



## Application of Multivariate Principal Component Factor Analysis to Morphological Characterization of Camels in Ethiopia

Keefelegn Kebede<sup>1,a,\*</sup>, Berhanu Bekele<sup>1,b</sup>, Sisay Tilahun<sup>2,c</sup>, Biresaw Serda<sup>3,d</sup>

<sup>1</sup>School of Animal and Range Sciences, Haramaya University, Ethiopia,

<sup>2</sup>Ethiopian Institute of Agricultural Research (EIAR), Addis Ababa, Ethiopia

<sup>3</sup>College of Agriculture and Environmental Sciences, Arsi University, Ethiopia

\*Corresponding author

### ARTICLE INFO

#### Research Article

Received : 27/11/2020

Accepted : 14/03/2022

#### Keywords:

Camel

Linear body measurements

Multivariate analysis

Factor analysis

Ethiopia

### ABSTRACT

This study was conducted to assess variability among linear body measurements (LBMs), deduce components that describe these traits, and investigate the inter-relationship among them. For this purpose, seventeen LBM traits namely heart girth, body-length, wither-height, ear-length, forelimb-length, hindlimb-length, barrel-girth, face-length, hip-width, chest-width, chest-depth, tail-length, neck-length, hump-length, hump-circumference, forehoof-circumference and hindhoof-circumference were measured on 300 (51 males and 249 females) camels. PC factor analysis was used to describe the variation in LBM traits where extracted factors were varimax rotated to enhance interpretability. Pearson's correlation coefficients among the traits were positive and very highly significant. From the factor analysis, two principal components (PCs) were extracted, which accounted for 63.2% of the total variance. PC1 accounted for 57.0% of the total observed variance and was loaded by EL, BL, FL, HL, FLL, WH, CD, NL, and HC; while PC2 contributed 6.1% of the total observed variance and had its loading on HG, BG, and HW. The results obtained from this study could be useful in designing appropriate husbandry, selection, and breeding programs for utilization of camel genetic resources.

<sup>a</sup> [kebede123@yahoo.de](mailto:kebede123@yahoo.de)

<sup>b</sup> <https://orcid.org/0000-0001-9278-9173>

<sup>c</sup> [sisayt9@gmail.com](mailto:sisayt9@gmail.com)

<sup>d</sup> <https://orcid.org/0000-0002-1648-5642>

<sup>b</sup> [berhanub06@gmail.com](mailto:berhanub06@gmail.com)

<sup>d</sup> <https://orcid.org/0000-0002-6659-2382>

<sup>d</sup> [berhanub06@gmail.com](mailto:berhanub06@gmail.com)

<sup>d</sup> <https://orcid.org/0000-0002-9437-1771>



This work is licensed under Creative Commons Attribution 4.0 International License

## Introduction

World Camel population is estimated to be around 35.5 million; where Somalia has the highest population of 7.2 million followed by Sudan 4.9 million and Ethiopia 1.3 million (FAOSTAT, 2018). The camel ecotypes in Ethiopia serve numerous functions (e.g., milk, meat, riding, packing) and thereby contribute significantly to the livelihood of the pastoralists and agro-pastoralists living in fragile environments (Abbas et al., 2000; Tura et al., 2010). In addition to this, pastoralists have kept and bred camels owing to exploit their extraordinary power to withstand thirst and hunger for long duration in the most inhospitable ecological conditions (Al-Dahash and Sassi, 2009).

Despite the high population of camel and its extremely considerable importance in Ethiopia, it has not received adequate attention from research and development institutions as well policymakers. Studies on camel production system, phenotypic and genetic characterization

are scanty (Yohannes et al., 2007) and show a serious lack of information (Gifford-Gonzalez and Hanotte, 2011). This hindered the design of an appropriate strategy for optimal utilization of the existing potential of camel genetic resources and the establishment of breeding programs. Taking into consideration of the current importance of camels in contributing to the livelihoods of pastoralists in marginal areas, and the role it plays towards resilience to the present climate change, it is thus imperative to identify and differentiate the phenotypic characteristics of camel populations in Ethiopia based on FAO guidelines.

In Ethiopia, the inter-relationship among LBMs of camels treated as multivariate (analyzed simultaneously) have not been widely exploited. Therefore, the present study makes use of PC factor analysis to assess variability among LBM traits, deduce components that describe these traits and investigate inter-relationship among them.

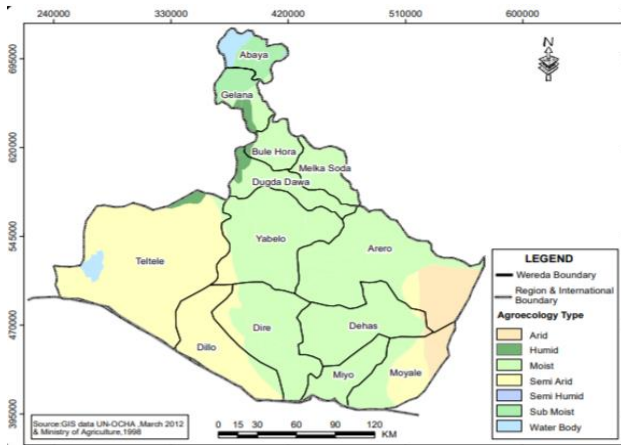


Figure 1. Map of the study area

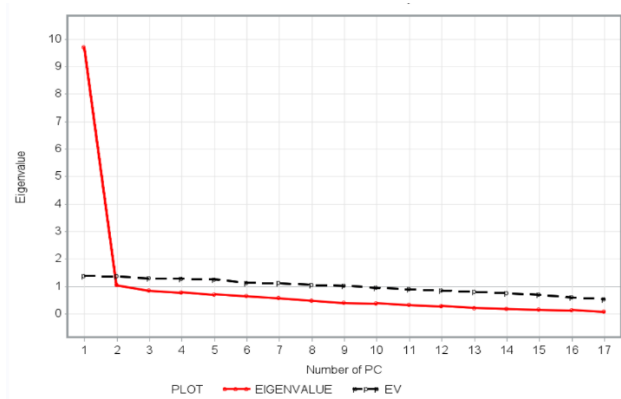


Figure 2. Scree and parallel analysis plots

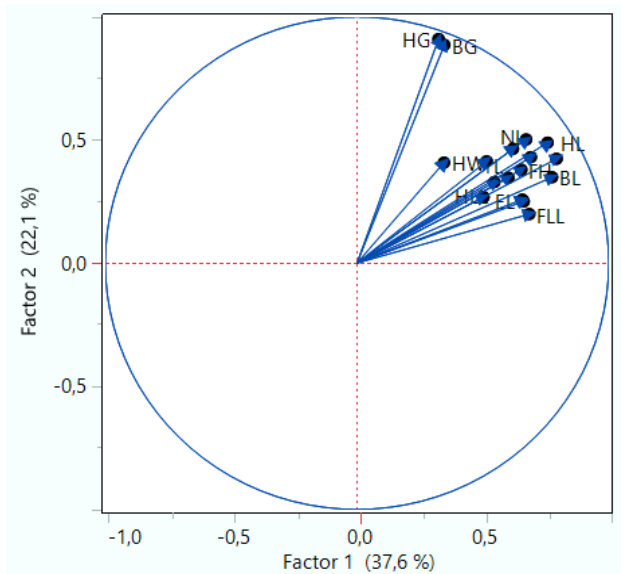


Figure 3. Factor pattern plot between the first two PC loadings.

## Materials and Methods

### Locations of the Study Area

The study was conducted in two districts namely, Yabello and Melka Soda in Southern Oromia regional state. Yabello district is one of the districts of the Borena zone. The district is situated in Latitude/Longitude: N 4° 52' 59.99" E 38° 4' 59.99". Melka Soda district is located in the northeastern part of West Guji zone. The location of Melka Soda Woreda is between 35° East and 30° West (Figure 1).

### Methods of Sampling and Data Collection

Discussions were held with the experts in the zonal and district pastoral development offices and representative pastoral community on the present production system and present condition and concentration of the Borena camels. A total of 300 mature camels (150 camels (27 males and 123 females) from Yabello district and (150 camels (24 males and 126 females) from Melka Soda district) were randomly selected. Seventeen morphometric traits namely heart girth, body length, wither height, ear length, forelimb length, hind limb length, barrel girth, face length, hip-width, chest width, chest depth, tail length, neck length, hump length, hump circumference, forehoof circumference and hind hoof circumference were recorded following the recommended FAO descriptors for camel genetic resources (FAO, 2012). Measuring tapes were used to measure the respective LBM traits.

### Statistical Data Analysis

All statistical analyses were performed using the SAS 9.4 software (version 9.4; SAS Institute Inc., NC). due to the larger numbers of females than males in livestock, such studies are mainly carried out on females (Bene et al., 2007; Ndumu et al., 2008; and Traore' et al., 2008). Thus, in this study, in order to avoid potential sampling bias due to the low number of males, only females were considered in the analysis.

### Exploratory Data Analysis

LBM traits were subjected to exploratory data analysis to get results of descriptive statistics and correlation matrices using the PROC UNIVARIATE and PROC CORR procedures of SAS (version 9.4; SAS Institute Inc., NC).

### PC Factor Analysis

#### Estimating the number of PCs

Several criteria are available for determining the number of PCs to be extracted. In this study, the criteria Kaiser–Guttman rule, the screen test, and parallel analysis plot were used.

#### Kaiser–Guttman Rule

This rule states that the number of PCs to be extracted should be equal to the number of PCs having an eigenvalue greater than 1. An eigenvalue greater than 1 indicates that PCs account for more variance than accounted for by one of the original traits in standardized data. Eigenvalues measure the amount of the variation explained by each PC and will be largest for the first PC and smaller for the subsequent PCs.

#### Scree Test

Plotting the eigenvalues against the corresponding PC produces a screen plot that illustrates the rate of change in the magnitude of the eigenvalues for an increasing number of PCs. The rate of decline tends to be fast first and then levelled off. The “elbow,” at which the curve bends, is considered to indicate the maximum number of PCs to extract.

Parallel analysis: To aid the decision making in the selection of the number of PCs extracted, a graphical method known as parallel analysis is suggested to enhance the interpretation of the scree plot. The optimum number of PCs is selected at the cut-off point, where the scree plot and the parallel analysis curve intersect.

Table 1. Descriptive statistic results of the traits

Nr.	Trait	Label	Min.	Max.	Mean±SE	STD	CV
1	Ear length	EL	6	14	10.4±0.11	1.66	16.00
2	Body length	BL	89	156	132.4±0.71	11.19	8.45
3	Face length	FL	99	162	143.8±0.71	11.18	7.78
4	Hump length	HL	136	194	179.6±0.58	9.18	5.11
5	Forelimb length	FLL	20	50	35.3±0.33	5.29	15.01
6	Wither height	WH	139	199	183.0±0.56	8.83	4.82
7	Heart girth	HG	154	230	203.8±0.85	13.48	6.62
8	Barrel girth	BG	180	250	226.3±0.82	12.99	5.74
9	Hip width	HW	25	50	38.4±0.21	3.26	8.50
10	Chest width	CW	28	43	37.1±0.19	3.02	8.13
11	Chest depth	CD	32	65	51.4±0.44	7.04	13.60
12	Neck length	NL	87	129	114.9±0.56	8.87	7.73
13	Tail length	TL	38	62	51.4±0.28	4.47	8.70
14	Hump circumference	HC	94	129	117.4±0.49	7.80	6.64
15	Hind limb length	HLL	18	39	28.1±0.26	4.05	14.44
16	Fore hoof circumference	FH	40	68	59.4±0.33	5.23	8.80
17	Hind hoof circumference	HH	34	62	52.2±0.36	5.63	10.79

SE = standard error of mean; STD = standard deviation; CV = coefficient of variation; Min. = Minimum; Max. = Maximum

Table 2. Phenotypic correlations and their statistical significance levels among body weight and morphometric traits of camels\*\*\*

Trait	EL	BL	FL	HL	FL2	WH	HG	BG	HW	CW	CD	NL	TL	HC	HL2	FH	HH
BW																	
EL	1																
BL	0.57	1															
FL	0.62	0.78	1														
HL	0.66	0.75	0.83	1													
FLL	0.42	0.63	0.65	0.62	1												
WH	0.56	0.70	0.73	0.75	0.58	1											
HG	0.44	0.57	0.69	0.64	0.40	0.63	1										
BG	0.46	0.58	0.70	0.65	0.41	0.61	0.92	1									
HW	0.36	0.34	0.45	0.45	0.27	0.37	0.47	0.49	1								
CW	0.45	0.48	0.50	0.53	0.37	0.48	0.41	0.41	0.38	1							
CD	0.36	0.64	0.53	0.52	0.50	0.56	0.56	0.53	0.33	0.39	1						
NL	0.55	0.62	0.73	0.65	0.58	0.59	0.62	0.63	0.50	0.42	0.47	1					
TL	0.48	0.60	0.58	0.63	0.47	0.56	0.52	0.50	0.41	0.48	0.57	0.55	1				
HC	0.57	0.69	0.75	0.74	0.53	0.65	0.67	0.68	0.45	0.40	0.57	0.69	0.57	1			
HLL	0.51	0.45	0.57	0.59	0.41	0.46	0.47	0.48	0.40	0.35	0.41	0.51	0.45	0.66	1		
FH	0.52	0.64	0.63	0.66	0.45	0.59	0.56	0.56	0.36	0.42	0.54	0.53	0.53	0.65	0.47	1	
HH	0.52	0.60	0.57	0.60	0.49	0.59	0.44	0.45	0.33	0.36	0.47	0.50	0.43	0.57	0.44	0.81	1

\*\*\* significant at P<0.001 for all correlation coefficients; ear length (EL), body length (BL), face length (FL), hump length (HL), forelimb length (FLL), wither height (WH), heart girth (HG), barrel girth (BG), hip-width (HW), chest width (CW), chest depth (CD), neck length (NL), tail length (TL), hump circumference (HC), hind limb length (HLL), fore hoof circumference (FH) and hind hoof circumference (HH)

Table 3. Eigenvalues and shares of total variance along with factor loadings after varimax rotation and communalities

Trait	EL	BL	FL	HL	FLL	WH	HG	BG	HW	CW	CD	NL	TL	HC	HLL	FH	HH
PC1	0.66	0.77	0.76	0.79	0.68	0.69	0.32	0.35	0.35	0.50	0.51	0.62	0.60	0.67	0.55	0.65	0.66
PC2	0.26	0.35	0.49	0.42	0.20	0.43	0.91	0.89	0.41	0.27	0.41	0.46	0.35	0.50	0.33	0.38	0.25
Communalities	0.49	0.72	0.81	0.81	0.51	0.66	0.93	0.90	0.29	0.32	0.44	0.60	0.48	0.70	0.41	0.60	0.50
		PC1				PC2				Total							
Eigenvalue		9.76				1.04											
% of total variance		57.04				6.13											

Ear length (EL), body length (BL), face length (FL), hump length (HL), forelimb length (FLL), wither height (WH), heart girth (HG), barrel girth (BG), hip-width (HW), chest width (CW), chest depth (CD), neck length (NL), tail length (TL), hump circumference (HC), hind limb length (HLL), fore hoof circumference (FH) and hind hoof circumference (HH)

*PC Loading*

They are correlation coefficients between the PC scores and the original traits. A high positive correlation between PC1 and a trait indicates that the trait is associated with the direction of the maximum amount of variation in the

dataset. A strong correlation between a trait and PC2 indicates that the trait is responsible for the next largest variation in the data perpendicular to PC1, and so on.

## Results and Discussion

Table 1 shows the mean  $\pm$  standard error, standard deviation, coefficient of variation, minimum, and maximum estimates of body weight and morphometric traits of the camels.

The mean values estimated for the traits were 10.4 (EL), 132.4 (BL), 143.8 (FL), 179.6 (HL), 35.3 (FLL), 183.0 (WH), 203.8 (HG), 226.3 (BG), 38.4 (HW), 37.1 (CW), 51.4 (CD), 114.9 (NL), 51.4 (TL), 117.4 (HC), 28.1 (HLL), 59.4 (FH), and 52.2 cm (HH). Ear length varied most (CV = 16 %) while wither height (CV = 4.8 %) varied the least. The descriptive statistics results found in this study agree with earlier reports by Yosef et al., (2014) and Tandoh et al. (2018).

The degree of linear association among body weight and morphometric traits measured by the Pearson correlation coefficient (r) and their statistical significance are presented in table 2. The correlation coefficients varied from 0.27 (between HW and FLL) to 0.92 (between BG and HG). All pairs of correlations were found positive and very highly significant ( $P < 0.0001$ ), indicating that the data is suitable for performing PC factor analysis. Such positive and very highly significant correlation coefficient values have also been reported in chickens by the studies of Yosef et al. (2014) and Gebremariam et al. (2014).

### PC Factor Analysis

Anti-image correlations computed showed that partial correlations were low, indicating that true factors existed in the data. This was further supported by the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy studied from the diagonal of partial correlation, revealing the proportion of the variance in the morphometric traits caused by the underlying factor. This was found to be sufficiently high with a value of 0.94. Eydurán et al. (2010) reported that a KMO measure of 0.60 and above is considered adequate. Bartlett's sphericity test for testing the null hypothesis that the correlation matrix is an identity matrix was used to verify the applicability of PCA. The value of Bartlett's sphericity test was significant ( $p$ -value = 0.001), implying that the PCA is applicable to the data set.

### Eigenvalues, Percentage of Total Variance with Rotated Component Matrix and Communalities

The eigenvalue of the total variance, the rotated component matrix and communalities of the traits investigated are presented in Table 3. Table 3 shows how much of the total variance of the observed traits is explained by each of the PCs after varimax rotation of the component matrix. Two PCs were identified with eigenvalues of 9.76 (PC1) and 1.04 (PC2). PC1 explained 57 % of the total variance while PC2 explained only 6 %. Accordingly, the first two PC factors combined accounted for 63 % of the total variability present in the parameters measured. The communalities are the proportion of variance that each variable has in common with other variables. Thus, if the communality of a trait is high, it means that the extracted factors explained a big proportion of the variance in the trait. The communality values ranged from 0.29 (HW) to 0.93 (HG) indicating that the data are conformable to PC factor analysis.

PC loadings presented in table 3 are the correlation coefficient between the first two PC scores and the original traits. They measure the importance of each LBM trait in accounting for the variability in the PC. That is, the larger the loadings in absolute terms; the more influential the variables are in forming the new PC and vice versa. The first factor (PC1) loaded heavily on EL, BL, FL, HLL, WH, HH, TL, FLL, HC, HL, NL, and FH while the second factor (PC2) loaded heavily on HG, and BG. The loading classification found in this study is somewhat similar to those reported by Uda (Yakubu et al., 2009), and immature Uda (Salako, 2006).

A scree-parallel analysis plot of eigenvalues against their PCs is shown in figure 2 below. The plot demonstrates the distribution of variance among the components graphically. For each PC, the corresponding eigenvalue is plotted on the y-axis. By definition, the variance of each component is less than the preceding one. Here there appears to be a marked decrease in downward slope after the second PC implying that we can summarize the nine morphometric traits by the first two PCs.

### Investigating Inter-Relationship Between Trait

A bi-plot display of PC1 and PC2 scores and PC loadings (Figure 3) is very useful in studying the relationships within observations, between traits, and the inter-relationship between observations and the traits. The X-Y axis of the bi-plot represents the standardized PC1 and PC2 scores, respectively.

From the above figure, one can see that all the traits are positive in PC1 and PC2. Traits with similar characteristics are displayed together in the bi-plot observational space since they have similar PC1 and PC2 scores. Similarly, traits with different attributes are displayed far apart since their PC scores are different. Since all the traits are positive in PC1 and PC2, those which constrain the system the most are EL, BL, FL, HLL, WH, HH, TL, FLL, HC, HL, NL, and FH (in PC1 axis), whereas those which constrain the system the most in PC2 axis are HG and BG.

## Conclusions

The use of independent orthogonal indices (PC1 and PC2) has proven to be more appropriate than the use of high patterns of original variables. In this study, PCA was explored in identifying patterns in seventeen measured original dimensions of camels to eliminate redundancy. The technique extracted two components. The first factor contributes effectively to explain the general body conformation of camels. The positive and significant correlation among different biometrical traits also makes them amenable for analysis. The two factors extracted from the present investigation could be used to select animals based on a group of traits rather than isolated traits. The study also revealed that factors extracted from the present investigation could be used in breeding programs with a sufficient reduction in the number of biometric traits to be recorded to explain body confirmation.

## Acknowledgement

The authors express their sincere gratitude to EIAR, Camel Project for its financial support and also thanks go to Camel owners for allowing their camels for experiment

## Authors' Contributions

Kefelegn Kebede; Conceptualization; Methodology; Formal analysis; Investigation, Writing - Original Draft, Supervision. Berhanu Bekele; Graduate fellow; Data curation; Writing - Original Draft. Sisay Tilahun: Conceptualization; Investigation; Supervision; Funding. Biresaw Serda; Conceptualization; Investigation; Supervision.

## References

- Abbas B, Al-Qarawi A, Al-Hawas A. 2000. Survey on camel husbandry in Qassim region, Saudi Arabia: herding strategies, productivity and mortality. *Magazine of Animal Husbandry and Veterinary Medicine of the Tropical Countries*, 53(3): 293- 298.
- Al-Dahash SA, Sassi MF. 2009. A preliminary study on management, breeding and reproductive performance of camel in Libya. *Iraqi Journal of Veterinary Sciences*, 23(2): 276-281.
- Bene S, Nagy B, Nagy I, Kiss B, Polgár JP, Szabó F. 2007. Comparison of body measurements of beef cows of different breeds. *Arch. Tierz. Dummerstorf*, 50: 363–373, 2007.
- Eyduran E, Topal M, Sonmez AY. 2010. Use of factor scores in multiple regression analysis for estimation of body weight by several body measurements in brown trouts (*Salmo trutta fario*). *International Journal of Agriculture and Biology*, 12: 611- 615.
- FAOSTAT, 2018. Food and agricultural organization of the United Nations, statistical division. <http://faostat3.fao.org/browse/Q/QA/E>. Last accessed on September 18, 2018.
- FAO, (Food and Agriculture Organization) 2012. Phenotypic Characterization of Phenotypic Resources. *Animal Production and Health Guideline No.11*. FAO, Rome.
- Gebremariam BH, Hailemicheal A, Gebru M. 2020. Breeding Practice and Correlation of Conformation Traits with Milk Offtake of Camel in Afar Region. *Research Square*; doi: 10.21203/rs.3.rs-32055/v1.
- Gifford-Gonzalez D, Hanotte O. 2011. Domesticating animals in Africa: implications of genetic and archaeological findings. *J. World Prehist.* 24:1-23.
- Ndumu DB, Baumung R, Hanotte O, Würzinger M, Okeyo MA, Jianlin H, Ibogo HK and Sölkner J. 2008. Genetic and morphological characterization of the Ankole Longhorn cattle in the African Great Lakes region. *Genet. Sel. Evol.*, 40: 467–490, doi:10.1186/1297-9686-40-5-467.
- SAS, (Statistical Analysis System) 2016. SAS for Windows. Release 9.4 SAS Institute, Inc., Cary, NC, USA.
- Salako AE. 2006. Application of morphological indices in the assessment of type and function in sheep. *International Journal of morphology*. 24(1):13-8.
- Tandoh G, Gwazal DS. Addass PA. 2018. Phenotypic Characterization of Camels (*Camelus dromedarius*) in Selected Herds of Katsina State. *Journal of Applied Life Sciences International* 18(3): 1-11; Article no. JALSI. 37787 ISSN: 2394-1103.
- Traoré A, Tamboura HH, Kabore A, Royo LJ, Fernández I, Álvarez I, Sangare M, Bouchel D, Poivey JP, Sawadogo L, Goyache F. 2008. Multivariate analyses on morphological traits of goats in Burkina Faso. *Archiv. Anim. Breed.*, 51: 588–600.
- Tura I, Kuria G, Walaga HK. Lesuper J. 2010. Camel Breeding Management among the Somali, Sakuye, Gabbra and Rendille Pastoralists of Northern Kenya. *Tropentag*, September 14-16, 2010, Zurich, Switzerland. Abstract.
- Yakubu A, Idahor KO, Agade YI. 2009. Using factor scores in a multiple linear regression model for predicting the carcass weight of broiler chickens using body measurements. *Revista UDO Agricola* 9(4): 963-967.
- Yohannes M, Mekuriaw Z, Getachew G. 2007. Camel and camel product marketing in Babilie and Kebribeyah woredas of the Jijiga Zone, Somali Region, Ethiopia. *Livestock Research for Rural Development*, 19:4.
- Yosef T, Kefelegn K, Mohammed K, Mengistu U, Solomon A, Tadelle D, Han J. 2014. Morphological diversities and eco-geographical structuring of Ethiopian camel (*Camelus dromedarius*) populations. *Emir. J. Food Agric.* 2014. 26(4): 371-389. doi: 10.9755/ejda.info/