Plant Genetic Resources: Characterization and Utilization

cambridge.org/pgr

Phenotypic characterization of Amaro coffee (*Coffea arabica* L.) local accessions using multi-variate techniques at Awada, Southern Ethiopia

Desalegn Alemayehu[1], Weyessa Garedew[2] and Abush Tesfaye Abebe[3]

[1]Jimma Agricultural Research Center, Jimma, Ethiopia; [2]Jimma University, College of Agriculture and Veterinary Medicine, Jimma, Ethiopia and [3]International Institute of Tropical Agriculture, Ibadan, Nigeria

Research Article

Cite this article: Alemayehu D, Garedew W, Tesfaye Abebe A (2022). Phenotypic characterization of Amaro coffee (*Coffea arabica* L.) local accessions using multi-variate techniques at Awada, Southern Ethiopia. *Plant Genetic Resources: Characterization and Utilization* 1–9. https://doi.org/10.1017/S1479262122000119

Received: 19 June 2021
Revised: 12 April 2022
Accepted: 20 April 2022

Key words:
Clustering; *Coffee arabica*; phenotypic characterization; principal component; Shannon–Weaver index

Author for correspondence:
Abush Tesfaye Abebe, E-mail: At.Abebe@cgiar.org

Abstract

As a country of origin of coffee, Ethiopia is endowed with an immense diversity of the crop in its diverse coffee-growing agro-ecologies. Amaro Kelo is one of the major coffee production agro-ecologies in Ethiopia, where the genetic diversity of its landrace coffee germplasm was not properly characterized previously. The study aimed to characterize 64 Amaro Kelo local coffee accessions to understand the potential of the accessions for utilization in future coffee genetic improvement efforts. The experiment was laid out in an 8×8 simple lattice design with two replications at Awada Agricultural Research Sub-Center. Data were collected on 19 quantitative and 10 qualitative traits, and subjected to multivariate analyses, i.e. cluster and principal component analyses. The cluster analysis identified five clusters based on the quantitative characters, and the distances between most of the clusters were highly significant at $P < 0.01$. Principal component analysis revealed the first six principal components with Eigenvalues greater than one accounted for 77.7% of the total variation. The first two principal components with respective contributions of 23.32 and 18.85% cumulatively accounted for 42.2% of the total variation in the accessions. In addition, high values of Shannon-diversity index were found for the qualitative traits: branching habit, growth habit, fruit shape, overall appearance and stem habit. In general, the multivariate analyses confirmed the presence of high variation among the studied Amaro-Kelo coffee accessions that might serve as an important genetic resource for future coffee genetic improvement or conservation efforts.

Introduction

Arabica coffee (*Coffea arabica* L.) is predominantly grown in the tropical and subtropical regions of the world (Berthaud and Charrier, 1988; Davis *et al.*, 2012). It is the only allotetraploid species of the genus ($2n = 4x = 44$) (Lashermes *et al.*, 1999; Benti *et al.*, 2020). Coffee is one of the most important beverages globally (Labouisse *et al.*, 2008), ranking second after oil in international trade (Geromel *et al.*, 2006; Hameed *et al.*, 2018). More than 2.5 billion cups of coffee are consumed daily across the globe (Coffee Industry Statistics, 2020), with 1.8–2.0% annual growth in consumption globally (Hameed *et al.*, 2018; ICO, 2020), and 17% yearly growth of global green bean production and consumption in the past decade (Hameed *et al.*, 2018).

Ethiopia is one of Africa's most important Arabica coffee-producing and exporting countries and the sixth most important coffee producer worldwide (ICO, 2018; USDA, 2018). The crop is mainly produced in the southern, southwestern and eastern parts of the country. The total area coverage of coffee in Ethiopia is estimated to be 758,523.29 ha, of which about 95% is produced by 4 million smallholder farmers (CSA, 2020). The estimated annual national production of coffee was 8.04 million of 60 kg bags, while the average national productivity was about 636 kg/ha (CSA, 2020).

The plateaus of central Ethiopia are reported as the origin of Arabica coffee from a relatively recent hybridization between Robusta coffee (*Coffea canephora* Pierre ex A. Froehner) and *C. eugenioides* S. Moore or their ecotypes (Tesfaye *et al.*, 2007; Lashermes *et al.*, 2011). Arabica coffee has its primary centre of origin and diversity in the highlands of southwestern Ethiopia (Charrier and Berthaud, 1985; Wrigley, 1988; Belachew, 2000). It is the species naturally occurring in Ethiopia and South Sudan (Davis *et al.*, 2012). As the centre of origin of Arabica coffee, Ethiopia is endowed with rich genetic resources of coffee in its diverse coffee-growing agro-ecologies (Hindorf and Omondi, 2011; Benti *et al.*, 2020). Amaro Kelo, where the local coffee accessions used in this study were collected, is one of the major coffee production agro-ecologies with high diversity of the crop in the country.

Agro-morphological traits, such as size and shape of leaf and plant form, colour of the shoot tip, characteristics of the fruit, angle of branching and length of internodes, are vital

to explore variations in coffee (De Vienne *et al.*, 2003). Biotic stress tolerance traits such as coffee leaf rust (CLR) caused by *Hemileia vastatrix* and coffee berry disease (CBD) caused by *Colletotrichum kahawae* are very significant diseases of coffee in Ethiopia and other major coffee-producing countries (Hindorf and Omondi, 2011; Daba *et al.*, 2019) that need to be given high priority in coffee improvement programmes. Hence, characterizing Amaro Kelo accessions based on these important diseases may provide important implications on the potential of the accessions to utilize for resistance breeding for these diseases. Multivariate analyses based on agro-morphological traits have been widely used to characterize and estimate the diversity of several crops, including coffee (Gichimu and Omondi, 2010; Olika *et al.*, 2011; Gessese *et al.*, 2015). These multivariate techniques have also been used to characterize the genetic diversity of different Arabica coffee accessions of Ethiopia, such as Harerge (Adem, 2009) and Limmu (Olika *et al.*, 2011). However, such characterization and diversity studies have not been carried out on Amaro Kelo local coffee accessions using agro-morphological traits, and information on the diversity of Amaro Kelo coffee accessions is lacking. Hence, the objective of this study was to characterize and estimate the extent of genetic diversity of Amaro Kelo coffee accessions based on 17 quantitative and 10 qualitative traits using multivariate techniques that might help identify superior coffee genotypes for direct release as a variety or use as a parental line for future coffee hybridization programmes after subjecting some of the best performing genotypes to multilocations and seasons evaluation.

Materials and methods

Experimental design and plant materials

A total of 58 local coffee accessions collected from 10 Peasant Associations (PAs) (the lowest Government Administrative structure in the rural areas) representing the major coffee-growing agro-ecologies of the Amaro woreda of Segen zone of the Southern Nations and Nationalities and Peoples (SNNP) Regional State in Ethiopia in the year 2013 (online Supplementary Table S1) were used for the study along with six released pure line varieties, i.e. Angafa (1377), Feyate (971), Koti (85,257), and Odicha (974), 74112 and 7440 were included as reference (standard check) varieties to compare the performance of the studied accessions. The trial was laid out in an 8×8 simple lattice design and planted in July 2014. This study was conducted in October 2018 by collecting data on the studied traits on a 4-year-old established trial. Each plot consisted of six coffee trees planted in a spacing of 2×2 m for both between rows and plants in a single row. All the experimental plots were established under uniform *Sesbania sesban* temporary shade trees, and all the other coffee management practices such as pruning and slashing were also uniformly applied to all the plots, as per the recommended coffee agronomic practices (Tesfaye *et al.*, 2006; Alemseged *et al.*, 2015).

Description of the experimental site

The experiment was carried out at Awada Agricultural Research Sub-Center located at Yirgalem town, 45 km south of Hawassa and 319 km south of the capital city Addis Ababa in Ethiopia. The sub-centre is located at 06°44′57″N latitude and 038° 23′16″E longitude, and an altitude of 1738 masl. The mean annual rainfall of the area is 1342 mm, with an average maximum and minimum temperatures of 28.4 and 11 °C, respectively.

Methods of data collection

Data on 19 agro-morphological (quantitative traits) and an additional 10 qualitative traits were collected on four sample trees of 4-year-old trees per row on each plot, as per the standard IPGRI (1996) coffee descriptors (list of the traits and their description is presented in online Supplementary Table S2).

Data collection on incidence of coffee diseases

The disease incidence (%) of CBD and CLR were also recorded on four sample trees. Visual disease estimation per individual sample coffee tree was performed as described by Van der Graaff (1981). Each of the four sample trees per accessions was randomly taken and first diagnosed for the presence or absence of the disease, then the disease incidence was calculated using the formula below:

$$\text{Incidence of coffee berry/leaf rust disease (\%)} = \frac{\text{Total number of diseased trees}}{\text{Total number of observed trees}} \times 100$$

Statistical analysis

Cluster analysis

Cluster analysis was performed using the proc cluster procedure of SAS software (SAS, 2013), employing the method of average linkage clustering strategy of the observations. The number of clusters was determined following the approach suggested by Copper and Milligan (1988), considering the three statistics: Pseudo F, Pseudo t^2 and Cubic Clustering Criteria (CCC) using SAS 9.4 software (SAS, 2013). Genetic divergence between clusters was determined using the generalized Mahalanobis's D^2 statistics (Mahalanobis, 1936). The D^2 value obtained for pairs of clusters were considered as the calculated value of χ^2 and were tested for significance at 5 and 1% level of probability against the tabulated values of χ^2 for P degrees of freedom, where P is the number of characters considered (Singh and Chaudhary, 1987).

Principal component analysis (PCA)

The PCA was performed based on correlation matrix using SAS version 9.3 (SAS, 2013). Principal components with Eigenvalue greater than one were considered important in explaining the total variations in the studied accessions (Cliff, 1988).

Shannon diversity index (H′)

Shannon diversity index (H') was calculated using 10 qualitative traits to analyse the phenotypic diversity of the studied coffee accessions. It was calculated using the formula:

$$H' = \sum_{i=1}^{s} p_i \ln(p_i) E_H = H/H_{\max} = H/\ln S,$$

where S is the number of traits category, E_H is Shannon's equitability, H is Shannon diversity index, $H_{\max}$ is maximum diversity possible and p_i is the relative proportion of the total number of entries (N) in the ith class (Shannon, 1948).

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press

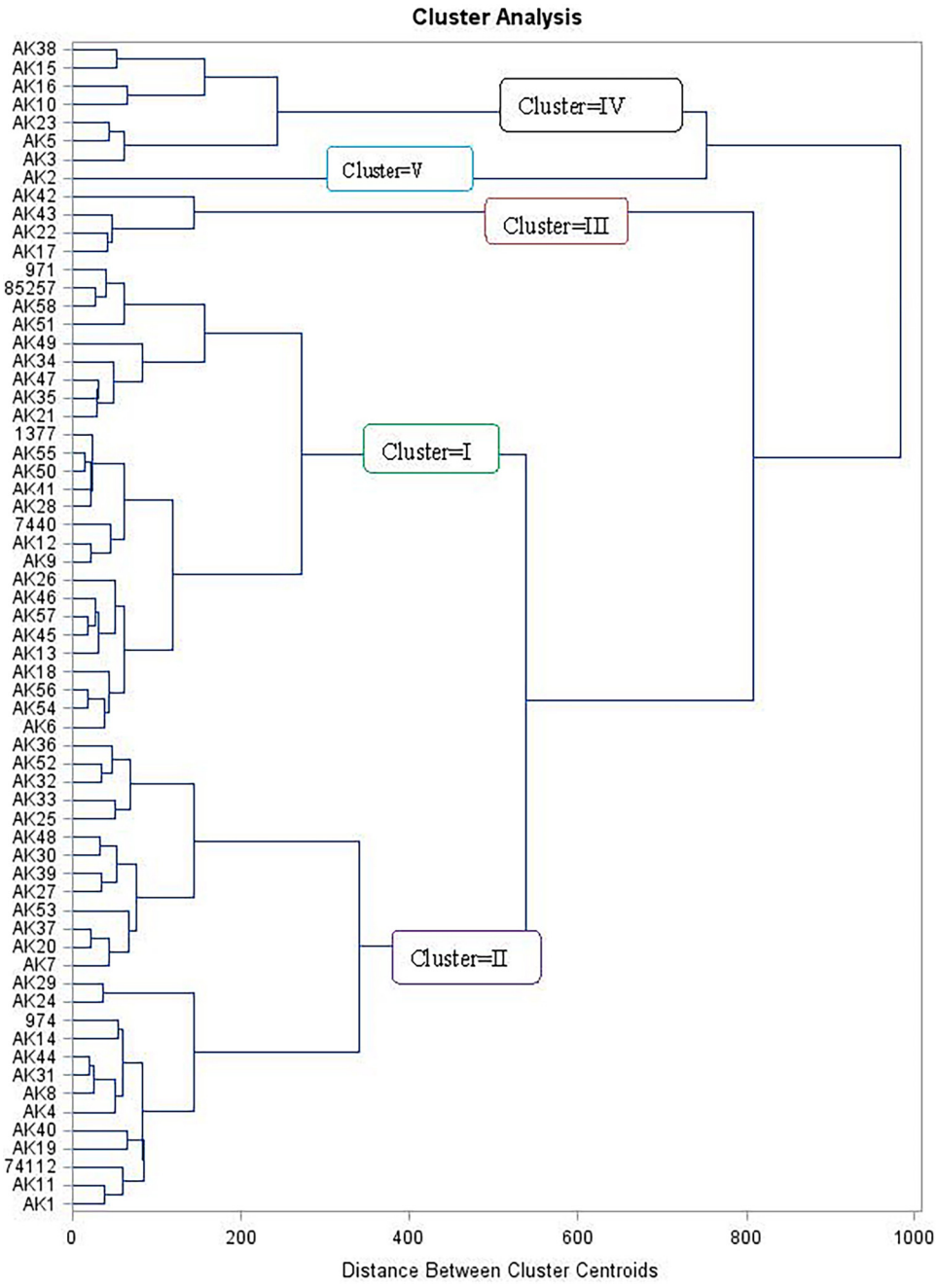

Fig. 1. Dendrogram for the association of 64 coffee genotypes based on 19 quantitative coffee traits.

Results

Cluster analysis of genotypes based on quantitative traits

The phenotypic similarities of 64 coffee accessions were analysed using cluster analysis based on 19 quantitative characters. Cluster analysis confirmed the presence of variation among the local Amaro coffee accessions. The 64 coffee accessions were grouped into five clusters (Fig. 1). The genotypes used as checks, Angafa, 7440, Feyate and Koti were grouped into cluster-I; whereas genotypes 74112 and Odicha were grouped into cluster-II. The majority of accessions 59 (92.20%) were classified into three clusters, with 26, 26 and 7 accessions in clusters I, II and IV; while clusters III and V had four and one genotypes, respectively.

Cluster analysis based on qualitative characters

Cluster analysis has also confirmed variations among the 64 coffee accessions for the qualitative traits, and accordingly, Amaro Kelo coffee accessions were grouped into six distinct groups (Table 1). Cluster-I was the largest and consisted of 33 accessions (51.6%), followed by cluster-II (31.3%), cluster-III (7.8%), cluster IV (4.7%), V (3.1%) and cluster VI (1.6%). Accessions grouped into cluster-I predominately have triangular stipule and lanceolate

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press

Table 1. Clustering patterns of 64 coffee accessions based on 10 qualitative characters

Cluster no.	Number of accessions	Per cent	Accessions
I	33	51.6	AK15, AK32, AK50, AK49, AK4, AK5, AK1, AK13, AK7, AK34, AK42, AK14, AK52, 974, AK27, AK20, AK23, AK3, AK54, K56, 74112, AK31, AK26, AK6, AK39, AK38, AK17, AK12, AK2, K36, AK28, AK48, 7440
II	20	31.3	AK46, AK30, AK47, AK57, AK22, AK44, AK41, AK25, AK21, AK24, AK43, AK51, 971, AK53, AK19, AK55, AK8, AK58, AK11, AK10
III	5	7.8	AK16, AK9, AK45, AK29, AK35
IV	3	4.7	AK33, 1377, AK37
V	2	3.1	85257, AK18
VI	1	1.6	AK40

leaf shape. Cluster II, on the other hand, comprised of 20 accessions that were characterized by lanceolate leaf shape. In addition, five accessions were grouped under cluster III, and these accessions predominantly possess typically open growth habit with spreading branch, flexible stem, bronze young leaf tip colour and ovate stipule shape. Cluster IV was characterized by apiculate leaf apex shape, flexible stem, deltate stipule shape, elliptic fruit shape and many branches (primary) with few secondary branches. Only one accession was grouped into cluster VI (AK40), characterized by open growth habit, strong stem, elliptic fruit shape, many branches (primary) with many secondary branches, ovate leaf shape, spreading angle of insertion on the primary branch and bushy type in overall appearance. Cluster V had compact growth habit, semi-erect insertion on primary branch, flexible stem, acuminate leaf shape, apiculate leaf apex and pyramidal in overall appearances.

Cluster means analysis

The mean value of the 19 quantitative characters in each cluster is presented in Table 2. The cluster means analysis revealed high variation for various characters. A closer look into the cluster means revealed that the clusters differ for the different characters measured. Cluster V consisted of one accession having the characteristic feature of relatively shorter average internodes length on the main stem (4.4 cm), a longer length of the longest primary branches (118.5), the lowest incidence of CLR (9.7%), short bean length (9.3 mm), short bean width (6.2 mm), short bean thickness (3.6 mm), short fruit width (10.7 mm) and light 100 bean weight (12.8 g). Moreover, this cluster is characterized by a high yield (2474.5 kg/ha), shorter plant height (210 cm), relatively longer height up to first primary branch (24.2 cm) and wider stem diameter (5.9 cm), as compared to other clusters.

Cluster IV, which contained seven accessions, was known for its unique characteristics of having high number of primary branches (71), high number of bearing primary branch (53), relatively wider bean thickness (3.7 mm), high coffee bean yield (1722.1 kg/ha), shorter height up to first primary branch (23.1 cm) and the lowest incidence of CBD (11.72%). Cluster III that consisted of four accessions was characterized by a low number of primary branches (57), a low number of bearing primary branch (40), the highest incidence of both CBD (61.9%) and CLR (16.9%), wide leaf area (44.3 cm^2), narrow bean thickness (3.6 mm), narrow fruit length (14.8 mm), shorter stem diameter (4.6 cm), lower coffee bean yield (86.0 kg/ha) and wider fruit width (11.3 mm). Cluster II was characterized by shorter length of the longest primary branch (112.7 cm), narrow bean thickness (3.7 mm), narrow leaf width (5.2 cm) and narrow leaf size (41.6). Accessions in cluster I consisted of about 26 coffee accessions with longer average internodes length on the main stem (5.0 cm), tall plant height (232.0 cm), heavier 100-bean weight (15.2 g), longer bean length (10.0 mm), wider bean width (6.5 mm), long fruit length (15.6 mm) and wide fruit width (11.3 mm). However, it is important to note in calculating cluster means that the superiority of a particular genotype for a given character might be diluted by other genotypes related and grouped into the same cluster that are inferior or intermediary for that character under consideration.

Inter-cluster distance (D^2) analysis based on quantitative traits

The χ^2 test for the five clusters indicated that there were highly significant differences ($P < 0.01$) among each other (Table 3). The smallest inter-cluster distance ($D^2 = 23.62$) was observed between clusters I and IV; while the highest ($D^2 = 371.29$) was found between clusters III and V. In most cases, the accessions among the clusters are significantly ($P < 0.01$, $x^2 = 34.80$) divergent from each other.

Inter-cluster distance (D^2) analysis based on qualitative characters

All the cluster distances based on the qualitative traits showed a highly significant ($P < 0.01$) difference among each other (online Supplementary Table S3). The highest inter-cluster distance ($D^2 = 254.80$) was found between clusters II and VI, followed by between clusters V and VI ($D^2 = 212.05$), I and VI ($D^2 = 188.70$) and clusters III and VI ($D^2 = 162.25$).

Principal component analysis

The relative contribution of each attributes to the observed variability was evaluated based on PCA, and the result is presented in Table 4. The first six principal components (PCs), i.e. PC1, PC2, PC3, PC4, PC5 and PC6 with Eigenvalues greater than one (4.431, 3.581, 2.511, 1.710, 1.424 and 1.110, respectively), accounted for 77.7% of the total variation were considered important for reporting in this study. The first principal component (PC1) explained 23.32% of the total variation, followed by PC2 (18.85%), PC3 (13.22%), PC4 (8.98%), PC5 (7.49%) and PC6 (5.84%). Almost all the studied traits, except stem diameter and CLR disease, showed high (Eigenvector value of >0.3)

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press

Table 2. Cluster mean value of 19 quantitative characters for five clusters of Amaro coffee accessions at Awada

	Cluster means					Cluster mean differences				
Variable	I	II	III	IV	V	I	II	III	IV	V
YLD	1161.40	623.87	86.00**	1722.14	2474.50**	203.74	−333.80	−871.66	764.48	1516.84
PH	232.02**	229.17	226.55	227.08	210.00*	2.39	−0.46	−3.08	−2.55	−19.63
HUFPB	23.76	23.32	24.00	23.14*	24.15**	0.22	−0.22	0.46	−0.40	0.61
SD	5.16	4.97	4.64*	5.31	5.85**	0.08	−0.11	−0.44	0.23	0.77
CD	207.32	199.28	194.26*	210.25	211.45**	3.70	−4.35	−9.36	6.63	7.83
AINL	4.97**	14.82	4.88	4.79	4.40*	0.10	−0.06	0.00	−0.08	−0.47
NBPB	50.98	50.29	40.50*	53.07**	52.00	0.69	0.00	−9.79	2.78	1.71
NPB	69.42	68.33	57.13*	70.57**	70.00	1.08	−0.02	−11.22	2.23	1.66
LLPB	116.73	112.68	113.43	117.44	118.50**	1.75	−2.30	−1.56	2.46	3.52
CBD	8.71	25.92	61.89**	6.48*	7.50	−10.06	7.16	43.12	−12.29	−11.27
CLR	14.00	14.13	16.86**	11.14	9.65*	0.15	0.28	3.02	−2.71	−4.20
HBW	15.18**	14.42	13.78	14.16	12.80*	0.55	−0.21	−0.86	−0.48	−1.84
BL	10.02**	9.94	9.73	9.79	9.34*	0.09	0.01	−0.20	−0.14	−0.59
BW	6.54**	6.36	6.36	6.33	6.16*	0.12	−0.07	−0.06	−0.09	−0.26
BT	3.78**	3.68	3.62	3.65	3.59*	0.06	−0.03	−0.09	−0.06	−0.13
FL	15.61**	15.13	14.78*	15.18	15.05	0.31	−0.17	−0.53	−0.13	−0.25
FW	11.30**	11.11	11.26	11.05	10.70*	0.12	−0.07	0.08	−0.13	−0.48
LW	5.41	5.25*	5.43**	5.39	5.35	0.07	−0.09	0.09	0.05	0.01
LS	43.48	41.59*	44.31**	43.95	43.35	0.66	−1.22	1.50	1.13	0.53

Where * is the lowest cluster mean difference; ** is the highest cluster mean difference; YLD, coffee bean yield (kg/ha); PH, plant height; HUFPB, height up to first primary branch; SD, stem diameter (cm); CD, canopy diameter (cm); AINL, average internode length (cm); NBPB, number of bearing of primary branch; NPB, number of primary branch; LLPB, length of longest primary branch (cm); CBD, coffee berry disease; CLR, coffee leaf rust; HBW, hundred bean weight (gm); BL, bean length (mm); BW, bean width (mm); BT, bean thickness; FL, fruit length (mm); FW, fruit width (mm); LW, leaf width; LS, leaf size.

Table 3. Inter cluster genetic divergence (D^2) based on 19 quantitative traits

Clusters	I	II	III	IV	V
I		24.57	87.47**	23.62	139.13**
II			34.48*	87.38**	260.67**
III				174.30**	371.29**
IV					53.17**
V					

*Significant ($P < 0.05$), **highly significant ($P < 0.01$) $\chi^2 = 34.80$, ($P < 0.05$) $\chi^2 = 28.87$.

contributions to the first PC that explained the highest percentage of the total variation. The traits: coffee bean yield, stem diameter, canopy diameter, number of bearing primary branches, number of primary branches, length of longest primary branches, CBD, hundred bean weight, seed length, seed width, seed thickness and fruit width showed high contributions to the second PC that explained 18.8% of the total variation.

Shannon diversity indices

In this study, Shannon diversity index values were variable among traits ranging from 0.35 to 1.18 (Table 5). Traits such as branching habit, fruit shape, growth habit, overall appearance and stem habit showed high Shannon diversity index values of 0.66, 1.18, 1.05, 0.87 and 0.54, respectively, and exhibited high Shannon diversity percentage contributions to the total variation compared to the other traits. The overall mean of H' value of 0.70 confirmed the high level of diversity among the studied Amaro Kelo coffee accessions.

Discussion

Based on the 19 quantitative traits, cluster analysis classified the 64 Amaro Kelo coffee accessions into five cluster groups (Fig. 1). Four (Angafa, 7440, Feyate and Koti) out of the six (67%) of the standard checks were grouped into the first cluster (cluster-1), which might indicate that these released varieties might have high similarity. This cluster also showed the third highest mean cluster bean yield, while Cluster V with one accession and cluster-IV with about seven accessions showed the highest and the second highest mean cluster bean yields, respectively, which might indicate these clusters possibly contain genotypes with better bean yield potential than the checks. Similar clustering studies with a comparable number of clusters were reported on Ethiopian coffee of different geographic regions, such as Harerghe (Adem, 2009), and Limu (Olika *et al.*, 2011) coffee germplasm, which provides additional evidence of the existence of high genetic diversity in the different Ethiopian coffee germplasm. The clustering pattern using the quantitative traits in this study showed that accessions collected from the different

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press

Table 4. Eigenvalues and Eigenvectors of the first six principal components (PCs) based on the 19 quantitative characters of the 64 Amaro Coffee Arabica germplasm at Awada

	Eigenvectors					
Character	PC1	PC2	PC3	PC4	PC5	PC6
Coffee bean yield (kg)	0.40	0.50	−0.28	−0.51	0.13	−0.08
Plant height (cm)	0.56	0.27	0.30	0.58	−0.33	−0.06
Height up to first primary branch (cm)	0.37	0.02	0.06	0.09	−0.14	0.47
Stem diameter (cm)	0.20	0.68	0.08	−0.02	0.35	0.43
Canopy diameter (cm)	0.36	0.72	0.17	−0.01	0.34	0.24
Internodes length (cm)	0.50	0.12	0.50	0.45	−0.30	−0.17
Number bearing of primary branches	0.47	0.51	−0.43	0.11	−0.09	−0.26
Number of primary branches	0.45	0.56	−0.50	0.05	−0.17	−0.21
Length of longest primary branch (cm)	0.48	0.54	0.16	0.27	0.30	−0.07
Coffee berry disease (%)	−0.37	−0.30	0.39	0.50	0.27	0.20
Coffee leaf rust (%)	−0.08	−0.26	−0.25	0.48	0.62	−0.30
Hundred bean weight (g)	0.63	−0.56	−0.32	−0.04	−0.12	0.05
Seed length (mm)	0.63	−0.46	0.03	−0.01	−0.27	0.27
Seed width (mm)	0.65	−0.43	−0.10	−0.02	0.13	−0.18
Seed thickness (mm)	0.72	−0.46	−0.07	0.01	0.25	−0.17
Fruit length (mm)	0.66	−0.25	−0.07	−0.08	−0.03	0.32
Fruit width (mm)	0.45	−0.51	−0.26	0.00	0.41	0.17
Leaf width (cm)	0.33	−0.03	0.78	−0.38	0.14	−0.23
Leaf size (cm^2)	0.34	−0.10	0.76	−0.42	0.10	−0.20
Eigenvalues	4.43	3.58	2.51	1.71	1.42	1.11
Difference	0.85	1.07	0.81	0.28	0.31	0.13
Percent of variation	23.32	18.85	13.22	8.98	7.49	5.84
Cumulative	0.23	0.42	0.55	0.64	0.72	0.78

Table 5. Shannon diversity indices for 10 qualitative morphological characters of 64 evaluated *C. arabica* accessions

Plant traits	H'	H_{max}	(%) contributed to variation
Growth habit	1.05	1.10	95.45
Stem habit	0.54	0.69	78.26
Branching habit	0.66	0.69	95.65
Angle of insertion	0.48	0.69	69.57
Leaf tip colour	1.01	1.61	62.73
Leaf shape	0.35	0.69	50.72
Leaf apex shape	0.38	0.69	55.07
Stipule shape	0.45	1.10	40.91
Fruit shape	1.18	1.39	84.89
Overall appearance	0.87	1.10	79.09
Overall mean of H'	0.70	0.98	71.23

PAs were clustered together in the same group; for instance, accessions collected from all the PAs were grouped in clusters I and II (Table 1, Fig. 1). This result was supported by Seyoum (2003), who reported grouping of the accessions into the same cluster regardless of their geographic origins in the study conducted on Arabica coffee accessions collected from different coffee-growing regions of Ethiopia. Similarly, accessions collected from the same PAs were also classified into different clusters implying the diversity in the accessions obtained from the same PA. It is likely that geographic origin does not affect the clustering pattern and diversity of the accessions, which might be due to the exchange of seeds from one PA to the other. Such diversity serves as an important source of accessions for future coffee breeding efforts to broaden the crop's genetic base, which may help develop new varieties utilizing Amaro Kelo accessions.

The qualitative traits clustered the Amaro Kelo coffee accessions into six distinct clusters (Table 1), showing high genetic diversity in this accessions. In line with this, other studies reported the presence of high genetic diversity of Ethiopian coffee for the qualitative traits using cluster analysis (Atinafu *et al.*, 2017; Masreshaw, 2018). It was also revealed that the clustering pattern of Amaro Kelo coffee accessions using quantitative traits was independent and distinct from the clustering pattern based on

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press

the qualitative traits. This suggests that the selection of genotypes or parental lines in improving coffee through hybridization needs to consider the clusters based on the quantitative and qualitative traits independently of each other.

The inter-cluster distance (D^2) analysis for both of the quantitative and qualitative traits revealed highly significant differences ($P < 0.01$) among most of the clusters (Table 3; online Supplementary Table S3). The maximum genetic recombination and heterosis in the subsequent progenies are expected from crosses involving parents from the clusters with the highest cluster distances (Singh and Chaudhary, 1987). Accordingly, crosses between parental lines selected from cluster V with cluster III, cluster II with cluster V and cluster III with cluster IV based on the quantitative traits are expected to produce relatively higher genetic recombination and hybrid vigour in their progenies. Similarly, highly significant inter-cluster distances (D^2) based on qualitative traits were also found between cluster II and cluster V, cluster V and cluster VI, cluster I and cluster VI and cluster III with VI (online Supplementary Table S3). The high inter-cluster distance might indicate a high chance of obtaining transgressive segregates and maximizing heterosis from crossing parents belonging to different clusters, as there is a higher chance that distinct accessions would contribute unique desired alleles at different loci (Ghaderi *et al*., 1984). However, parental selection needs to take the specific merits of each cluster and a genotype within a cluster into consideration that could potentially contribute desirable genes for traits of interest towards the objectives of the hybridization programme.

Characters with the largest absolute values closer to unity within the first few principal components influence the clustering more than those with lower absolute values closer to zero (Chahal and Gosal, 2002). Accordingly, traits with relatively higher influence in the first principal component (PC1) were hundred bean weight, bean length, bean width, bean thickness, fruit length, plant height, stem diameter and average internodes length on the main stem, and hence had more contributions to the total variation and were the ones that most differentiated the clusters (Table 4). The traits coffee bean yield, number of primary branches, number of bearing primary branches, length of the longest primary branch, fruit width, hundred bean weight, canopy diameter and stem diameter mainly influenced the variation in the second principal component (PC2). Traits such as average internode length, number of primary branches, leaf width and leaf size were the major traits associated with the third principal component (PC3), indicating their important role in the total variability of Amaro Kelo coffee accessions. In line with this finding, Akpertey *et al*. (2019) reported six traits, i.e. plant height, stem diameter, span, number of laterals per tree, diameter of laterals and number of nodes per lateral had high contributions to the first PC that accounted for 39.2% of the total variation in the 71 Robusta coffee accessions. This indicated the importance of these traits in classifying Robusta coffee accessions into different clusters in Ghana. The authors also recommended considering these traits in selecting diverse parents for a hybridization programme in improving Robusta coffee. Similar findings were also reported by Kebede and Bellachew (2008), Olika *et al*. (2011), Gessese *et al*. (2015) and Masreshaw (2018), where the grouping of coffee accessions from different geographic origins was performed using PCA in Ethiopia. Yigzaw (2005) also reported the high contributions of inter-node length, tree height, canopy diameter, number of branches, bean and fruit characters to the total variations in the Ethiopian coffee germplasm. Likewise, Masreshaw (2018) reported that average inter-node length of primary branches, the average length of primary branches, canopy diameter, fruit width, fruit thickness, bean width, bean thickness and hundred bean weight showed predominant contributions to the total variation among Yayo coffee germplasm. This finding agrees with Olika *et al*. (2011), who reported bean length, hundred bean weights and leaf width had major contributions to the total variation of Limmu coffee germplasm.

Shannon diversity indices (H') were used to compare phenotypic diversity among the qualitative characters, and the mean diversity index for the qualitative traits was 0.78, with mean contribution of 71% to the total variation (Table 5). The qualitative traits that showed more than average contributions to the total variation included branching habit (95.6%), growth habit (95.4%), fruit shape (84.9%), overall appearance (79.1%) and stem habit (78.3%). In line with this finding, Masreshaw (2018) also reported high contributions of leaf tip colour, unlikely stipule shape, leaf shape and leaf apex shape to the total variations in the studied 64 Yayo coffee germplasm. This author also reported a high Shannon diversity index (H') for fruit colour, young leaf tip colour, stipule shape and leaf shape. Similarly, Yigzaw (2005) and Adem (2009) reported that the Shannon diversity values were variable among the different coffee germplasm and qualitative traits and ranged from 0.41 to 0.99 for the coffee germplasm sourced from various parts of the country and 0.17 to 0.39 for the West Harerge coffee germplasm, respectively. Gizachew and Mohammed (2017) also reported a minimum value of H' (0.67) for branching habit and a maximum value (0.98) for leaf shape. The lowest diversity index in this study was 0.35 for the trait leaf shape, which, according to Hennink and Zeven (1990), indicates unbalanced frequency classes and a lack of diversity for the trait. In general, the diversity indices of all the qualitative traits suggested the presence of high variability for these traits among the evaluated Amaro Kelo coffee accessions.

The cluster analysis employed dendrogram, cluster mean analysis and inter-cluster distance to characterize the Amaro Kelo local coffee accessions. The dendrogram classified the accessions into five clusters, and the cluster means characterized each of the cluster groups based on the 19 quantitative traits that helped identify the cluster groups with high potential for the different studied traits. Moreover, the inter-cluster distance displayed how the cluster groups are closely or distantly related. On the other hand, PCA decomposed the total variation into several principal components in which the first few principal components showed the highest contributions to the total variation. Eigenvalues combined with Eigenvectors were used to identify the traits that showed the highest contribution to the total variation. All the techniques used, i.e. cluster, principal component and Shanon–Weaver diversity index provided unique and complementary information that confirmed the local Amaro Kelo coffee accessions showed high variations for the studied characters, which might provide opportunities for genetic improvement through selection, hybridization and conservation of the genotypes for future utilization. Several other workers have also reported similar high genetic diversity in Arabica coffee (Seyoum, 2003; Yigzaw, 2005; Kebede and Bellachew, 2008) and in Robusta coffee (Akpertey *et al*., 2019). Hence, the availability of genetic diversity in the local Amaro Kelo coffee accessions may provide an opportunity to improve coffee through selection and hybridization. This accession needs to be conserved appropriately using both *in situ* and *ex situ* (in field gene bank, *in vitro* or

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press

cryopreservation) techniques (Engelmann *et al.*, 2007), and can serve as a resource for potential future coffee genetic improvement. In addition, the observed variability for some of the important quality and disease resistance traits might be exploited to improve the productivity, marketability and consumption of coffee.

The results of this study for the quantitative traits are based on data recorded only for one year because the scope of the study was to perform an initial characterization of the accessions to implicate the potential of the accessions for future utilization and identify superior genotypes for advanced multilocation variety trials. Hence, other similar analyses based on data collected in subsequent years may change the study outcomes due to genotype by seasons interaction.

Summary and conclusion

Cluster analysis using the significant quantitative traits confirmed high variation among the local Amaro Kelo coffee accessions, and grouped them into five clusters. The smallest inter-cluster distance ($D^2 = 23.62$) was observed between clusters I and IV; while the highest ($D^2 = 371.29$) was between clusters III and V. Moreover, the studied coffee accessions were grouped into six distinct groups based on 10 qualitative characters. Since maximum genetic recombination and variation in the subsequent generation are expected from crosses that involve parents from the clusters characterized by maximum distances, considering the distance between cluster groups in selecting parental lines for hybridization is important in achieving good genetic recombination and segregation in the progenies.

The PCA revealed that the first two principal components, i.e. PC1 and PC2, with respective PC values of 23.32 and 18.85%, respectively, contributed more to the total variation. Thus, almost all the studied traits contributed to the discrimination of the tested accessions. However, hundred bean weight, bean length, bean width, bean thickness, fruit length and plant height contributed to the total variation and differentiated the clusters. The qualitative traits with high Shanon diversity values (H'), such as branching habit, fruit shape, growth habit, overall appearance and stem habit, revealed high contributions to the overall variability, indicating the importance of Amaro Kelo coffee accessions to improve these traits.

Generally, the present study indicated enormous genetic variability among Amaro Kelo coffee collections for various important morphological traits. Hence, there is an opportunity to exploit these accessions to improve the various quantitative traits identified as important contributors to the total variation in PC-I, II and III, i.e. hundred bean weight, bean length, bean width, bean thickness, fruit length, plant height, stem diameter, average internodes length on the main stem, coffee bean yield, number of primary branches, number of bearing primary branches, length of longest primary branch, fruit width, canopy diameter, stem diameter, average internode length, leaf width and leaf size.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1479262122000119.

Financial support. This study was funded by the Ethiopian Institute of Agricultural Research.

Conflict of interest. None.

Ethical standards. The authors would like to acknowledge the smallholder farmers of the Amaro Kelo district for maintaining the coffee accessions used in this study. The coffee accessions were collected by Jimma Agricultural Research Center (a public research centre) for use in coffee genetic improvement work that will ultimately made available for public use, including smallholder farmers.

References

Adem A (2009) *Agro-Morphological Characterization of Coffee (Coffea arabica L.) Landrace Collected from Mesela, West Harerge, Ethiopia* (M.Sc. Thesis). Submitted to Graduate Studies of Hawassa University, Hawassa, Ethiopia, 88pp.

Akpertey A, Anim-Kwapong E and Ofori A (2019) Assessment of genetic diversity in Robusta coffee using morphological characters. *International Journal of Fruit Science* **19**, 276–299.

Alemseged Y, Tesfaye S, Taye K and Endale T (2015) *Manual on Modern Coffee Tree Management. Amharic Version*. EIAR, Addis Ababa, Ethiopia, 66p.

Atinafu G, Mohammed H and Kufa T (2017) Genetic variability of Sidama coffee (*Coffea arabica* L.) landrace for agro-morphological traits at Awada. *Southern Ethiopia Academic Research Journal of Agricultural Science and Research* **5**, 263–275.

Belachew B (2000) *Arabica Coffee Breeding for Yield and Resistance to Coffee Berry Disease (Colletotrichumkahawae sp.)* (Ph.D. Thesis). Submitted to the Imperial College at Wye University of London, London, UK, p. 201.

Benti T, Gebre E, Tesfaye K, Berecha G, Lashermes P, Kyallo M and Yao NK (2020) Genetic diversity among commercial arabica coffee (*Coffea arabica* L.) varieties in Ethiopia using simple sequence repeat markers. *Journal of Crop Improvement* **35**, 147–168. doi: 10.1080/15427528.2020.1803169

Berthaud J and Charrier A (1988) Genetic resources of Coffea. In Clarke RJ and Macrae R (eds), *Coffee Agronomy*. London: Elsevier Applied Science, pp. 1–41.

Central Statistical Agency (CSA) (2020) *Agricultural Sample Survey: Report on Area and Production of Major Crops of Private Peasant Holdings for Meher Season of 2017*. Addis Ababa, Ethiopia: Central Statistical Agency.

Chahal GS and Gosal SS (2002) *Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches*. New Delhi, India: Narosa Publishing House, p. 604.

Charrier A and Berthaud J (1985) Botanical classification of coffee. In Clifford MN and Wilson KC (eds), *Coffee Botany, Biochemistry and Production of Beans and Beverage*. London: Croom Helm, pp. 13–47.

Cliff N (1988) The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin* **103**, 276.

Coffee Industry Statistics (2020) *Global Coffee Industry Statistics and Consumption Trends – Updated for 2020*. Coffee Industry Statistics. Available at https://en.wikipedia.org/wiki/Economics_of_coffee (Accessed 29 January 2021).

Copper MC and Milligan GW (1988) The effect of error on determining the number of clusters. *Proceeding of International Workshop on Data Analysis, Decision Support and Expert Knowledge Representation in Marketing and Related Areas of Research*, pp. 319–328.

Daba G, Helsen K, Berecha G, Lievens B, Debela A and Honnay O (2019) Seasonal and altitudinal differences in coffee leaf rust epidemics on coffee berry disease-resistant varieties in Southwest Ethiopia. *Tropical Plant Pathology* **44**, 244–250.

Davis AP, Gole TW, Baena S and Moat J (2012) The impact of climate change on indigenous Arabica coffee (*Coffea arabica*): Predicting future trends and identifying priorities. *PLoS One*, **7**, 1–13.

De Vienne D, Santon S and Falque M (2003) Principal sources of molecular markers. In Vienne DD (ed.), *Molecular Markers in Plant Genetics and Biotechnology*. Plymouth, UK: Science Publishers, Inc., pp. 3–41.

Engelmann F, Dulloo ME, Astorga C, Dussert S and Anthony F (eds) (2007) *Complementary Strategies for Ex Situ Conservation of Coffee (Coffea arabica L.) Genetic Resources. A Case Study in CATIE, Costa Rica*. Rome, Italy: Topical reviews in Agricultural Biodiversity. Bioversity International, x + 63pp.

Geromel C, Ferreira LP, Guerreiro SMC, Cavalari AA, Pot D, Pereira LFP and Marraccini P (2006) Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *Journal of Experimental Botany* **57**, 3243–3258.

Gessese MK, Bellachew B and Jarso M (2015) Multivariate analysis of phenotypic diversity in the South Ethiopian coffee (*Coffea arabica* L.) for quantitative traits. *Advances in Crop Science and Technology* S1: 003.

Ghaderi A, Adams MW and Nassib AM (1984) Relationship between genetic distance and heterosis for yield and morphological traits in dry edible bean and faba bean. *Crop Science* **24**, 37–42.

Gichimu BM and Omondi CO (2010) Morphological and 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.

Gizachew A and Mohammed H (2017) Agro-morphological characterization of sidama coffee (*Coffea arabica* L.) accessions accession under its specialty coffee growing area, Awada, Southern Ethiopia. *International Journal of Research Studies in Science, Engineering and Technology* **4**, 11–23.

Hameed A, Hussain SA, Ammar S, Ansar H and Suleria R (2018) 'Coffee bean-related' agroecological factors affecting the coffee. In: J.-M. Merillon, K. G. Rmawat (Eds.), *Co-Evolution of Secondary Metabolites*. Reference series in Phytochemistry, Springer Nature, Switzerland AG 2018, pp. 1–67.

Hennink S and Zeven AC (1990) The interpretation of Nei and Shannon Weaver within population variation indices. *Euphytica* **51**, 235–240.

Hindorf H and Omondi CO (2011) A review of three major fungal diseases of *Coffea arabica* L. in the rainforests of Ethiopia and progress in breeding for resistance in Kenya. *Journal of Advanced Research* **2**, 109–120.

International Coffee Organization (ICO) (2018) *Total Production by all Exporting Countries*. International Coffee Organization, 222 Gray's Inn Road, London, WC1X 8HB. Available at http://www.ico.org/trade_statistics.asp.

International Coffee Organization (ICO) (2020) *Total Production by All Exporting Countries*. International Coffee Organization. 222 Gray's Inn Road, London WC1X 8HB Available at http://www.ico.org/trade_statistics.asp.

IPGRI (1996) *Descriptors for Coffee (Coffea spp and Psilanthus spp)*. Rome, Italy: International Plant Genetic Resources Institute.

Kebede M and Bellachew B (2008) Phenotypic diversity in the hararge coffee (*Coffea arabica* L) accessions for quantitative traits. *East African Journal of Sciences* **2**, 13–18.

Labouisse J-P, Bellachew B, Kotecha S and Bertrand B (2008) Current status of coffee (*Coffea arabica* L.) genetic resources in Ethiopia: implications for conservation. *Genetic Resources and Crop Evolution* **55**, 1079–1093.

Lashermes P, Combes MC, Robert J, Trouslot PD, Hont PA, Anthony F and Carries A (1999) Molecular characterization and origin of the *Coffea arabica* L. geneome. *Molecular and General Genetics MGG* **26**, 259–266.

Lashermes P, Combes MC, Ansaldi C, Gichuru E and Noir S (2011) Analysis of alien introgression in coffee tree (*Coffea arabica* L.). *Molecular Breeding* **27**, 223–232.

Mahalanobis PC (1936) On the generalized distance in statistics. *Proceedings of the National Institute of Science* **2**, 49–55.

Masreshaw (2018) *Genetic Variability Study in Yayu Coffee (Coffea arabica L.) Geremplasm Using Morphological Traits at Metu, Southwestern, Ethiopia* (Thesis). Submitted to Graduate Studies of Jimma University, Jimma, Ethiopia.

Olika K, Sentayehu A, Taye K and Weyessa G (2011) Genetic diversity analysis of limmu coffee (*Coffea arabica* L.) collection using quantitative traits in Ethiopia. *International Journal of Agricultural Research* **6**, 470–481.

SAS (2013) *Statistical Analysis System (Version 9.4)*. Cary, NC, USA: SAS Institute.

Seyoum S (2003) *Genetic Divergence for Seedling Parameters and Associations Among Agronomic Traits in the Ethiopian Coffee (Coffea arabica L.) Accessions* (M.Sc. Thesis). submitted to the School of graduate studies of Alemaya University, Alemaya, Ethiopia.

Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423.

Singh RK and Chaudhary BD (1987) *Biometrical Methods in Quantitative Genetic Analysis*. New Delhi-Ludhiana, India: Kalyani publishers, p. 318.

Tesfaye S, Alemseged Y, Taye K, Endale T and Anteneh N (2006) *Coffee Seedling Management and Production*. Addis Abeba, Ethiopia: Ethiopian Agricultural Research Organization, 17p (Amharic version).

Tesfaye K, Borsch T, Govers K and BeKelo E (2007) Characterization of Coffea chloroplast microsatellites and evidence for the recent divergence of *C. arabica* and *C. eugenioides* chloroplast genomes. *Genome* **50**, 1112–1129.

United States Department of Agriculture (USDA) (2018) *Coffee: World Markets and Trade*. Foreign Agricultural service. United States Department of Agriculture (USDA), Office of Global Analysis, Washington D.C., USA.

Van der Graaff NA (1981) *Selection of Arabica Coffee Types Resistant to Coffee Berry Disease in Ethiopia*. Wageningen: Mededelingen Londbovwhogeschola, p. 110.

Wrigley G (1988) *COFFEE: Tropical Agriculture Series, London*. New York, USA: John Wiley & Son. Inc., 639p.

Yigzaw D (2005) *Assessment of Cup Quality, Morphological, Biochemical and Molecular Diversity of Coffee arabica Genotypes of Ethiopia Thesis* (Ph.D. Plant sciences (Plant breeding)). University of the Free State, Bloemfontein, South Africa, p. 197.

https://doi.org/10.1017/S1479262122000119 Published online by Cambridge University Press