5.468	0.00482
Fallahi	17	9	7	0.824	2.066	0.00182
Total	111	39	29	0.814	4.074	0.00359

Table 3: Average pairwise differences between six sheep breeds

	Barki	Rahmani	Ossimi	Saidi	Sohagi	Fallahi
Barki	-----	0.00275	0.00527	0.00286	0.00472	0.00298
Rahmani	3.116	-----	0.00417	0.00127	0.00338	0.00135
Ossimi	5.971	4.729	-----	0.00434	0.00534	0.00456
Saidi	3.248	1.436	4.921	-----	0.00351	0.00166
Sohagi	5.352	3.830	6.050	3.981	-----	0.00381
Fallahi	3.381	1.533	5.162	1.887	4.326	-----

Average number of nucleotide difference between breeds, D (below)

Average number of nucleotide substitution per site between breeds, Dxy (above)

The sequences of 111 analyzed samples were aligned with references sequences of different haplogroups and the Neighbour-Joining tree was constructed using Mega 5.0 software (**Fig. 3**). The phylogeny result showed the presence of four haplogroups; HapA, HapB, HapC and HapE in the examined animals whereas the haplogroup D was absent (**Table 4**). The result showed that 86 out of 111 tested animals cluster with haplogroup B

(77.48%), whereas 12 tested animals cluster with each of both haplogroups A and C (10.81%) and one animal belongs to haplogroup E (0.90%).

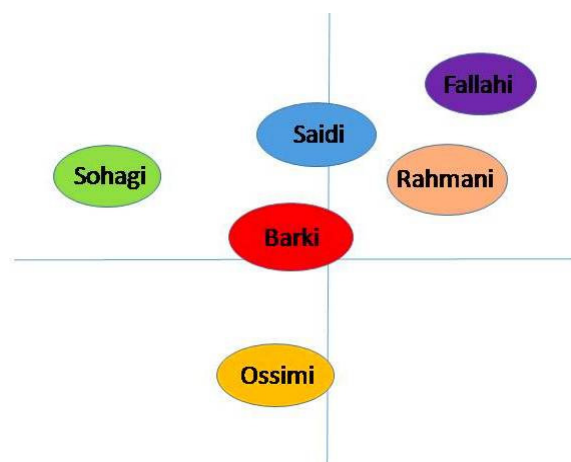


Fig. 2: A schematic diagram showed the genetic distances between sheep breeds

Table 4: The different haplogroups to which the tested animals are belonging

Breed	No. of samples	haplogroup A		Haplogroup B		haplogroup C		haplogroup E	
		No.	Freq.	No.	Freq.	No.	Freq.	No.	Freq.
Barki	19	1	5.3%	16	84.2%	2	10.5%	0	00.0%
Ossimi	20	1	5.0%	13	65.0%	6	30.0%	0	00.0%
Rahmani	19	1	5.3%	18	94.7%	0	00.0%	0	00.0%
Fallahi	17	2	11.8%	15	88.2%	0	00.0%	0	00.0%
Saidi	14	3	21.4%	11	78.6%	0	00.0%	0	00.0%
Sohagi	22	4	18.2%	13	59.1%	4	18.2%	1	4.5%
Total	111	12	10.81%	86	77.48%	12	10.81%	1	0.9%

MtDNA sequence studies help in answer to some important queries around the sheep domestication process like the phylogenetic relationships among wild and domestic sheep [22], divergence times between the five major maternal lineages [6,10], demographic history [23-24], origins of different mtDNA lineages [8,11,25] and the patterns of gene flow [26,27].



Fig. 4: A schematic diagram showed the entrance ways of different haplogroups to Egypt

Moreover, the absence of haplogroup D and the presence of only one animal cluster with haplogroup E is a logical result because these haplogroups are present in Caucasus area which is far away from Egypt and there is no any sheep migration was done from this area.

The appearance frequencies of the different haplotypes in Egyptian sheep breeds depend on the existence places of these breeds (**Fig. 5**). For example, the haplotype B is dominant in Rahmani, Fallahi and Barki breeds which are present in Northern and Middle Egypt near to the entrance road of this haplogroup from Central Asia via Sinai. Also the presence of haplotypes A and C at high frequencies in sheep breeds present in Southern region of Egypt (Saidi and Sohagi breeds) due to the entrance of these haplogroups from Eastern and Southern Asia to Egypt via Red Sea close to these regions.

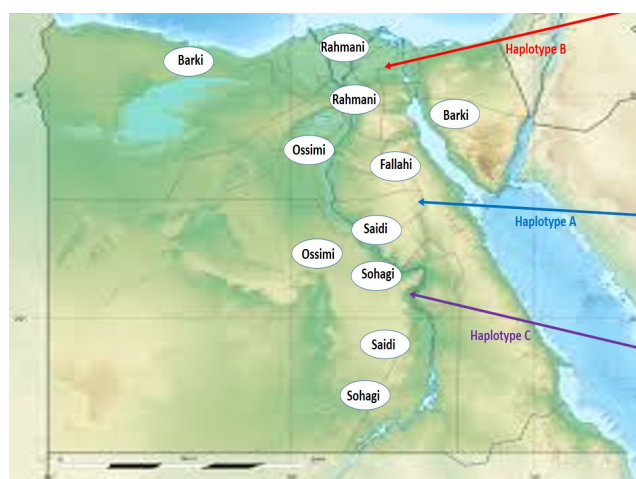


Fig. 5: A schematic diagram showed the geographic distribution of tested sheep breeds

References

1. Ryder ML. Sheep, pp. 63-84. In: *Evolution of Domesticated Animals*. Edited by I.L. Mason. Longman Group, London/New York; 1984.
2. Colledge S, Conolly J, Shennan S. The evolution of Neolithic farming from SW Asian origins to NW European limits. *Eur J Archaeol* 2005; 8: 137-156.
3. Chessa B, Pereira F, Arnaud F. Revealing the history of sheep domestication using retrovirus integrations. *Sci* 2009; 324: 532-536.
4. Scherf BD. *World Watch List for Domestic Animal Diversity*. Ed. 3. Food and Agriculture Organization of the United Nations, Rome, 2000.
5. Warmuth V, Eriksson A, Bower MA, Barker G, Barrett E, Hanks BK, Li S, Lomitashvili D, Ochir-Goryaeva M, Sizonov GV. Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proc Natl Acad Sci* 2012; 109: 8202-8206.
6. Meadows JR, Hiendleder S, Kijas JW. Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity* 2011; 106: 700-706.
7. Kijas JW, Lenstra JA, Hayes B, Boitard S, Porto Neto LR. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol* 2012; 10(2): e1001258.
8. Demirci S, Koban BE, Dag ND, Pi E, Engin A, O€ F, Y€unc€u E, Do SA, Togan I. Mitochondrial DNA diversity of modern, ancient and wild sheep (*Ovis gmelinii anatolica*) from Turkey: new insights on the evolutionary history of sheep. *PLoS One* 2013; 8:e81952.
9. Hiendleder S, Kaupe B, Wassmuth R, Janke A. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc. Royal Soc. London Series B Biol Sci* 2002; 269: 893-904.
10. Pedrosa S, Uzun M, Arranz JJ, Gutierrez-Gil B, San Primitivo F, Bayon, Y. Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. *Proc Biol Sci* 2005; 272(1577): 2211-2217.
11. Tapio M, Marzanov N, Ozerov M, Cinkulov M, Gonzarenko G. Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas. *Mol Biol Evol* 2006; 23: 1776-1783.
12. Zhao Y, Zhao E, Zhang N, Duan C. Mitochondrial DNA diversity, origin, and phylogenetic relationships of three Chinese large-fat-tailed sheep breeds. *Trop Anim Health and Prod* 2011; 43: 1405-1410.
13. Haus T, Akom E, Agwanda B, Hofreiter M, Roos C, Zinner D. Mitochondrial diversity and distribution of African green monkeys (*chlorocebus gray*, 1870). *Am J Primatol* 2013; 75(4): 350-360.

14. Zhong X, Wang N, Hu D, Wang J. Sequence analysis of cytb gene in *Echinococcus granulosus* from Western China. *Korean J Parasitol* 2014; 52: 205-209.
15. Othman OE, Abd El-Kader HA, Sally SS, Abd El-Aziem SH. Cytochrome b conservation between six camel breeds reared in Egypt. *J. Genet. Eng. Biotech* 2017; 15(1): 1-6.
16. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; 16: 1215.
17. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 1999; 41: 95-98.
18. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 2009; 25:1451-1452.
19. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular (evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. *Mol Biol Evol* 2011; 28: 2731-2739.
20. Peter C, Bruford M, Perez T. Genetic diversity and subdivision of 57 European and Middle-Eastern sheep breeds. *Anim Genet* 2007; 38: 37-44
21. Taberlet P, Valentini A, Rezaei HR. Are cattle, sheep and goats endangered species?. *Mol Ecol* 2008; 17: 275-284.
22. Meadows JR, Cemal I, Karaca O, Gootwine E, Kijas JW. Five ovine mitochondrial lineages identified from sheep breeds of the near East. *Genet* 2007; 175(3): 1371-1379.
23. Dobney K, Larson G. Genetics and animal domestication: new windows on an elusive process. *J Zool* 2006; 269: 261-271.
24. Zeder MA. Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact. *Proc Natl Acad Sci* 2008; 105:11597-11604.
25. Singh S, Kumar S, Kolte AP, Kumar S. Extensive variation and sub-structuring in lineage A mtDNA in Indian sheep: genetic evidence for domestication of sheep in India. *PLoS One* 2013; 8: e77858
26. Tapio M, Ozerov M, Tapio I, Toro MA, Marzanov N, Cinkulov M, Goncharenko G, Kiselyova T, Murawski M, Kantanen J. Microsatellite-based genetic diversity and population structure of domestic sheep in northern Eurasia. *BMC Genet* 2010; 11:76.
27. Cai DW, Tang ZW, Yu HX, Han L, Ren XY, Zhao XB, Zhu H, Zhou H. Early history of Chinese domestic sheep indicated by ancient DNA analysis of Bronze Age individuals. *J Archaeol Sci* 2011; 38: 896-902.
28. Othman OE, Pariset L, Balabel EA, Marioti M. Genetic characterization of Egyptian and Italian sheep breeds using mitochondrial DNA. *J Genet Eng Biotech* 2015; 13(1): 79-86.