Scientific Research Publishing

Genetic Diversity Analysis of Coffee (*Coffea arabica* L.) Germplasm Accessions Growing in the Southwestern Saudi Arabia Using Quantitative Traits

Taieb Tounekti[1][*], Mosbah Mahdhi[1], Turki Ali Al-Turki[2], Habib Khemira[1]

[1]Centre for Environmental Research and Studies, Jazan University, Jazan, Saudi Arabia
[2]Natural Resources and Environmental Research Institute, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia
Email: *tounekti_taiebb@yahoo.com

Abstract

Coffee (*Coffea arabica*) is an economically important crop cultivated on the terraced slopes and narrow valleys of the mountains in the south-western Saudi Arabia. The objective of the study was to evaluate genetic diversity among nineteen coffee accessions collected from different coffee growing areas based on quantitative morphological and pomological traits. Data of 17 quantitative traits were collected and subjected to various statistical analyses. The analysis of variance showed significant differences ($P < 0.05$) among the accessions for the majority of quantitative traits measured. A principal component analysis (PCA) and hierarchical cluster analysis (HCA) were used to determine the nature and degree of divergence among accessions. PCA showed that four principal components explained about 85.62% of the total variation. Differentiation of germplasm into different clusters was because of cumulative effects of a number of characters. Cluster and distance analysis of quantitative traits revealed the existence of five different groups. The maximum distance was seen between clusters IV and V (76.4) while the minimum was seen between I and II (16.13). Accessions like KSA6 and KSA8 from Tallen valley and KSA10 and KSA11 from Hada and Shada valleys collection are recommended for the next breeding work as they are high yielder accessions compared to the others. However, it is suggested to complete the present results by additional assessment of the *in situ* collection and variety trials. These results are of great interest for future breeding programs for this species.

Keywords

Coffea arabica, Cluster Analysis, Genetic Diversity, Principal Component Analysis

1. Introduction

Coffea, a member of the Rubiaceae family, has more than 70 species, but only two species, arabica (*Coffea arabica* L) and robusta (*Coffea canephora* Pierre), have commercial value. *C. arabica* is originated from Ethiopia and was propagated and dispersed all over the world from a low number of plants, which in addition to its autogamous reproduction led to a narrow genetic basis within Arabica coffee cultivars [1] [2]. In the southwestern tip of the Arabian Peninsula which encompasses southwestern Saudi Arabia and Yemen, *C. arabica* L. has been cultivated for the last 3 to 4 centuries [3] on the terraced slopes and narrow valleys of the Mountains at altitude ranging mostly from 1200 to 1800 m above sea level [4]. In south-western Saudi Arabia, trees of more than 100 years of age were found in traditional coffee growing areas in Jazan, Asir and Al-Baha regions. According to Fyfa Development Authority (FDA, governmental organism), the number of trees being about 78,000 with 84% was in Addayer district (Jazan region). These trees produce about 500 tonnes of coffee beans each year. Still the number of trees is most likely underestimated; since FDA operates a large nursery which offers annually thousands of seedlings for free for the local farmers (more than 40,000 coffee seedlings were distributed in 2016 alone). Therefore, Saudi Arabia continues to produce some of the best coffees in the world [5] [6], as coffee is grown under mostly organic practices without use of synthetic fertilizers, pesticides or herbicides and growers often use goat manure and stone mulching [7]. Regrettably, coffee genetic resources are under threat mostly due to the scarcity of water for irrigation and the erosion of the terraces and the spread of the cultivation of guat (*Catha edulis*) [7]. Besides, coffee farmers have replaced coffee by other higher priority monoculture crops and/or there has been a shift in coffee cultivation from productive to marginal sites [7]. To reduce such genetic erosion, efforts to collect and conserve Saudi coffee germplasm were carried out by the Centre for Environmental Research and Studies (CERS, Jazan University) in cooperation with Fyfa Development Authority (FDA) and about 20 germplasm accessions have been collected from different coffee growing areas of the southwestern Saudi Arabia and maintained at the experimental field of the FDA. In fact, our aim is to build gene pool by collecting coffee accessions of various geographical origins and centres of diversity and so to guarantee preservation of co-adapted gene complexes [8] [9]. Interestingly, the germplasm collected from the southwestern Saudi Arabia, where stressful conditions prevail, could provide a base material for developing cultivars tolerant to environmental stresses.

The economic value of a population is allied to morphology, agronomic performance, seed quality and nutritional qualities as well as on the efficient use of native germplasm [10]. Still, the convenient use of the potential held in germplasm collections needs detailed knowledge about the collections, including genetic diversity studies, evaluation and classification [9] [10]. Despite the appropriate environmental conditions, the productivity and production of coffee per unit area in the highlands of the southwestern Saudi Arabia remain very low

as compared to world average. This is mainly attributed to lack of adaptable cultivars for each ecological zone of the region. Previous reports have valued the extent of genetic diversity of Arabica coffee germplasm collected from different locality in Ethiopia [11] [12]. However, no systematic diversity analysis has been carried out to quantify and verify the level of genetic diversity of Arabica coffee in the southwestern Saudi Arabia. A good understanding of genetic variability is a useful tool in the genetic improvement of the crop. This will certainly boost the identification and use of interesting genes in breeding program. In fact, crop improvement through breeding depends on the disposal of genetic variability as well as on how easily this variability could be fixed in genotypes [13] [14].

In the southwestern Saudi Arabia, coffee has been for centuries propagated mainly from seeds, thus we expect that most of what growers here call cultivars are actually populations of a multitude of landraces with a degree of genetic variability. The relative isolation between coffee valleys could be considered as an outstanding occasion for the fixation of characters and varieties in these regions. Therefore the coffee gene pool might be unique to this region. It is, therefore, interesting to safeguard. Besides the only information we have concerning coffee accessions in the southwestern Saudi Arabia is that almost all Saudi and Yemeni coffee types arises from ancient, "heirloom" varieties of *Coffea arabica* first naturalized hundreds of years ago [4] [15] [16]. The characteristics of "Typica" variety seem to prevail over those of "bourbon" in these regions. Keeping this in view, this present study aimed to estimate the genetic diversity among coffee germplasm accession based on several quantitative traits. Combined multivariate analyses including the principal component analysis (PCA) and hierarchical cluster analysis (HCA) were employed to achieve such objective.

2. Material and Methods

2.1. Description of the Study Area and Plant Material

We carried out an-area-wide survey during 2015-2016 and examined several coffee accessions from 19 localities in the southwestern Saudi Arabia where the species was known to be cultivated (Figure 1). The main coffee producing valleys with old plantations visited were depicted in the Table 1. These sites have similar environmental conditions. In the most coffee farm we visited, most soils are sandy-loam, with slightly alkaline pH (7.5 to 8).

The Mountains of the southwest Saudi Arabia receive the highest rainfall in the country. The highest average annual rainfall recorded in this area includes 577 mm at Fyfa Mountain and 357 mm at Abha. Since the area falls within the boundaries of an arid region, rainfall is always inconsistent. Most of the rain is between April and September. Occasional showers may also fall during the cold months. Still the west facing slopes receive the maximum amount of rainfall, while in the eastern slopes, the intensity of the rain decreases and rains are unpredictable throughout the year. The climate of the mountains is characterized by cold winters and cool summer. Climatic data achieved from the Fyfa Meteorological Station during a period from 1984 to 1998 shows that the highest
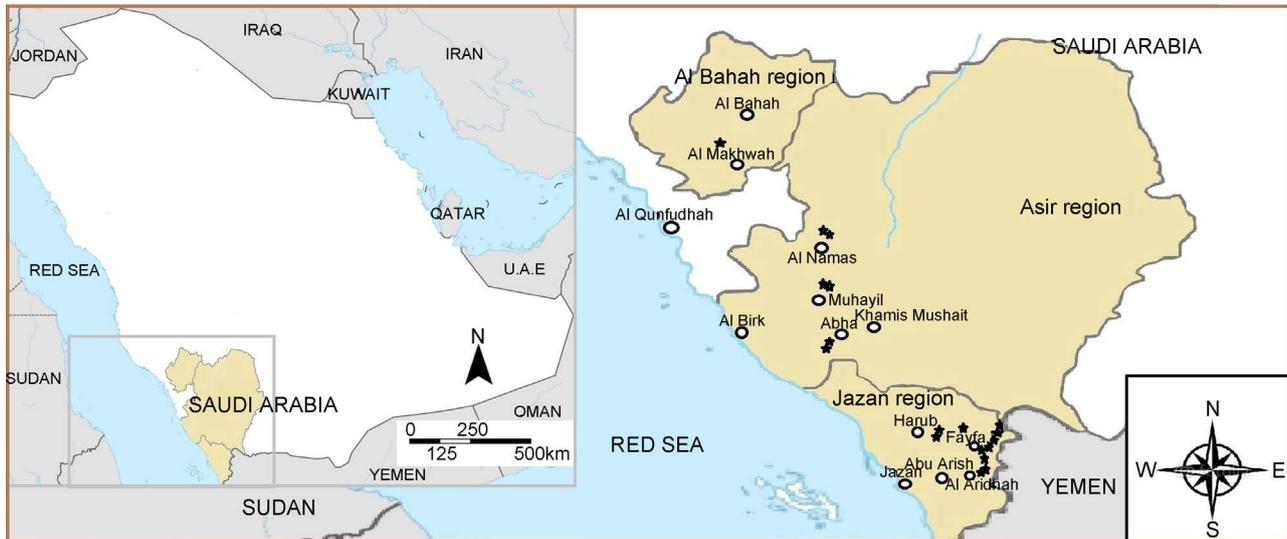
Figure 1. Map of the study area showing sampling sites in the south-western Saudi Arabia (Jazan, Asir and Al-Baha regions). The study areas are marked with black stars.

monthly maximum and the lowest monthly minimum temperature were recorded in June (33°C) and January (14°C) respectively. The mean maximal rainfall (220 mm) during the same period was recorded in April and the mean minimum in November. Mean relative humidity in summer varies from 60.9% in June to 79% in July, while in winter it ranges from 74.1% in December to 70.9% in March.

2.2. Data Collection

As many as possible the existing variability in several morphological and pomological quantitative traits was recorded (Table 2). Such traits concern the tree (plant height (m), and plant yield (kg)), leaves (length (cm), width (cm) and leaf area (cm²)), fruits (100 of fruits weight (g), length (mm), width (mm), and thickness (mm) and geometric diameter (mm)), and beans (50 beans weight (g), length (mm), width (mm), and thickness (mm), and percent seed defects (%)). Most samples were a mixture of fruits picked from more than one tree. The bulk of fruits samples were collected before January. The color of flushes and adult leaves was also used as a qualitative trait. The quantitative traits were complemented with more information provided by the farmers, such as the local names of the different coffee types and productivity. For the estimation of the productivity, a number of trees from each germplasm accession were harvested. Coffee production in each coffee farm was divided by the number of plants harvested per farm, thus the average production of a plant was gotten. Afterwards, the average production of each plant per farm was multiplied by the number of plants per hectare, which allowed obtaining the productivity, in kg·ha⁻¹, in each population.

2.3. Data Analysis

A digital Vernier caliper with 0.01 mm accuracy (Series 500, Mitutoyo, Japan)

Table 1. Geographical origin of the studied coffee (*Coffea arabica* L.) germplasm accessions used in the study. The main coffee producing valleys with old plantations in Jazan, Asir and Al-Baha regions were visited. Coffee plants were cultivated on the terraced slopes and narrow valleys of the Mountains at altitude ranging mostly from 1200 to 1800 m above sea level.

No	Accession	Region	District	Valley/place	Collection place	Altitude (m.a.s.l)
1	KSA1	Jazan	Addayer	Al-Zoughli	17°18'03.4"N 43°09'44.3"E	1254 m
2	KSA2	Jazan	Addayer	Al-Guatil	17°19'00.6"N 43°11'30.1"E	1484 m
3	KSA3	Jazan	Addayer	Al-Guatil	17°19'00.6"N 43°11'30.1"E	1484 m
4	KSA4	Jazan	Fyfa	Fyfa	17°15'20.6"N 43°06'21.1"E	1541 m
5	KSA5	Jazan	Addayer	Wadi Dafa	17°25'40.9"N 43°10'25.7"E	1254 m
6	KSA6	Jazan	Addayer	Tallan	17°23'11.5"N 43°09'47.8"E	1672 m
7	KSA7	Jazan	Addayer	Tallan	17°23'11.5"N 43°09'47.8"E	1672 m
8	KSA8	Jazan	Addayer	Tallan	17°23'00.9"N 43°09'47.3"E	1546 m
9	KSA9	Jazan	Addayer	Tallan	17°23'11.5"N 43°09'47.8"E	1672 m
10	KSA10	Asir	Mahayel Asir	Hada	18°38'43.8"N 42°09'10.4"E	1503 m
11	KSA11	Albaha	Al Makhwah	Shada	19°50'53.9"N 41°18'22.6"E	1548 m
12	KSA12	Jazan	Haroub	Maaddi	17°26'36.6"N 42°59'15.0"E	1287 m
13	KSA13	Jazan	Haroub	Maaddi	17°26'40.6"N 42°59'23.9"E	1344 m
14	KSA14	Jazan	Al-Aridha	Sala	17°03'21.7"N 43°08'25.1"E	1186 m
15	KSA15	Asir	Abha	Rayda	18°11'44.6"N 42°23'06.5"E	1594 m
16	KSA16	Asir	Abha	Rayda	18°11'44.6"N 42°23'06.5"E	1594 m
17	KSA17	Asir	Abha	Rayda	18°11'37.1"N 42°22'58.1"E	1519 m
18	KSA18	Asir	Al-Majarda	Wadi Ghill	19°09'35.3"N 42°04'18.8"E	1329 m
19	KSA19	Asir	Al-Majarda	Wadi Ghill	19°09'35.3"N 42°04'18.8"E	1329 m

Data of 18 quantitative traits were collected from each accession.

was used to measure the dimensions of the 100 samples of coffee berries and beans. The length (*L*), width (*W*) and thickness (*T*) measurements were taken. The geometric diameter (*FGD*) of the fruits was calculated as follow according to Mohsenin [17].

Table 2. Quantitative morphological and pomological traits used to characterize different accessions of *Coffea arabica*.

Trait	Description
PH	Plant height (m)
BL	Bean length (mm)
BW	Bean width (mm)
BT	Bean thickness (mm)
HBW	Hundreds of bean weight (g)
FL	Fruit length (mm)
FW	Fruit width (mm)
FT	Fruit thickness (mm)
FFW	Fifty fruits weight (g)
FGD	Fruit geometric diameter
FSI	Fruit shape index
LL	Leaf length (cm)
LW	Leaf width (cm)
LA	Leaf area (cm^2)
FC	New flush colour
BD	Beans defect (percent of elephant beans)
PR	Plant productivity

$$FGD = \left(LWT \right)^{1/3}$$

As well the fresh mass of 100 coffee fruits and 50 coffee beans were determined. The data collected were processed by different methods of statistical analysis using the SAS software [18]. Analysis of variance (ANOVA) was used to assess the significant differences by studied traits. Genetic divergence between clusters was determined using the generalized Mahalanobis D^2 statistics. Mahalanobis [19] developed this method to determine divergence prevailing among population in terms of generalized group distance [20]. The generalized distance between any two population is given by formula:

$D^2 p = \left(Xi - Xj \right)^1 S^{-1} \left(Xi - Xj \right)$ where, $D^2 p$ = total generalized distance based on p characters. Xi and Xj are the p mean vectors of accessions i and j, respectively. $S = p \times p$ pooled error variance-covariance matrix [19]. Testing the significance of D^2 values obtained for a pair of clusters were taken as the calculated value of Chi-square (χ^2) and tested against the tabulated value of χ^2 for p degrees of freedom (p is the number of characters) [21] at appropriate probability level that was considered.

In order to assess the extent of genetic variation and percentage similarity within the accessions and its nature, the collected data were then subjected to a multivariate analysis by PCA followed by HCA. PCA was achieved using correlation matrix in order to study the relationship among quantitative traits that are correlated among each other's by converting into uncorrelated traits called PCs [22]. The PCA produced Eigen-vectors and factor scores that were used respectively to

measure the relative discriminative power of the axes and their associated characters. The HCA constitutes a simple and less demanding model and is employed to summarize the position of accessions by sorting them into distinct groups [23]. Clustering was done using the proc cluster of SAS [18] by employing the method of average linkage clustering called Un-weighted pair group methods with arithmetic average (UPGMA).

3. Results

3.1. Analysis of Variance

Analysis of variance for the measured quantitative morphological and pomological traits reveal a significant ($P < 0.05$) differences among the coffee germplasm accessions for all traits (data not shown). Because of this all traits were included in multivariate analysis like cluster analysis and PCA.

3.2. Cluster Analysis

The 19 coffee germplasm accessions were classified into five distinct groups (Table 3, Figure 2). Cluster II had the highest number of accessions with eight accessions (44% of the total population) followed by cluster I with four accessions (22%). Clusters III had three accessions (17%), while IV and V had two

Table 3. Grouping of accessions of *the C. arabica* analyzed for 17 quantitative morphological and pomological traits and some qualitative traits, distributed in five groups.

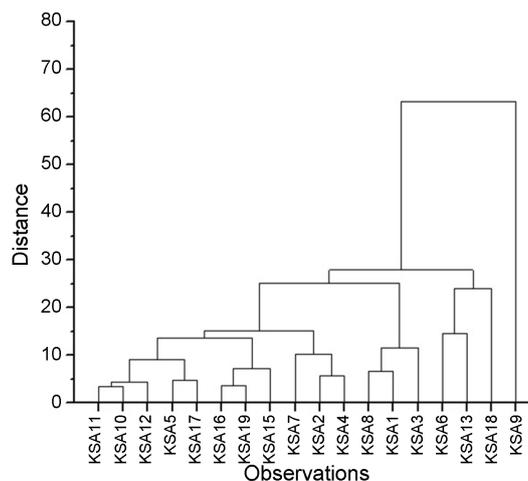| Cluster	Number of accession per group	Accessions
I	4	KSA16, KSA15, KSA19, KSA18
II	8	KSA11, KSA7, KSA10, KSA2, KSA4, KSA5, KSA12, KSA17
III	3	KSA8,KSA3, KSA1
IV	2	KSA6, KSA13
V	1	KSA9

Figure 2. UPGMA dendrogram depicting the genetic relationship of coffee germplasm based on 17 quantitative and qualitative characters evaluated in the field. The symbols used are indicated in the Table 1.

(11%) and one (6%) accessions, respectively. The cluster I includes the accessions KSA15 and KSA16 from Rayda valleys while KSA18 and KSA19 from Wadi Ghill. These accessions were characterized by their low to medium productivity. They have bold fruits with medium to large discs and the beans were of large size. This type contains low percentages of elephant beans (beans defect). The new leaf flushes were bronze to light bronze, and adult leaves were of average size. According to our observation and talk with the coffee growers, these accessions have a relatively flexible branches, resulting in many stem and heads for old plants. Some coffee growers named some accessions as "Tufahi". The Cluster II includes accessions KSA11, KSA7, KSA10, KSA2, KSA4, KSA5, KSA12 and KSA17. These accessions were collected from Shada, Tallan, Hada, Khacher, Fyfa, Wadi Dafa, Maadi, and Rayda valleys. Most of these accessions were of great vigour (more than 6 m height), moderate productivity and their compact cylindrical growth. The new leaf flushes were green to brown and adult leaves were green and of medium size shape. The fruits were bold with medium discs and large irregularly formed seeds, due to high percentage of elephant beans (beans defect) present in several representative of this cluster. The fruits and seeds were smaller than those of the cluster I. Some representatives of this type were named by local coffee growers as "Tessawi". When the spacing was not enough, the branches of different plants inter-cross making a dense canopy. The cluster III includes KSA8, KSA1, and KSA3. All accessions of this cluster are highly productive. This type was seen in different valleys such as Tallan and Khacher in Jazan region. Trees were of medium size (about 4 m high). This type has smaller and more rounded fruits. The beans were of medium size with elliptic rounded shape. Seeds contain only a little proportion of elephant beans. The pulp was very sweet. The shape of young and old trees with one stem is conical, with relatively open growth. All the representative of these cluster had a green leaf flushes, and mature leaves were of average size. The stems are of great flexibility. Some representatives of this type were named by some farmers as "Kholani". The Cluster IV includes the accessions KSA6 and KSA13. All accessions of this cluster are vigorous and very highly productive (the highest productive cluster). The trees were slightly taller than those of the cluster III (4 - 5 m height). It has a large elongated elliptic fruits, with small to medium sized discs. The beans were also large, regularly elongated elliptic with low fraction of elephant beans. New leaf flushes are brown to green, and adult leaves were large. The leaves were of narrow and pointed shape. The leaf surface is quite flat (this type has the highest leaf area). Some representatives of this type were named by some coffee growers "Dawairi or Dohaïri". Finally the cluster V includes only one accession KSA9. This cluster has relatively the lowest vigour (short stature) and productivity. This type was only found in Tallan valleys and with very limited number of trees. The tree shape was conical. Fruits and beans were small, elliptic rounded and very few defected beans (elephant beans) were present. Leaf flushes were light bronze, and old leaves were small and dark green. The most interesting character of this accession is it capacity to bears fruit all year round. This type is

named "Udaini" by some coffee growers.

The distance between the clusters of the 19 coffee accessions based on 17 quantitative traits is given in the Table 4. The inter cluster distance analysis showed significance ($P < 0.05$) differences between all cluster. The minimum distance was recorded between cluster I and II (16.13) flowed by cluster I and IV (20.23). The highest inter cluster distance was shown between cluster IV and V flowed by cluster I and V (75.7).

The dendrogram drawn from the HCA shows that all the accessions were distinct from each other at 100% level of similarity. At about 39% level of similarity, all the accessions had formed a single cluster except KSA9 which was distinct from all other accessions. At 72% level of similarity a new cluster was formed including the accessions KSA6, KSA13, and KSA18. At 75% level of similarity KSA3, KSA1 and KSA8 had joined to form one single cluster. At 85% level of similarity, three other accessions represented by KSA4, KSA2, and KSA7 have formed a new cluster and the other eight accessions (KSA11, KSA10, KSA12, KSA5, KSA17, KSA16, KSA19 and KSA15) have formed a separate cluster (the most heterogeneous cluster).

3.3. Principal Component Analysis

The principal component (PC) and percentage contribution of each component to the total variation in quantitative morphological and pomological traits were represented in the Table 5. The results showed that 4 PCs accounted for 85.62% of the total variation in the population. The relative discriminating power of the PCA as revealed by the Eigen values was high in PC1 (8.67) and lower in PC4 (1.45). The first component (PC1) accounted for 51.01% of the total variation. The highest contributions to the variations were shown by HBW (0.32) and FFW (0.32), followed by FGD (0.31). The FL (0.29), FW (0.30), FT (0.30), BL (0.30), BW (0.26) and BT (0.22) contribute greater to variations. As well the LL (0.27), LW (0.24) and LA contribute considerably in the PC1. All other characters contributed slightly to the PC1 (Table 5). The PC2 contributed by 15.0% of the total variation. Characters that contributed to this component include BD (0.53) and PH (0.38). However, LW (−0.36), LA (−0.33), PR (−0.32), and LL (−0.25) contributed negatively. The PC3 accounted for 11.1% of the total variation, and FC contributed the highest (0.59) flowed by PR (−0.48) and PH (−0.37). The FSI (0.66) contributed more to the variation in PC4. FW contributed negatively to this component (−0.28) (Figure 3).

Table 4. Distance between clusters of the coffee germplasm accessions.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 2	16.13			
Cluster 3	34.69	21.05		
Cluster 4	20.23	26.03	32.60	
Cluster 5	75.74	60.08	44.41	76.41

Table 5. Eigenvalue, factor scores and contribution of the first four principal component axes to variation in the coffee accessions.

| Parameter	Coefficients of PC1	Coefficients of PC2	Coefficients of PC3	Coefficients of PC4
FC	−0.08	0.10	0.59	0.02
BD	0.08	0.53	−0.15	0.26
PR	0.09	−0.32	−0.48	−0.13
PH	0.14	0.38	−0.37	0.27
HBW	0.32	0.05	0.07	−0.08
BL	0.30	−0.01	0.04	0.23
BW	0.26	−0.05	0.26	−0.07
BT	0.22	−0.13	0.26	−0.31
FFW	0.32	0.11	−0.02	−0.06
FL	0.29	0.08	0.15	0.21
FW	0.30	0.15	0.01	−0.28
FT	0.28	0.22	−0.04	−0.21
FGD	0.31	0.16	0.04	−0.09
FSI	0.05	−0.14	0.26	0.66
LL	0.27	−0.25	−0.05	0.06
LW	0.24	−0.36	−0.13	0.20
LA	0.25	−0.33	−0.06	0.16
Eigen value	8.67	2.55	1.87	1.45
Percentage	51.01%	15.01%	11.04%	8.50%
Cumulative	51.01%	66.02%	77.06%	85.62
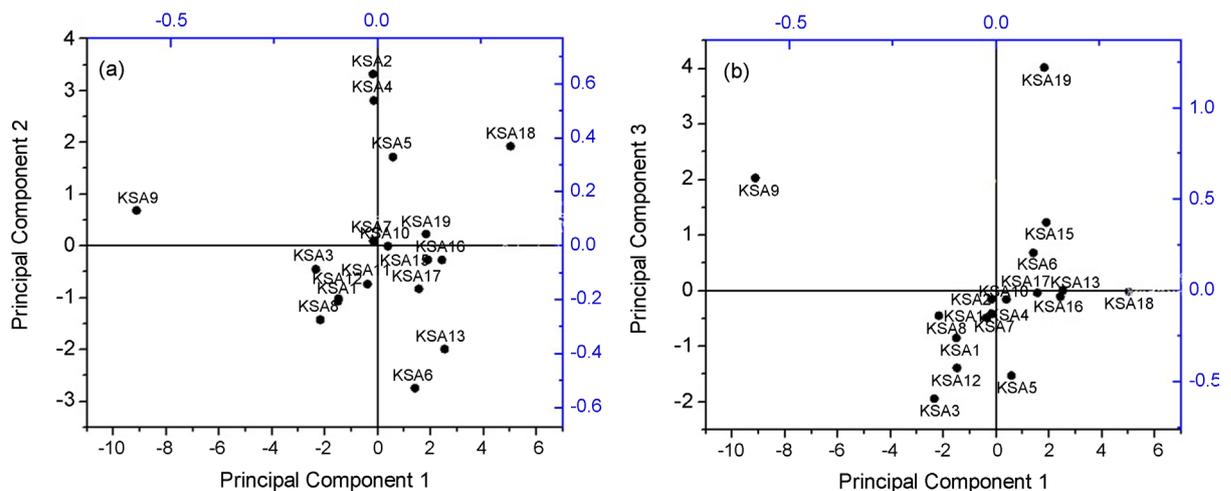
Figure 3. Plots of the first and second components (a and the first and third component (b) resulting from a PCA of the coffee accessions using 17 quantitative traits.

4. Discussion and Conclusions

Analysis of variance for the measured quantitative morphological and pomological traits reveals a significant ($P < 0.05$) differences among the coffee germplasm accessions for all traits. Therefore all traits were included in the multivariate

analysis like HCA and PCA. Such variability in several interesting traits clearly evidenced the option to bring significant perfection for coffee yield and/or coffee fruits and beans characteristics through selection and crossing. Variability between Saudi coffee accessions may be credited either to the evolutionary trends or to the natural mutation occurring to the population [24] [25]. The existence of board morphological diversity among Arabica coffee accessions was further confirmed by many authors who reported substantial morphological and agronomic variability among coffee germplasm accessions [11] [12] [26] [27] [28] [29]. Still, the major variations observed for most of morphological traits in a population were the result of the combination of the genotypic and environmental effects [30].

Cluster analysis based on coffee quantitative traits grouped 19 coffee genotypes into five distinct clusters (Table 3, Figure 2), in which the first cluster consisted of four accessions (22%). The second cluster consisted of eight accessions (44%), the third cluster consisted of three accessions (17%), the fourth cluster contained two accessions (11%) and the fifth cluster consisted of only one accession (6%) from the total accessions. This indicates that coffee accessions of the same cluster group were at least morphologically similar. The distribution pattern of genotypes in five clusters confirmed the existence of diversity among the coffee accessions. Interestingly, our results showed that several coffee accessions were clustered together despite they were collected from different geographic locations (sometimes very far from each other), as shown in the clusters I, II, III and IV. For example accessions collected from different place such Rayda and Wadi Ghill were clustered in cluster I. Likewise accessions collected from Shada, Tallan, Hada, Khacher, Fyfa, Wadi Dafa, Maadi, and Rayda valleys were also clustered together in cluster II. Cluster III includes accessions from Tallan and Khacher. Finally accession collected from Tallan and Maadi were grouped in cluster IV. This could be attributed to the unrestricted movement of the coffee seed from different valleys by man as well as wild animals [11] [31]. The same results were reported by Seyoum *et al.* [32] for coffee accessions from Ethiopia. Besides, previous report showed that morphological variation was more significant as an indicator of genetic diversity in coffee than variation in geographic origin [33]. Still our study reveal that coffee accessions collected from the same valley were clustered into different clusters which suggest a high genetic diversity within each collection area. The best example of this is Tallan valley where four different accessions were growing in the same place and under those same conditions and all of these accessions were clustered in four different clusters. This diversity could be exploited in order to increase the genetic base of coffee varieties.

The distance between the clusters of the 19 coffee accessions based on 17 quantitative traits is given in the Table 4. The inter cluster distance analysis showed significant ($P < 0.05$) differences between all clusters. The minimum distance was recorded between cluster I and II (16.13) flowed by cluster I and IV (20.23). The highest inter cluster distance was shown between cluster IV and V

flowed by cluster I and V (75.7). This confirms that the coffee accessions among clusters were divergent from each other. The lowest inter cluster distance observed indicated that the accessions in these clusters were not genetically diverse or there was little genetic diversity between them. However, a high inter cluster distance indicated that there is a high chance for obtaining transgressive segregates and maximizing heterosis by crossing accessions belonging to different clusters as there is a higher chance that distinct accessions would contribute by unique desired alleles at different loci [34] [35]. According to our results, the maximum recombination and segregation of the progenies is expected from crosses involving parents selected from cluster IV and V followed by cluster I and V and cluster II and V. Still, the breeder needs to specify his objectives in order to make best use of the characters where the characters are divergent. Falconer [36] reported that genetic diversity has probably arisen through diversity in origin (geographical separation), ancestral relationship, gene frequencies and morphology. These workers indicated that plants differing in either one or more of these factors would differ by a significant number of genes.

The present results showed that clusters III and IV consisted of 3 and 2 germplasm accessions respectively were the most interesting cluster regarding productivity and fruits and beans characteristics. The accessions grouped under the cluster IV gave the best characters of interest (highest productivity with big fruits and beans). This cluster is followed by the cluster III with also a very good productivity; while the fruits and beans were smaller than the cluster IV. The biggest fruits and beans were shown by the cluster I; however theses accessions were less reproductive (low to medium productivity). On the other hand, the cluster II consisted of the majority of the accessions (44%). This cluster was the most heterogeneous including accessions with moderate productivity and compact cylindrical growth. The fruits were bold with medium discs and large irregularly formed seeds, due to high percentage of elephant beans (beans defect) present in all representative of this type. However, cluster V consisted of only one accession (KSA9) has lowest productivity and fruits were of very small size. Still this accession has an interesting character which is its capacity to bears fruit all year round. Therefore, we can conclude that different clusters have different breeding value that enable breeders to improve different traits and potential selection should be made based on the relative merits of each cluster for each trait depending on the objectives of the breeding program.

The dendrogram drawn from HCA shows the relationship between the 19 accessions based on the 17 quantitative traits (Figure 2). This graph showed that all the accessions were distinct from each other at 100% level of similarity, however at 85% level of similarity, five different clusters were formed. The first cluster one include the accession KSA9, the second one include KSA6, KSA13, and KSA18, while the third cluster include KSA3, KSA1 and KSA8, the fourth cluster include KSA4, KSA2, and KSA7 and finally the fifth cluster include KSA11, KSA10, KSA12, KSA5, KSA17, KSA16, KSA19 and KSA15 (the most heterogeneous cluster). These results corroborate our above results.

Eigenvalues, percentage of total variance, percent of total cumulative variance and eigenvectors for 17 morphological and pomological quantitative traits in 19 Saudi coffee accessions were given in Table 5. PCA was performed to assess the relative importance of each trait for characterization of accessions (Figure 3). The results showed that about 85.62% of the variations present among accessions were explained by four principal components (PC). The first 2 principal components PC1 and PC2 with values of 51.0% and 15.0% respectively contributed more to total variations among coffee germplasm accessions. According to Chahal and Gosal [37], characters with largest absolute values closer to the unity within the first principal component (PC1) influence the clustering more than those with lower absolute values closer to zero. So, morphological and pomological traits mainly those contributed to the PC1 and PC2 played a major role in classifying coffee accessions into different clusters and should be considered in selecting diverse parents in crossing program. This finding is partly in agreement with the finding of Olika *et al.* [29] who have reported bean length, hundred bean weight, leaf length and leaf width contributed to the variation among Limmu coffee accessions. This finding is similar with the finding of Tikader *et al.* [38].

In conclusion, Saudi Arabia is endowed with immense potential of diverse coffee materials and suitable ecological conditions for coffee cultivation. The present study confirmed that coffee accessions showed variations for several morphological and pomological traits. Characterization of germplasm accessions based on 17 quantitative traits using the linkage method of Hierarchical cluster analysis resulted in grouping of the germplasm accessions into five clusters. Most of the inter-cluster distances were significantly different indicating the presence of variability which can be exploited through selection and hybridization. The PCA reveals that the first four components accounted for 85.62% of the total variation. The traits with the largest values in the first two PCs played major role in clustering the accessions and should be considered for selection program. Characters with high variability are expected to provide high level of gene transfer during breeding programs [39]. Interestingly, the germplasm collected from the southwestern Saudi Arabia, where stressful conditions prevail, could provide a base material for developing cultivars tolerant to environmental stresses.

Acknowledgements

References

[1] Steiger, D.L., Nagai, C., Moore, P.H., Morden, C.W., Osggod, F.V. and Ming, R.

(2002) AFLP Analysis of Genetic Diversity within and among *Coffea arabica* Varieties. *Theoretical and Applied Genetics*, **105**, 209-215.
https://doi.org/10.1007/s00122-002-0939-8

[2] Anthony, F., Quiros, O., Ropart, P., Bertrand, B. and Lashermes, P. (2002) Detection by Simple Sequence Repeat Markers of Introgression from *Coffea canephora* in *Coffea arabica* Varieties. *Plant Breeding*, **121**, 542-544.
https://doi.org/10.1046/j.1439-0523.2002.00748.x

[3] Eskes, A.B. (1989) Identification, Description and Collection of Coffee Types in P. D. R. Yemen. IPGRI, Rome.

[4] Al-Zaidi, A.A., Baig, M.B., Shalaby, M.Y. and Hazber, A. (2016) Level of Knowledge and Its Application by Coffee Farmers in the Udeen Area, Governorate of IBB—Republic of Yemen. *The Journal of Animal and Plant Sciences*, **26**, 1797-1804.

[5] Lewin, B., Giovannucci, D. and Varangis, P. (2004) Coffee Markets: New Paradigms in Global Supply and Demand. The International Bank for Reconstruction and Development, Agriculture and Rural Development Discussion Paper 3.

[6] FAO. (2015) FAO Statistical Pocketbook Coffee 2015.
http://www.fao.org/3/a-i4985e.pdf

[7] Al-Turki, T.A. (2002) An Initiative in Exploration and Management of Plant Genetic Diversity in Saudi Arabia. In: Engels, J.M., Rao, V.R., Brown, A.H. and Jackson, M.T., Eds., *Managing Plant Genetic Diversity*, CAB International, Wallingford, 339-349. https://doi.org/10.1079/9780851995229.0339

[8] Ghafoor, A., Ahmad, Z., Hashmi, N.I. and Bashir, M. (2003) Genetic Diversity Based on Agronomic Traits and SDS-PAGE Markers in Relation to Geographic Pattern of Blackgram [*Vigna mungo* (L.) Hepper]. *Journal of Genetics and Breeding*, **57**, 5-14.

[9] Beuselinck, P.R. and Steiner, J.J. (1992) A Proposed Framework for Identifying and Utilizing Plant Germplasm Resources. *Field Crop Research*, **29**, 261-272.

[10] Beer, S.C., Goffreda, J., Phillips, T.D., Murphy, J.P. and Sorrells, M.E. (1993) Assessment of Genetic Variation in *Arena sterilis* Using Morphological Traits, Isozymes, and RFLPs. *Crop Science*, **33**, 1386-1393.

[11] Yigzow, D. (2005) Assessment of Genetic Diversity of Ethiopian Arabica Coffee Genotypes Using Morphological, Biochemical and Molecular Markers. Ph.D. Dissertation, University of the Free State, South Africa.

[12] Kebede, M. and Bellachew, B. (2008) Phenotyoic Diversity in the Hararge Coffee (*Coffee arabica* L) Germplasm for Quantitative Traits. *East African Journal of Sciences*, **2**, 13-18.

[13] Ariyo, O.J. (1990) Measurement and Classification of Genetic Diversity in Okra (*Abelmoschus esculentus*). *Annals of Applied Biology*, **116**, 335-341.
https://doi.org/10.1111/j.1744-7348.1990.tb06615.x

[14] Kiran Patro, T.S. and Ravisankar, C. (2004) Genetic Variability and Multivariate Analysis in Okra [*Abelmoschus esculentus* (L.) Moench]. *Tropical Agricultural Research*, **16**, 99-113.

[15] Eskes, A.B. and Mukred, A.W.O. (1989) Coffee Survey in PDR Yemen. ASIC, 13, Colloque, Piape, 582-590.

[16] Al Hakimi, A.A.S. and Allard, B. (2005) Collection, Characterization and Evaluation of Yemeni Landraces of Coffee. *Zagazig Journal of Agricultural Research*, **32**, 1-12.

[17] Mohsenin, N.N. (1986) Physical Properties of Plant and Animal Materials. Taylor and Francis, New York.

[18] SAS (1999) Statistical Analysis System, Statistical Methods. SAS Institute Inc., Cary,

NC.

[19] Mahalanobis, P.C. (1936) On the Generalized Distance in Statistics. *Journal of Genetics*, **41**, 159-193.

[20] Sharma, J.R. (1998) Statistical and Biometrical Techniques in Plant Breeding. New Age International (P) Ltd., New Delhi, 178-203.

[21] Singh, R.K. and Chaudhary, B.D. (1987) Biometrical Methods in Quantitative Genetic Analysis. Kalyani Publishers, New Delhi, Ludhiana, India, 318.

[22] Johnson, R.A. and Wichern, D.W. (1988) Applied Multivariate Statistical Analysis. 2nd Edition, John Wiley & Sons Inc., New York.

[23] Ariyo, O.J. and Odulaja, A. (1991) Numerical Analysis of Variation among Accessions of Okra (*A. esculentus* L. Moench). Malvaceae. *Annals of Botany*, **67**, 527-531.

[24] Avise, J.C. and Hamrick, J.L. (1997) Conservation Genetics: Case Histories from Nature. Chapman & Hall, New York.

[25] Hedrick, P.W. (2000) Genetics of Populations. 2nd Edition, Jones and Bartlett Publishers, Sudbury, MA.

[26] Bayetta, B. (1997) Arabica Coffee Breeding in Ethiopia: A Review. ASIC, Nairobi, Kenya, 17, 406-414.

[27] Ermias, H. (2005) Evaluation of Wellega Coffee Germplasm for Yield, Yield Component and Resistant to Coffee Berry Disease at Early Bearing Stage. Ph.D. Dissertation, School of Graduate Studies of Alemaya University.

[28] Gichimu, B.M. and Omondi, C.O. (2010) Morphological Characterization of Five Newly Developed Lines of Arabica Coffee as Compared to Commercial Cultivars in Kenya. *International Journal of Plant Breeding and Genetics*, **4**, 238-246. https://doi.org/10.3923/ijpbg.2010.238.246

[29] Olika, K., Sentayehu, A., Taye, K. and Weyessa, G. (2011) Variability of Quantitative Traits in Limmu Coffee (*Coffea arabica* L.) in Ethiopia. *International Journal of Agricultural Research*, **6**, 482-493. https://doi.org/10.3923/ijar.2011.482.493

[30] Welsh, R.J. (1990) Fundamentals of Plant Genetics and Breeding. John Wiley & Sons, New York.

[31] Esayas, A. (2005) Molecular Genetic Diversity Study of Forest Coffee Tree (*Coffea arabica* L.) Populations in Ethiopia: Implications for Conservation and Breeding. Ph.D. Dissertation, Swedish University of Agricultural Sciences.

[32] Seyoum, S. (2003) Genetic Divergence for Seedling Parameters and Associations among Agronomic Traits in the Ethiopian Coffee (*Coffea arabica* L.) Germplasm. Ph.D. Dissertation, School of Graduate Studies of Alemaya University.

[33] Bayetta, B. (2001) Arabica Coffee Breeding for Yield and Resistance to Coffee Berry Disease (*Colletotichum kahawae* sp.nor). Ph.D. Thesis, Imperial College of Wye, University of London.

[34] Ghaderi, A., Adams, M.W. and Nassib, A.M. (1884) Relationship between Genetic Distance and Heterosis for Yield and Morphological Traits in Dry Edible Bean and Faba Bean. *Crop Science*, **24**, 37-42.

[35] Peeters, J.P. and Martinelli, J.A. (1989) Hierarchical Clustering Analysis as a Tool to Manage Variation in Germplasm Collections. *Theoretical and Applied Genetics*, **78**, 42-48. https://doi.org/10.1007/BF00299751

[36] Falconer, D.S. (1981) Introduction to Quantitative Genetics. 2nd Edition, John Wiley & Sons, Inc., New York, UK.

[37] Chahal, G.S. and Gosal, S.S. (2002) Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches. Alpha Science International, UK,

604.

[38] Tikader, A., Reo, A.A., Ravindran, S., Naik, V.G., Mukherjee, P. and Thangavelu, K. (1999) Divergence Analysis in Different Mulberry Species. *Indian Journal of Genetics and Plant Breeding*, **59**, 62-88.

[39] Gana, A.S., Shaba, S.Z. and Tsado, E.K. (2013) Principal Component Analysis of Morphological Traits in Thirty-Nine Accessions of Rice (*Oryza sativa* L.) Grown in a Rainfed Lowland Ecology of Nigeria. *Journal of Plant Breeding and Crop Science*, **5**, 120-126.