Genetic Diversity of Coffee (*Coffea arabica* L.) Collections for Morpho-agronomic Traits in Southwestern Ethiopia

Masreshaw Yirga [1*], Wosene Gebreselassie [2], and Abush Tesfaye [3]

[1]Jimma Agricultural Research Center, EIAR, P.O. Box 192, Jimma, Ethiopia

[2]College of Agriculture and Veterinary Medicine, Jimma University, P.O. Box 307, Jimma, Ethiopia

[3]International Institutes of Tropical Agriculture, Oyo Road, P.M.B. 5320, Ibadan, Nigeria

Abstract

Background: Information on the genetic variability of plants on the basis population is important for conservation and utilization of genetic resources. However, information on such genetic diversity is not still yet available at individual level in Yayu coffee germplasm, southwestern Ethiopia.

Objective: The study was conducted to estimate the genetic variability among coffee collections with respect to morpho-agronomic traits.

Materials and Methods: Sixty-two Coffee (*Coffea arabica* L.) collections with two standard check varieties (74110 and **74112**) were evaluated using 8 x 8 simple lattice design at Metu Agricultural Research Sub Center. The experiment was conducted on six-year old coffee trees during the 2018 main cropping seasons. The coffee trees were managed as per the recommendation for coffee production practices.

Results: Cluster analysis was employed using 19 quantitative traits and 64 coffee collections grouped into seven clusters. Significant inter cluster-distance was found between most of the paired clusters. The results revealed the chance of developing hybrids by crossing coffee collections from cluster-V and VI followed by cluster-IV and VI. Principal component analysis revealed that, the first seven principal components with Eigen values exceeding one were responsible for about 74.94 % of the observed variation among the coffee collections. Out of the entire variations, the first and the second principal components accounted for more than one-third of the total variation (35.32 %).

Conclusion: The information and genetic variability obtained in the present study could be used to plan conservation, effective pure line selection, and crossing of coffee germplasm in future coffee improvement programs.

Keywords: Germplasm; Cluster; Genetic divergence; Principal component

1. Introduction

Coffee (*Coffea arabica* L.) belongs to the genus *Coffea,* in the family *Rubiaceae. Coffea arabica* is one the most important commercial species in the world market (Gray *et al*., 2013). It is the most widely drunk beverages in the world due to its best cup quality and source of income for million people in coffee growing countries (Lashermes *et al*., 2011; Mishra and Slater, 2012). Ethiopia is the fifth major exporter of Arabica coffee in the world next to Brazil, Vietnam, Colombia and Indonesia; while it is the highest producer among African country. Coffee plays significant role in Ethiopian cultural and socio-economic life of the nation. It contributes about 35% of the country's foreign currency earnings and about 25%, the population directly or indirectly drives their income from coffee value chain (USDA, 2020).

Ethiopia is also the origin and center of genetic diversity for *Coffee arabica* species. There is a high genetic diversity of coffee in the country, which is confirmed by several phenotypic and molecular studies (Sylvain, 1958; Meyer, 1965; Kassahun Testate, *et al*., 2013; Mesfin Kebede and Bayetta Belachew, 2005; Tadesse Benti *et al*., 2021). The entire genetic diversity of indigenous (wild) Arabica is confined mainly in the Afromontane rain forest located in the West and East of Great Rift Valley (Kassahun *et al.,* 2013). Hence, the existence of wide genetic variability is expected to safeguard coffee production from dangers posed by possible biotic and abiotic stresses (Tadesse Woldemariam, 2003). Regardless of the presence of substantial genetic diversity in the crop species, the country is still not yet fully utilizing its coffee genetic resources as expected in terms of improving coffee

*Corresponding Author. yirgamasresha@gmail.com

ISSN 1993-8195 (Online), ISSN 1992-0407(Print)

productivity and the livelihood of the rural community (Paulos and Teketay, 2000). Consequently the national productivity of coffee per unit area remains very low (0.7 t ha^{-1}) (CSA, 2018). Major contributing factors for low production are uses of unimproved local landrace, Conventional husbandry and processing practices (Taye, 2010), and the direct and indirect potential impact of climatic variability (Davis *et al.*, 2012).

Arabica coffee gene pool has threatened by genetic erosion mostly attributed to deforestation of its natural habitat, establishment and expansion of modern plantation with illegal and legal settlements (Woldemariam *et al.*, 2002). Since 1973, considerable coffee germplasm collections have made to capture the available coffee genetic variability for the purpose of selecting and developing adaptable coffee varieties. Hence, about 12,452 indigenous and exotic coffee germplasm were collected and *ex-situ* conserved at the Ethiopian Biodiversity Institute (EBI) (5731 collections) (Taye, 2010) and Jimma Agricultural Research Center (6721 collections) field gene banks (Tadesse, 2017). However, some germplasm died in their maintenance fields due to climate change and adaptation problem, as they are forced to be grown outside of their original environment. It has been well understood that, varieties belonging to one region adapt differently when grown in another region (Bayetta *et al.*, 1993). Bearing in mind this fact, the national coffee breeding strategy has currently designed location specific adaptation for local landrace variety development and promotion under diverse coffee growing agro ecologies (Fikadu *et al.*, 2008).

Information on the nature and magnitude of genetic variability present in any crop species is a key resource for developing effective crop improvement program through selection or crossing of different parental lines (Dabholkar, 1996). These genetic variations can be enumerated at species, populations and individuals. Genetic variability is genetic differences among individuals within a population. This is the vital information for plant breeding activities, because proper management of diversity can produce permanent gain in the performance of plant and can safeguard against seasonal fluctuations (Sharma, 1998).

Therefore, the studies on coffee diversity are vital for conservation of genetic resources and improvement of crops. Multivariate analysis is a useful tool in quantifying the degree of genotypic divergence among biological populations and to assess the relative contribution of different components to the total divergence levels (Murty and Arunachalam, 1966; Das and Gupta, 1984). Such a study also permits to select the genetically diverse parents to obtain the desirable recombinant in the segregating populations upon crossing. In the hybridization programs, inclusion of more diverse parents has been observed to increase the chance for obtaining strong heterosis and giving broad spectrum of variability in segregating generations (Joshi and Dhawan, 1966).

Yayu forest is one of the Afromontane rainforest Biosphere Reserves, primarily designated for *in situ* conservation of wild *Coffea arabica* gene pool in Southwestern highlands of Ethiopian (,Dereje Likissa, 2014). Some morpho-physiological and molecular diversity assessment studies at population level suggested the existence of genetic variability in Yayu coffee gene pool (Tadesse Woldemariam and Feyera Senebeta, 2008; Taye Kufa, 2006; Esayas Aga, 2005; Kassahun Testate*, et al.*2013). Therefore, systematic and detail characterization of coffee collections at individual level is very important for effective conservation and efficient exploitation of its germplasm through selection and crossing in coffee variety development prorgam. Considering these facts, the study was conducted to estimate the extent of genetic variability among collected coffee germplasm in this region using multivariate analysis with respect to quantitative traits.

2. Materials and Methods

2.1. Description of the Experimental Site

The experiment was conducted at Metu Agricultural Research sub-Center during the 2018 cropping season. Metu is located 600 km away from Addis Ababa in the southwesterly direction in Illubabor zone of the Oromia Regional State. The sub center is situated at a distance of 3 km from Metu town. The geographical location of the sub center is 8°19' 0" N latitude 35°35' 0"E longitude and altitude of 1558 meters above sea level. The mean annual temperature ranges from 12.7 and 28.9 ^{0}C with annual rainfall of 1829 mm/annum. The major soil type is Nitosols with pH of 5.24 (Paulos Dubale, 2001).

2.2. Experimental Materials, Design and Field Management

Sixty-two *Coffea arabica* germplasm collected from Yayu woreda of Illubabor Zone and two commercially grown check varieties were used for this study (Table 1). The study was conducted on six-year old coffee trees during the 2018 main cropping seasons. Experiment was laid down in an 8 x 8 simple lattice design. Each collections was planted in a single row of six trees using spacing of 2 x 2 m. Collections were established under a fast growing *Sesbania sesban* legume shade tree. All other management practices were also uniformly applied for the coffee trees as per the Jima Agriculture Research Center recommendation for coffee production.

Table 1. Description of *Coffea arabica* germplasm collections used in the study.

Collection	District	Specific collection site	Collection	District	Specific collection site
Y63	Yayu	Dogi	Y95	Yayu	Geri geba
Y64	Yayu	Dogi	Y96	Yayu	Geri geba
Y65	Yayu	Dogi	Y97	Yayu	Geri geba
Y66	Yayu	Dogi	Y98	Yayu	Geri geba
Y67	Yayu	Dogi	Y99	Yayu	Geri geba
Y68	Yayu	Sembo	Y100	Yayu	Geri geba
Y69	Yayu	Sembo	Y101	Yayu	Geri geba
Y70	Yayu	Sembo	Y102	Yayu	Geri geba
Y71	Yayu	Sembo	Y103	Yayu	Geri geba
Y72	Yayu	Sembo	Y104	Yayu	Geri geba
Y73	Yayu	Sembo	Y105	Yayu	Gordeya
Y74	Yayu	Sembo	Y106	Yayu	Gordeya
Y75	Yayu	Sembo	Y107	Yayu	Gordeya
Y76	Yayu	Sembo	Y108	Yayu	Gordeya
Y77	Yayu	Sembo	Y109	Yayu	Gordeya
Y78	Yayu	Sembo	Y110	Yayu	Gordeya
Y79	Yayu	Sembo	Y111	Yayu	Gordeya
Y80	Yayu	Sembo	Y112	Yayu	Gordeya
Y81	Yayu	Geba	Y113	Yayu	Degitu
Y82	Yayu	Geba	Y114	Yayu	Degitu
Y83	Yayu	Geba	Y115	Yayu	Degitu
Y84	Yayu	Geba	Y116	Yayu	Degitu
Y85	Yayu	Geba	Y117	Yayu	Degitu
Y86	Yayu	Geba	Y118	Yayu	Degitu
Y87	Yayu	Achebo	Y119	Yayu	Degitu
Y88	Yayu	Achebo	Y120	Yayu	Degitu
Y89	Yayu	Achebo	Y121	Yayu	Degitu
Y90	Yayu	Achebo	Y122	Yayu	Degitu
Y91	Yayu	Achebo	Y123	Yayu	Degitu
Y92	Yayu	Achebo	Y124	Yayu	Degitu
Y93	Yayu	Achebo	74110	Metu	Bishari
Y94	Yayu	Achebo	74112	Metu	Bishari

2.3. Data Collection

During the course of this study data on 25 quantitative traits, included: Height up to first primary branch (cm), total tree height (cm), number of main stem node, Average Inter-node length on orthotropic branch (cm), main stem diameter (mm), canopy Diameter (cm), number of primary branches, number of secondary branches, Percentage of bearing primary branches (%), number of nodes on primary branches, length of primary branches (cm), average inter-node length on primary branches (cm),leaf length (cm), leaf width (cm), leaf area (cm^2), fruit length (mm), fruit width (mm), fruit thickness (mm)**,** bean length (mm),bean width (mm), bean thickness (mm), hundred bean weight(gm), yield per tree (kg), coffee berry disease and rust severity (%) were recorded on tree basis from each coffee collection using the standard procedures of the International Plant Genetic Resources Institute (IPGRI, 1996) coffee descriptor.

2.4. Data Analysis

2.4.1. Analysis of variance (ANOVA)

ANOVA of 8 X 8 simple lattice design was subjected using SAS software for each trait. The simple lattice design analysis of variance as structured is stated in Table 2 (Cochran and Cox, 1957).

Table 2. Analysis of variance (ANOVA) for simple lattice design.

Source of variations	Df	SS	MS	F-valus
Replications	(r–1)	SS_r	MS_r	MS_r/MSe
Genotype (adjusted)	(k^2–1)	SS_g	MS_g	MS_g / MSe
Blocks with in replication (adj.)	r (k–1)	SS_b	MS_b	MS_b/MSe
Intra block error	(k–1)(rk–k–1)	Sse	MSe	

Note: *r = Number of replication; g = Number of genotypes; Df = Degrees of freedom; k = Block sizes; SS = Sum squares; MS = Mean squares; SSr = Sum squares of replication; SSg = Sum square of genotypes; SSb = Sum square of block; SSe = Sum square of error; MSr = Mean of square due to replication; MSg = Mean of square due to genotypes; MSb = Mean square of block within replication; and MSe = Mean of square due to error.*

Simple lattice design ANOVA was computed using the following model:

$$Y_{ijk} = \mu + t_i + \beta_j + \chi_{k(j)} + \Sigma_{ijk}$$

Where, Y_{ijk} = response of Y trait from the i^{th} collection under j^{th} replication and k^{th} level of incomplete blocks within replications; μ = overall mean effects; t_i = effects of i^{th} level of collections; β_j = effects of j^{th} level of replication; $\chi_{k(j)}$ = effects of K^{th} level of incomplete blocks within replications; and Σ_{ijk} = the residual or random error component.

2.4.2. Multivariate analysis

Multivariate analysis techniques *viz*. cluster analysis and principal component analysis (PCA) was employed using SAS statistical package software. The numbers of clusters were determined by looking into three statistical approaches, namely, Pseudo-F, Pseudo-t^2 and cubic clustering criteria which is suggested by Copper and Miligan (1988). Accordingly, the number was decided where local peaks of Pseudo-F statistics and cubic clustering criteria combined with small values of Pseudo-t^2 statistics followed by a larger Pseudo-t^2 statistics for the next cluster fusion.

Divergence analysis (D^2) was used to estimate the genetic distance/divergence of the coffee germplasm collections or to classify the divergent collections into different groups and it also measures the forces of differentiation at inter-cluster levels and determines the relative contribution of each component trait to the total divergent (Sharma *et al.*, 1998). Genetic divergence between clusters was determined using the generalized Mahalanobis's D^2 statistics (Mahalanobis, 1936) formula: $D^2_{ij} = (x_i\text{-}x_j)\, s^{-1}\, (x_i\text{-}x_j)$; where, D_{ij}^2 = the distance between class i and j; $X_i\text{-}x_j$ = the difference in the mean vectors of the two populations (class i and j); and s^{-1} = the inverse of pooled variance covariance matrix. The D^2 values obtained for pairs of clusters were considered as the calculated values of Chi-square (χ^2) and tested for significance both at 1% and 5 % probability levels against the tabulated value of (χ^2) for 'P' degree of freedom, where P is the number of traits considered (Singh and Chaudhary, 1987).

The principal components analysis (PCA) was employed in order to minimize the traits into a new set of linearly combined measurements and to identify the traits contributing large part of the total variation among the collections. The analysis was performed using SAS software. In this analysis, only principal components with Eigen values greater than one were considered as important for the total variations.

3. Results and Discussion

3.1. Analysis of Variance

Analysis of variance (ANOVA) revealed the existence of significant ($p<0.05$) variation among coffee germplasm collections for most of the quantitative traits studied except for height up to first primary branch, number of main stem nodes, percentage of bearing primary branches, leaf width, leaf area and fruit length (Table 3). The existences of sufficient variability among the evaluated materials create immense opportunity to bring considerable improvement through selection and cross breeding in the future coffee improvement program.

Therefore, the possible reason for the existence of considerable genetic diversity in the present study will be attributed to either out crossing nature of the crop through different pollinators (Meyer, 1965; Gezahegn Berecha, *et al.*, 2014), or to the gene flow through dissemination of seeds and seedlings from place to place by means of wild animal and human being (Esayas Aga, *et al.* 2005; Feyera Senbeta, 2006). The significant difference observed for measured quantitative traits in this investigation were in agreement with the finding of earlier authors who reported considerable genetic variability within the Arabica coffee germplasm for yield, disease resistance and growth characters (Bayetta Belachew, 1997; Olika Kitila *et al.*, 2011; Getachew WeldeMichael *et al.*, 2013; Ermias H/Mariam, 2005; Yigzaw Desalegn, 2005; Lemi Beksisa and Ashenafi Ayano, 2016; Tadesse Benti *et al.*, 2021; Lemi *et al.*, 2021).

Table 3. Analysis of variance for 25 traits of 64 coffee collections studied at Metu during 2018.

Trait	Mean squares				RE	CV (%)
	Replication (1)	Treatment (adjusted)(63)	Blocks within rep.(adj.)(14)	Error (49)	(%)	
HUP	162.00	12.23 ^{ns}	13.97	9.35	103.38	11.46
TPH	182.41	603.20*	287.13	321.80	97.60	8.61
NMSN	328.64	8.00 ^{ns}	7.28	6.40	100.29	7.84
AILM	7.41	0.48**	0.52	0.17	111.01	7.12
SD	582.68	17.05**	9.24	2.15	101.24	3.10
CD	264.21	212.84**	350.75	63.65	135.14	4.80
NPB	498.49	25.12**	27.02	9.45	123.49	6.44
NSB	506.02	1469.03**	671.90	155.13	100.23	8.65
PBPB	2195.53	164.22^{ns}	133.71	130.6	100.01	34.75
NNPB	14.99	3.03**	5.68	0.34	145.75	3.27
ALPB	834.36	66.18*	173.39	39.57	149.5	8.21
AILPB	8.30	0.38**	0.49	0.04	122.00	4.48
LL	2.95	0.52**	0.46	0.18	105.44	3.47
LW	5.61	0.15 ^{ns}	0.44	0.10	151.16	5.37
LA	690.99	23.41 ^{ns}	35.72	17.13	111.25	8.70
FL	28.69	0.90 ^{ns}	1.54	0.58	120.16	4.57
FW	23.14	0.70**	1.31	0.36	136.86	4.14
FT	24.61	0.50**	1.24	0.28	150.11	4.30
BL	2.91	0.56**	0.24	0.13	98.46	3.31
BW	0.66	0.09**	0.07	0.03	109.27	3.00
BT	0.13	0.04**	0.03	0.02	100.78	3.70
HBW	20.08	4.34**	1.93	0.98	101.56	5.60
CBD	45.55	165.48**	153.72	86.54	106.87	89.45
CLR	129.38	61.14**	49.25	26.21	108.29	49.21
YLD	0.045	0.010*	0.007	0.006	100.420	21.00

Note: *HUP = Height up to first primary branches; TPH = Total plant height; NMSN = Number of main stem nodes; AILMS = Average inter-node length of main stem; SD = Stem diameter; CD = Canopy diameter; NPB = Number of primary branches; NSB = Number of secondary branches; PBPB = Percentage of bearing primary branches; NNPB = Number of nodes of primary branches; ALPB = Average length of primary branches; AILPB = Average inter node length of primary branches; LL = Leaf length; LW = Leaf width; LA = Leaf area; FL = Fruit length; FW = Fruit width; FT = Fruit thickness; BL = Bean length; BW = Bean width; bean thickness; HBW = Hundred bean weight; CBD = Coffee berry disease; CLR = Coffee leaf rust; DF = YLD = Yield per tree; and RE = Relative efficiency. ** = highly significant at p<0.01; * = significant at p<0.05; and ns = non-significant. CV = Coefficient of variation.*

3.2. Cluster Analysis

Out of the 25 study traits, 19 were significantly different among *Coffea arabica* collections. Therefore, cluster analysis was employed using 19 quantitative traits to categorized 64 coffee collections into seven clusters (Table 4 and Figure 1).The distribution pattern revealed that, Cluster-I contained the highest number of collections (19) followed by cluster-III (16), cluster-II (13), cluster-IV (7), cluster-V (4), cluster-VI (4) and cluster-VII (1). In cluster analysis, if the categorization is successful, individuals within (homogenous) shall be closer and different clusters (heterogeneous) shall be farther apart.

Cluster-I contained two coffee collections collected from each of Dogi and Gerigeba site, five from Sembo, three collections from each of Geba, Achebo and Degitu collection sites and one collection from Gordiya. Cluster–II consisted of four collections from Degitu, three collections from Achebo, two collections each from Gerigeba and Sembo, one collection from each of Geba and Gordiya. Similarly, cluster-III was comprised of two standard check varieties (74110 and 74112)

originated from Metu and one collection collected from each Dogi, Geba and Achebo sites, three collections from each Geri Geba, Degitu and Goridya and two collections from each Sembo and Goridya. On the other hand, cluster-IV had seven collections collected from Dogi, sembo, Achibo and Degitu, each with one collection and Geri Geba with three collections. Moreover, cluster-V and VI each possessed four collections collected from Dogi, Sembo, Geba and Degitu. However, the last cluster (VII) was unique to single collection, which was collected from Degitu site.

In this clustering pattern, there were coffee collections collected from the same area of collection sites grouped in to different clusters; the possible cause might be difference in their genetic back ground and gene flow through exchange of seeds or seedlings. There were also collections collected from different site grouped in the same cluster, which was also probably due to originally from the same sources. Therefore, most of the collections were grouped together both from their source and outsource area of collection site, which might be due to selection pressure and genetic drift (Amsalu Ayana and Endashaw Bekele, 1999). This is almost certainly attributable to the continuous movement of coffee seed and seedlings from one site to another by humans or animals.

From this finding, it can be concluded that the selection of coffee collections for hybridization should be based on genetic diversity. This is more under lined by Bayeta Belachew (2001), who suggested as morphological variation is more considerable than collection area as an indicator of genetic diversity in coffee. The current finding is in agreement with Mesfin and Bayetta (2005) who grouped 100 Hararghe coffee accessions in to six clusters. Olika Kitila *et al.* (2011) and Getachew WeldeMichael *et al.* (2013) each grouped 49 Limu coffee collections into four and five clusters, respectively.Moreovere, Lemi *et al.*(2021) also grouped another set of Limu coffee collections in to three clusters.

Table 4. The distribution of 64 coffee collections in seven clusters based on D^2 analysis evaluated at Metu in 2018.

Cluster no.	No. acc.	%	Collections
I	19	30	Y101, Y77, Y67, Y72, Y91, Y88, Y63, Y68, Y107, Y94, Y82, Y114, Y83, Y76, Y112, Y84, Y79, Y122, and Y103
II	13	20	Y111, Y115, Y87, Y106, Y121, Y81, Y90, Y71, Y100, Y116, Y89, Y104, and Y80
III	16	25	Y69, Y86, Y105, Y66, Y108, Y97, Y78, Y117, Y98, Y93, Y109, 74112, 74110,Y99, Y110, and Y120
IV	7	11	Y96, Y124, Y95, Y75, Y102, Y92 and Y64
V	4	6	Y70, Y85, Y65 and Y123
VI	4	6	Y119, Y113, Y73 and Y74
VII	1	2	Y118

3.2.1. Cluster characterization using quantitative traits

Mean performance of different clusters for the 19 traits (Table 5) reflected that the coffee collections in cluster-IV were the high yielder (0.39 kg per tree) followed by cluster-II and III each produced the same yield (0.38 kg per tree). Besides, collections in cluster-VI exhibited the highest total plant height, stem diameter, canopy diameter, number of primary branches, number of secondary branches, number of nodes of primary branches, average length of primary branches, average inter-node length of primary branches, leaf length and bean width. Interestingly, collections in this cluster showed the lowest coffee berry disease severity level, which is an advantage for coffee breeder to develop improved varieties through making use of these collections. Furthermore, collections in this cluster also gave medium values for the rest of traits.

Collections in Cluster-I showed medium mean values for all the traits except for number of primary branches, which had the lowest, mean values. Similarly, the highest cluster mean was found for fruit width, fruit thickness, bean thickness and hundred bean weights in cluster-II and number of nodes of primary branches in cluster-III, while all the rest of the traits produced medium cluster mean value in these two clusters (cluster-II and III). Cluster-IV had the highest cluster mean for yield per tree and coffee leaf rust severity, while the remaining traits scored medium mean values except lowest in stem diameter. Even though, the collections in cluster-IV had relatively highest coffee leaf rust severity level, unfortunately the score value (13 %) lays in moderately resistant level, which will not be difficult for future improvement of those collections in disease-resistant point of view.

Cluster-V possessed collections with longest bean length, but with shortest total plant height, average inter-node length of main stem node, average length of primary branches, average inter-node length of primary branches, narrow canopy diameter, lowest number of

secondary branches and number of nodes of primary branches, while the remaining traits showed medium values. In contrast, cluster-VII which comprised of only one collection was mainly characterized by shortest leaf length and bean length, narrow fruit width, fruit thickness, bean width and bean thickness, lowest for hundred bean weight, coffee leaf rust severity and yield per tree, in contrast the remaining traits had moderate values except having the longest average internodes length of main stem node and highest coffee berry disease severity level. The highest value of coffee berry disease severity in this collection will be difficult to breeders for future improvement of traits in coffee.

Table 5. Mean values of 19 traits for seven clusters of 64 coffee collections evaluated at Metu in 2018.

Traits	Cluster						
	I	II	III	IV	V	VI	VII
TPH	197.7	214.07	222.29	201.81	171.44*	226.83**	223.35
AILMS	5.79	5.81	6.13	5.74	5.16*	6.03	6.30**
SD	48.92	50.3	47.87	47.79*	47.96	51.96**	50.25
CD	167.01	168.62	166.72	161.12	153.79*	181.72**	170.35
NPB	45.31*	49.54	50.38	45.62	43.08	50.52**	49.14
NSB	143.14	172	128.27	105.39	125.91*	194.28**	193.15
NNPB	17.65	17.83	18.22**	17.57	17.26*	18.01	18.1
ALPB	75.51	76.92	77.58	73.58	72.80*	85.72**	77.67
AILPB	4.46	4.48	4.41	4.36	4.31*	4.95**	4.39
LL	12.3	12.03	12.24	12.12	12.36	12.77**	11.02*
FW	14.48	14.64**	14.56	14.06	14.25	14.25	13.70*
FT	12.32	12.52**	12.38	11.91	12.16	12.09	11.56*
BL	10.96	10.99	10.87	10.69	11.03**	10.85	9.94*
BW	6.51	6.61	6.52	6.54	6.51	6.65**	6.28*
BT	3.93	4.00**	3.96	3.94	3.92	3.92	3.69*
HBW	17.67	18.30**	17.6	17.04	17.63	17.83	14.25*
CBD	9.31	10.37	10.06	8.59	15.19	6.23*	60.87**
CLR	9.06	11.73	9.74	13.00**	10.6	12.53	8.22*
YLD	0.37	0.38	0.38	0.39**	0.34	0.35	0.25*
Lowest mean	0.37	0.38	0.38	0.39	0.34	0.35	0.25
Highest mean	197.70	214.07	222.29	201.81	171.44	226.83	223.35
Overall mean	42.23	45.32	43.17	40.07	39.25	48.29	49.08

Note: *TPH = Total plant height; AILMS = Average inter-node length of main stem; SD = Stem diameter; CD = Canopy diameter; NPB = Number of primary branches; NSB = Number of secondary branches; NNPB = Number of nodes of primary branches; ALPB = Average length of primary branches; AILPB = Average inter-node length of primary branches; LL = Leaf length; FW = Fruit width; FT = Fruit thickness; BL = Bean length; BW = Bean width; BT = Bean thickness; HBW = Hundred bean weight; CBD = Coffee berry disease; CLR = Coffee leaf rust; and YLD = Yield per tree. **,* represents maximum and minimum values, respectively.*

3.2.2. Genetic divergence (D^2)

The values of pair wise average intra and inter-cluster divergence (D^2) among 64 coffee collections in seven clusters based on their 19 quantitative traits are presented in Table 6. Accordingly, the inter-cluster distances in all the cases were greater than the intra-cluster distances suggesting wider diversity among the collections of the distant clusters. The intra-cluster degree of diversity was relatively maximum in cluster V and VI (5.55), indicating that the collections in cluster V and VI were a little bit heterogeneous as compared to those in other clusters. Generally, the range of intra-cluster values indicated homogeneous nature of the genotypes within the clusters.

The chi-square test revealed the existence of highly significant differences among the paired inter cluster distance except cluster I and II, I and III, I and IV, I and V, II and III, II and VI, III and IV and VI and V. The maximum inter-cluster distance was found between cluster-IV and VII (396.99) followed by Cluster-V and VII (374.31), cluster-I and VII (260.90), and cluster- III and VII (245.11); while, the lowest inter-cluster distance was recorded between cluster-I and III (12.13) followed by cluster-II and VI (20.55), cluster-IV and V (22.01), and clusters-I and V (23.70). The highest value of inter-cluster distance indicated that the accessions belonging to these cluster were far diverged. On the other hand, the lowest cluster distance indicates a close relationship between the accessions.

Cluster-IV was found divergent from cluster-VII chiefly due to number of secondary branch, coffee berry disease, coffee leaf rust, plant height and yield per tree, indicating maximum contribution of these traits towards the divergence. Similarly, cluster-V was divergent from cluster-VII mainly for total plant height, average inter-node length of main stem node, bean length, and coffee berry disease severity. On the other hand, intra cluster mean performance for most of the traits in cluster-VI was maximum and greater than the mean of cluster-VII except for average inter-node length of main stem node. The highest mean value in cluster-VI indicates the role of those traits towards the divergence between cluster-VI and VII. In this perspective, Jagadev *et al.* (1991) stated that the traits contributing maximum towards the divergence should be given greater emphasis for deciding the type of cluster for the purpose of further selection and choice of the parents for hybridization.

The collections of distant clusters could be used in hybridization program to obtain a higher heterotic response in the hybrids and a wide range of variation among the segregate. The highest inter-cluster distance is found between clusters-IV and VII, suggesting that superior hybrids or recombinants can be realized by crossing between the desirable lines of these clusters. However, the progenies will be high yielder, but they might be susceptible to CBD, because the mean performance of collections in cluster-VII exhibited the highest CBD severity level. Hence, the possible alternative to develop superior hybrids or recombinants would be mating between collections found in Cluster-V & VI followed by cluster-IV and VI. Moreover, the heterosis could also be exploited by crossing between collections with moderate diversity like cluster-II and V, followed by cluster- II and IV, I and VI, III and VI and III and V.

The present result is in support of Bayeta Belachew (2001) who reported the significance of genetic diversity among parents with respect to geographical origin or morphological traits for maximum heterosis to occur in certain hybrid traits of coffee germplasm. Similar reports were also made by Szamosi *et al.* (2010) who found heterosis by crossing between melon genotypes with higher inter cluster distance.

Table 6. Pair wise average intra (bold) and inter cluster divergence values (D^2) among 64 coffee collections in seven clusters based on their 19 quantitative traits tested at Metu in 2018.

Cluster	I	II	III	IV	V	VI	VII
I	**(2.43)**	24.23	12.13	25.05	23.10	73.07**	260.90**
II		**(3.19)**	23.70	77.40**	81.23**	20.55	172.74**
III			**(2.77)**	29.14	49.84**	61.63**	245.11**
IV				**(4.43)**	22.01	141.41**	396.99**
V					**(5.55)**	160.73**	374.31**
VI						**(5.55)**	160.86**
VII							**(0.00)**

Note: ** = *highly significant at* $p<0.01$, $\chi^2 = 34.80$, $p<0.05$, $\chi^2 = 28.87$.

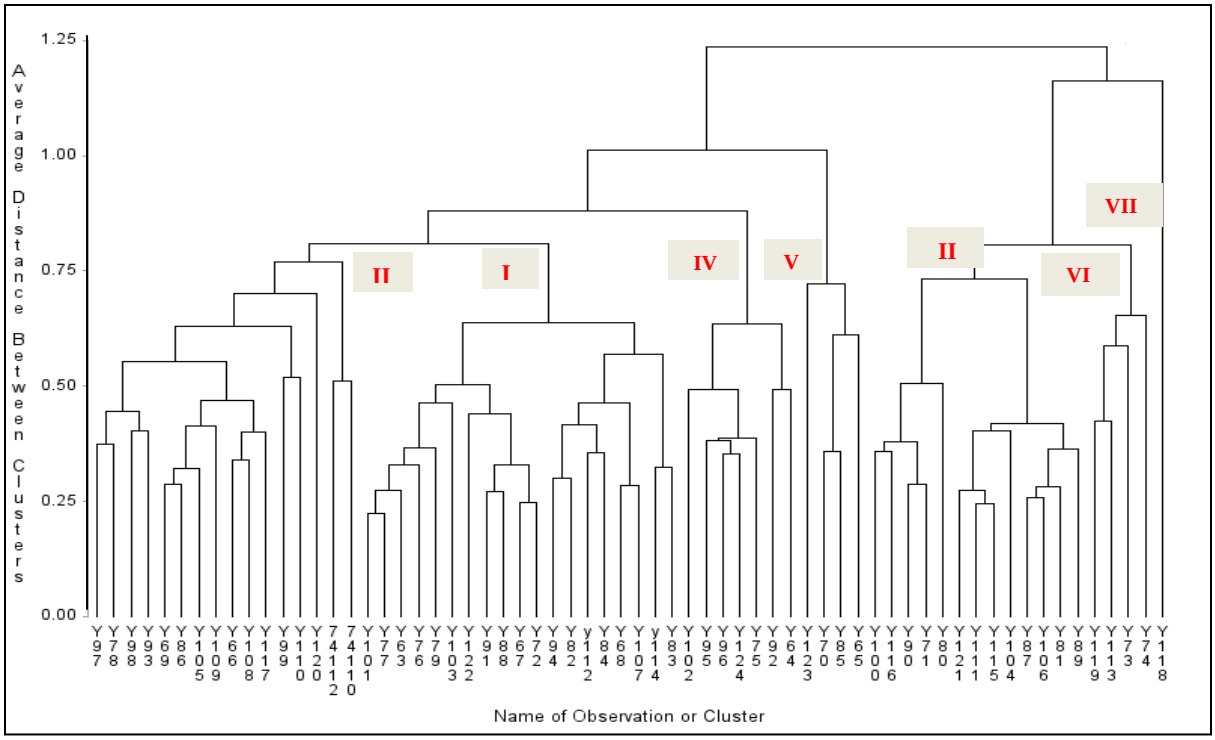

Figure 1. Dendrogram depict 64 coffee accession in seven clusters.

3.3. Principal Component Analysis (PCA)

The principal component analysis was done using 19 quantitative traits with the intention of minimizing the dimensionality of large number of interrelated traits in a given data set and retaining maximum information about the genetic variation. Accordingly, the first seven principal components with Eigen values exceeding one were responsible for about 74.94 % of the total variation among the coffee collections. Out of the entire variations, the first two principal components accounted for the maximum variation (35.32%) among the coffee collections (Table 7 and Figure 2).

The first principal component that accounted the highest total variation (21.99%) was due to the chief contribution of positive discriminatory traits like the average length of primary branches, fruit width, fruit thickness and hundred bean weights. The considerable variation observed in the second principal component (13.33% of the total variation) was attributed to average inter-node length of primary branches, average length of primary branches, canopy diameter, bean width and bean thickness. Quantitative traits, which had substantial contribution to the third principal component that accounted for 12.99% total variation, were total plant height, number of primary branches, and number of nodes of primary branches, leaf length and bean length. Similarly, variation in total plant height, coffee leaf rust, average inter-node length of main stem node and stem diameter had a great deal of contribution to the fourth principal component.

The variation in the fifth principal component was also, attributed to traits like number of secondary branches, number of nodes of primary branches, bean length and yield per tree. On the other hand, average inter-node length of main stem node, number of nodes of primary branches, average length of primary branches, fruit width and fruit thickness were predominantly influenced the variation in the sixth principal component. Moreover, quantitative traits like coffee berry disease, leaf length and number of nodes of primary branches had influenced the seventh principal component. Consistent with this finding many investigators also found comparable result from different Arabica coffee germplasm (Mesfin kebede and Bayetta Belachew, 2005; Olika Kitila *et al.*, 2011; Getachew WeldeMichael *et al.*, 2013; Lemi Beksisa and Ashenafi Ayano, 2016;Lemi *et al.*,2021)

The first and second principal components accounted for more than one third of the total variations (35.32%). Chahal and Goal (2002) inferred that characters with the largest absolute values closer to unit within the first principal component influence the clustering more than those with lower absolute values closer to zero. Therefore, in the current investigation discrimination of the coffee collections into different cluster was mainly due to average inter-node length of primary branches, average length of primary branches, canopy diameter, fruit width, fruit thickness, bean width, bean thickness and hundred-bean weight.

Table 7. Eigenvectors and Eigen values of the first seven principal components for 19 quantitative traits of 64 coffee collections evaluated at Metu in 2018.

Quantitative traits	PC						
	1	**2**	**3**	**4**	**5**	**6**	**7**
TPH	0.27	0.18	**0.32**	**0.37**	–0.06	–0.22	–0.03
AILMS	0.21	0.19	0.15	**0.31**	–0.27	**–0.40**	0.12
SD	0.05	0.23	0.28	**–0.40**	0.22	–0.16	–0.06
CD	0.24	**0.31**	0.15	–0.25	0.10	–0.18	0.08
NPB	0.18	0.09	**0.32**	0.28	0.02	0.06	–0.19
NSB	0.13	0.25	0.19	**–0.32**	**0.34**	0.15	0.12
NNPB	0.10	–0.12	**0.31**	0.24	0.29	**0.35**	**–0.31**
ALPB	**0.33**	**0.30**	–0.07	0.03	–0.02	**0.33**	–0.03
AILPB	0.24	**0.34**	–0.28	–0.14	–0.25	0.07	0.16
LL	0.12	0.26	**–0.30**	–0.17	–0.06	0.00	**–0.44**
FW	**0.36**	–0.23	0.04	–0.04	–0.16	**0.33**	0.08
FT	**0.34**	–0.22	0.06	–0.11	–0.18	**0.32**	0.13
BL	0.20	0.00	**–0.43**	0.14	**0.36**	–0.10	–0.07
BW	0.29	**–0.31**	0.01	–0.17	0.14	–0.05	0.19
BT	0.27	**–0.34**	–0.01	–0.06	0.23	–0.28	0.08
HBW	**0.32**	–0.16	–0.26	0.13	0.22	–0.29	0.05
CBD	–0.15	0.17	0.03	0.17	0.15	0.15	**0.73**
CLR	0.02	0.21	–0.31	**0.37**	0.24	0.23	0.04
YLD	0.15	–0.08	–0.08	–0.06	**-0.45**	0.00	0.01
Eigen values	**4.18**	**2.53**	**2.30**	**1.69**	**1.31**	**1.16**	**1.06**
Total variance (%)	**21.99**	**13.33**	**12.09**	**8.91**	**6.9**	**6.13**	**5.58**
Cumulative variance (%)	**21.99**	**35.32**	**47.42**	**56.33**	**63.23**	**69.35**	**74.94**

Note: *PC = Principal component; TPH = Total plant height; AILMS = Average inter-node length of main stem; SD = Stem diameter; CD = Canopy diameter; NPB = Number of primary branches; NSB = Number of secondary branches; NNPB = Number of nodes of primary branches; ALPB = Average length of primary branches; AILPB = Average inter node length of primary branches; LL = Leaf length; FW = Fruit width; FT = Fruit thickness; BL = Bean length; BW = Bean width; BT = Bean thickness; HBW = Hundred bean weight; CBD = Coffee berry disease; CLR = Coffee leaf rust; and YLD = Yield per tree.*

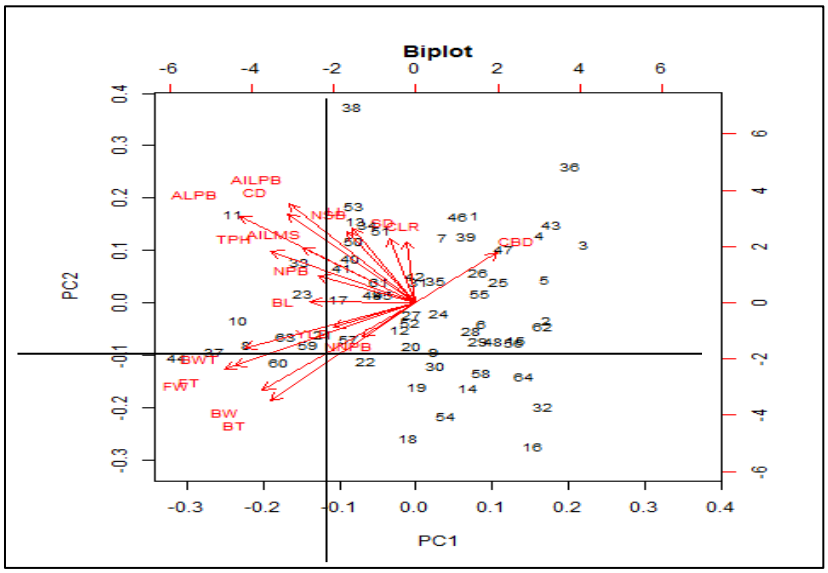

Figure 2. Biplot of the first two principal components for 64 Arabica Coffee collections.

4. Conclusions

Characterization of the coffee collections using a multivariate technique revealed the availability of sufficient genetic variability among coffee collections. The distribution pattern of the coffee collections into seven clusters demonstrated the presence of considerable genetic divergence among the tested collections. Hence, crossing of these collections in a breeding program would result in superior hybrids. Principal component analysis revealed that the first seven principal components were responsible for about 74.94 % of the total variation. Average internode length of primary branches, average length of primary branches, canopy diameter, fruit width, fruit thickness, bean width, bean thickness and hundred bean weights contributed to the observed variability among collections. The results imply that these traits could be used for selection of collections and crossing in coffee improvement program. Moreover, future evaluation of these coffee collections with respect to coffee quality traits would be indispensable using biochemical and molecular techniques.

5. Acknowledgments

The author acknowledges Jimma Agricultural Research Center and Ethiopian Institute of Agricultural Research for availing the resources and facilities for conducting the research. Special appreciation also goes to Metu Agricultural Research Sub-Center for material support, technical assistance and cooperation in the field during the entire period of data collection.

6. References

Amsalu Ayana and Endashaw Bekele. 1999. Multivariate analysis of morphological variation in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. *Genetic Resources and Crop Evolution*, 46: 273–284.

Bayetta Belachew, Ashenafi Ayano and Tadesse Benti. 1993. Screening of Arabica coffee collection for Bebeka environment. 21st International Conference on Coffee Science, 11th–15th September 2006. Montpelier, France.

Bayetta Belachew. 2001. Arabica coffee breeding for yield and resistance to coffee berry disease (*Colletotrichum kahawae*). PhD Dissertation, University of London, Britain.

CSA (Central Statistical Authority of Ethiopia). 2018. Area and production of major crops, Central Statistics Agency, Addis Ababa, Ethiopia.

Copper, M.C. and Milligan, G.W. 1988. The effect of error on determining the clusters. Proceedings on the International workshop on Data Analysis, Decision support and Expert Knowedge Representation in Marketing and Relayted Areas of Research, June 21–23,1987, University of Karlsruhe, WestGeremany. Pp. 319–328.

Chahal, G.S. and Gosal, S.S. 2002. *Principles and procedures of plant breeding: Biotechnology and conventional approaches*. Alpha Science International, United Kingdom.

Dabholkar, A.R. 1999. *Elements of Biometrical Genetics*. Concept Publishing Company, New Delhi.

Das, P.K. and Gupta, T.D. 1984. Multivariate analysis in Blackgram. *Indian Journal of Genetics*, 44: 243–247.

Davis, A.P., Tadesse Woldemariam, Susana, B. and Justin, M. 2012. The impact of climate change on indigenous Arabica Coffee (*Coffea arabica* L.): Predicting future trends and identifying priorities. *PLoS One*, 7(11):1–13.

Dereje Likissa. 2014. Assessing the impact of UNESCO biosphere reserves on forest cover change: The case of Yayu coffee forest biosphere reserve in Ethiopia. MSc Thesis, Center for Geo-Information, Wageningen University, The Netherlands. Pp. 102.

Ermias H/Mariam. 2005. Evaluation of Wellega coffee germplasm for yield, yield component and resistant to coffee berry disease at early bearing stage. MSc Thesis, Haramaya University, Ethiopia.

Esayas Aga. 2005. Molecular genetic diversity study of forest coffee tree (*Coffee arabica* L.) populations in Ethiopia: Implications for conservation and breeding. Doctoral, Dissertation Swedish University of Agricultural Science, Sweden.

Feyera Senbeta. 2006. Biodiversity and ecology of Afromontane rainforests with wild *Coffea arabica* L. populations in Ethiopia. *Ecology and Development Series* No. 38, Zentrum fur Entwicklungsforschung, Center for Development Research.

Fikadu Tefera, Bayetta Belchew, Behailu Atero, Ashenafi Ayano and Tadesse Benti. 2008. Germplasm collection and maintenance of coffee (*Coffee arabica* L.). Pp. 14–17. *In*: Girma Adugna, Bayetta Belachew, Tesfaye Shimber, Endale Taye and Taye Kufa (eds.). *Coffee Diversity and Knowledge. Proceedings of a National Workshop Four Decades of Coffee Research and Development in Ethiopia.* August 2007, Addis Ababa, Ethiopia.

Getachew WeldeMichael, Sentayehu Alamerew and Taye Kufa. 2013. Genetic diversity analysis of some Ethiopian specialty coffee (*Coffea arabica* L.) germplasm accessions based on morphological traits. *Time Journals of Agriculture and Veterinary Sciences,* 1: 47–54.

Gezahegn Berecha, Aerts, R., Vandepitte, K., Van Glabeke, S., Roldan-Ruiz, I. and Honnay, O. 2014. Effects of forest management on mating patterns, pollen flow and intergenerational transfer of genetic diversity in wild Arabica coffee (*Coffea arabica* L.) from Afromontane rainforests. *Biological Journal of the Linnean Society,* 112: 76–88.

Gray, Q., Abu Tefera and Teddy Tefera. 2013. Ethiopian coffee annual report. GAIN Report No.ET-1302. GAIN Report Assessment of Commodity and Trade by USDA.

IPGRI (International Plant Genetic Resource Institute). 1996. Diversity for development. International Plant Genetic Resources Institute, Rome, Italy.

Jagadev, P.N., Samal, K.M. and Lenka, D. 1991. Genetic divergence in rape mustard. *The Indian Journal of Genetics and Plant Breeding,* 51(4): 465–467.

Joshi, A.B. and Dhawan, N.L. 1966. Genetic improvement of yield with special reference to self-fertilizing crops. *Indian Journal of Genetics and Plant Breeding*, 26A: 101–113.

Kassahun Testate, Govers, K., Endashaw Bekele and Borsch, T. 2013. ISSR fingerprinting of *Coffea arabica* throughout Ethiopia reveals high variability in wild populations and distinguishes them from landraces. *Plant Systematics and Evolution*, 300(5): 881–897.

Lashermes, P., Combes, M.C. Ansaldi, C. Gichuru, E. and Noir, S. 2011. Analysis of alien introgression in coffee tree (*Coffe arabica* L.). *Molecular breeding*, 27(2): 223–232.

Lemi Beksisa and Ashenafi Ayano. 2016. Genetic variability, heritability and genetic advance for yield and yield components of Limmu coffee (*Coffea arabica* L.) accessions in South Western Ethiopia. *Middle-East Journal of Scientific Research,* 24 (6): 1913–1919.

Lemi Beksisa, Tadesse Benti, Getachew Weldemichael.2021. Phenotypic diversity of Ethiopian coffee (*Coffea arabica* L.) accessions collected from Limmu coffee growing areas using multivariate analysis. *American Journal of BioScience*, 9(3): 79–85.

Mahalanobis, P.C. 1936. On the generalized distance in statistics. *National Institute of Science,* 2: 49–55.

Mesfin Kebede and Bayetta Belachew. 2005. Genetic divergence of Hararge coffee (*Coffea arabica* L.) germplasm accessions at pre-bearing stage. *Proceedings of the 20th International conference on Coffee Science*. 11-15 October 2004, Bangalore, India. Pp. 1107–1112.

Meyer, F.G., Fernie, N.M., Narasimhaswamy, R.L., Monaco, L.C. and Greathead, D.J. 1965. Food and Agriculture Organization (FAO) coffee mission to Ethiopia, 1964/1965. Food and Agriculture Organization of the United Nations, Rome, Italy.

Mishra, MK. and Slater, A. 2012. Recent Advances in the genetic transformation of coffee. Review Article. *Biotechnology Research International*, 580857,doi:10.1155/2012/580857.

Murty, B.R .and Arunachalam, V. 1966. The nature of genetic divergence in relation to breeding system in crop plants. *Indian Journal of Genetics,* 26: 188–198.

Olika Kitila, Sentayehu Alamerew, Taye Kufa and Wosene Gebreselassie. 2011. Variability of quantitative traits in Limmu coffee (*Coffea arabica* L.) in Ethiopia. *International Journal of Agricultural Research,* 6: 482–493.

Paulos Dubale. 2001. Soil and water resources and degradation factors affecting their productivity in the Ethiopian highland agro-ecosystems. *Michigan State University Press*, 8(1): 1–18.

Paulos Dubale and Demel Teketay. 2000. The need for forest coffee germplasm conservation in Ethiopia and its significance in the control of coffee diseases. *Proceedings of the Workshop on Control of Coffee Berry Disease (CBD) in Ethiopia*. January 13-15, Addis Ababa, Ethiopia. Pp. 125–135.

Sharma, J.R. 1998. *Statistical and biometrical techniques in plant breeding*. New Age International (P) Limited Publishers, New Delhi. Pp. 432.

Singh, R.K. and Chaudhary, B.D. 1987. *Biometrical methods in quantitative genetic analysis*. Kalyani Publishers, New Delhi.

Sylvain, P.G. 1958. Ethiopian coffee and its significance to world coffee problems. *Economic Botany,* 12(2): 111–139.

Szamosi, C.I., Solmaz, N., Sari and Bársony, C. 2010. Morphological evaluation and comparison of Hungarian and Turkish melon (*Cucumis melo* L.) germplasm. *Scientia Horticulturae,* 124(2): 170–182.

Tadesse Benti, 2017. Progress in Arabica coffee breeding in Ethiopia: Achievements, challenges and prospects. *International Journal of Sciences: Basic and Applied Research,* 33: 15–25.

Tadesse Benti, Endale Gebre, Kassahun Tesfaye, Gezahegn Berecha, Lashermes, P., et al. 2021. Genetic diversity among commercial Arabica coffee (*Coffea arabica* L.) varieties in Ethiopia using simple sequence repeat markers. *Journal of Crop Improvement*: 35(2): 147–168.

Tadesse Woldemariam, Denich, M., Demel Teketay and Vlek, P.L.G. 2002. Human impacts on *Coffea arabica* L. genetic pools in Ethiopia and the need for its *in situ* conservation. Engels, J., Rao, M., Brown, V.R. and Jackson, M.T. (eds.). *Managing*

Plant Genetic Diversity, *IPGRI Publication*, Rome. Pp: 237–247.

Tadesse Woldemariam. 2003a. Vegetation of Yayu forest in SW Ethiopia: Impacts of human use and implications for in situ conservation of wild *Coffea arabica* L. populations. PhD Dissertation, Ecology and Development Series, No.10, Cuvillier Verlag, Gottingen.

Tadesse Woldemariam and Feyera Senebeta. 2008. Sustainable management and promotion of forest coffee in Bale, Ethiopia. Bale Eco-Region Sustainable Management Programme SOS Sahel/ FARM- Africa, Addis Ababa, Ethiopia.

Taye Kufa. 2006. Ecophysiological diversity of wild Arabica coffee populations in Ethiopia: Growth, water relations and hydraulic characteristics along a climatic gradient. PhD Dissertation, Ecology and Development Series, No.46.Cuvillier Verlag, Gottingen. Pp. 305.

Taye Kufa, 2010. Environmental sustainability and coffee diversity in Africa. ICO World Coffee Conference, Guatemala City. http://dev.ico.org/event_pdfs/wcc2010/presentationswcc-2010-kufa-notese-pdf. Accessed on 10 November 2011.

United States Department of Agriculture (USDA). 2020. Coffee annual report: Global Agricultural Information Network, USDA foreign Agricultural Service report Number: ET2020-0004.

Yigzaw Desalegn, 2005.Assessment of genetic diversity of Ethiopian Arabica coffee accessions using morphological, biochemical and molecular markers. A PhD Dissertation, University of the Free State, South Africa. Pp. 197.