

Phenotypic Diversity in the Hararge Coffee (*Coffea arabica* L) Germplasm for Quantitative Traits

Mesfin Kebede^{1*} and Bayetta Bellachew²

¹Awada Agricultural Research Center, P O Box 205, Yirgalem, Ethiopia

²Jimma Agricultural Research Center, P O Box 192, Jimma, Ethiopia

Abstract: A field experiment was conducted at Awada Agricultural Research Sub-Center, Ethiopia, to study the magnitude of phenotypic diversity among Hararge coffee (*Coffea arabica* L.) germplasm accessions based on quantitative traits. A total of 104 entries consisting of 100 accessions from Hararge and 4 standard cultivars were evaluated using nested design. Analyses of variance showed significant differences among the accessions for all the traits considered, indicating the presence of high genetic variability among the Hararge coffee germplasm accessions. Cluster analysis grouped the entries into 6 groups of different sizes, ranging from 5 entries in cluster VI to 44 in cluster III. The maximum inter-cluster distance was obtained between clusters II and VI while the minimum was observed between clusters I and III. Clusters I and V, I and VI, II and IV, II and V, II and VI, III and VI, IV and V and V and VI were significantly divergent. The first four principal components constituted 78.5 percent of the total variation prevalent within the germplasm accessions, while 38.5 percent was represented by the first principal component. The length of the longest primary branch, stem diameter, average length of primary branches, total number of internodes per plant and total number of primary branches per plant were the five important characteristics that contributed most to the total variation in the first principal component, implying that there is high potential to improve these traits through selection. The study revealed the presence of high genetic diversity among Hararge Coffee germplasm accessions and the possibility of developing improved varieties through selection and hybridization.

Keywords: Cluster Analysis; *Coffea arabica*; Genetic Diversity; Germplasm; Hararge; Quantitative Traits

1. Introduction

Ethiopia is well-known for being the home of arabica coffee which is highly-regarded for its very fine quality, unique aroma and flavor. The coffee types that are acclaimed for having such unique characteristics include Sidamo, Yirgacheffe, Hararge, Ghimbi and Limu (Workafes and Kasso, 2000). Since Ethiopia is the primary center of origin and genetic diversity for *C. Arabica*, there is high genetic variability for yield and yield components, disease and pest resistance, and other traits. This is substantiated by the fact that, within Hararge region itself, including the major coffee-producing districts such as Habro, Chercher, Wobera, Garamuleta, Harar Zuria and Gursum, which are known for the production of the best quality coffee (Bridge and Eyassu, 1968), there is high variability of yield and other characteristics. Furthermore, survey results in the past indicated the presence of considerable variations among coffee types such as Abadiro, Kubania, Shimbure, and Bunaqalla (Bayetta, 1987).

For any crop improvement program, a breeder depends on the variability present in the germplasm collections in order to advance in production, bring about stability in different biotic and abiotic stresses or changes in crop characteristics and meet breeding interest (IBPGR, 1987). In cognizant of this fact and in order to alleviate the production problems, concerted efforts were undertaken to collect coffee germplasm during 1998 from different coffee-growing areas in Hararge, eastern Ethiopia by Jimma Agricultural Research Center (JARC) in Ethiopia and, as a result, more than 900 accessions were collected and maintained at the center.

Several workers have estimated the extent of genetic diversity present from the different sources of arabica coffee germplasm collections. For instance, a study by Catter (1992) on second progeny arabica coffee collections of Ethiopian origin indicated the prevalence of a high level of variability in morphological, agronomic and biochemical characteristics. The genetic diversity analysis conducted by Lashermes *et al.* (1996) by employing RAPD markers on cultivated and sub-spontaneous accessions of arabica coffee confirmed the narrow genetic base of commercial cultivars (3 typica and 3 bourbon types). On the other hand, they also reported the existence of large genetic diversity within the sub-spontaneous material, which consisted of 11 samples representing the different coffee growing areas in Ethiopia. Furthermore, they have suggested the existence of east-west differentiation in the Ethiopian coffee germplasm.

Though there are indications of genetic variations in Hararge Coffee, such as the presence of a number of vernacular names (Bayetta, 1987), no systematic study has been carried out to quantify and verify the level of genetic diversity. In addition, it is necessary to extract detailed information about the individual accessions employed in the study so that they can be used in the ongoing breeding program. This was the impetus to conduct the present study, with the objective of estimating the genetic diversity among Hararge coffee germplasm accessions for quantitative traits and of facilitating their use in breeding programs.

*Corresponding author. E-mail: mesfin04@yahoo.com

2. Materials and Methods

The experiment was carried out at the Awada Agricultural Research Sub-Center, Ethiopia in the 2002 cropping season. Awada is characterized as a mid-altitude area with an altitude of 1750 meters above sea level, respective annual mean minimum and maximum rainfall of 858.1 mm and 1676.3 mm and annual mean minimum temperatures of 11.0°C and maximum 28.4°C. The major soil types at the research center are Eutric Nitosol and Chromic Cambisols that are highly suitable for coffee production.

A total of 928 Hararge coffee accessions collected from 16 districts in the Eastern and Western Hararge Zones of eastern Ethiopia were planted in July 2000 for maintenance. One hundred accessions taken at random

were considered for this study along with 4 coffee berry disease resistant cultivars as standard checks. A detailed description of the accessions is given in Table 1. Each of the accessions was planted in a single row of six plants using an augmented design with 29 blocks, where each block had 36 accessions including the 4 standard checks. A spacing of 1.5 m between plants and 2 m between rows was used. All field management practices were applied to all plots uniformly as recommended (JARC, 1996). Four plants were taken at random from each accession and labeled for data collection on different growth characters listed in table 2. Jima Agricultural Research Center's coffee breeding and genetics conventional methods were employed for data collection (Mesfin, 1982 and Bayetta, 2001).

Table 1. Details of germplasm accessions used in the study.

Serial No	Accession No.	Origin		Altitude range (masl)	
		Zone	District		
1	21298, 22098, 22898, 23498, 27098, 26798, 27498, 25298, 28898	East Hararge	Bedeno	1500-1900	
2	35198, 35398, 37598, 36798, 45998		Dedder	1400-1550	
3	15398, 13998, 13098, 16898, 14298, 17598, 20098, 20898, 11898, 16698, 17398, 18498,		Girawa	1500-1900	
4	1998, 4298, 5398, 5498		Gursum	1600-1800	
5	7098		Jarso	1300-1900	
6	8898, 9598, 9198, 10598, 9998, 8598		Kombolcha	1500-1700	
7	29798, 30998, 31398, 31698, 33398		Meta	1500-2000	
8	48998, 47398, 49698, 50398, 51198, 48198, 54398, 55098, 51698, 46698		Boke	1550-1700	
9	67298, 67598, 69098, 71398, 66798, 69298		Chiro	1600-1900	
10	92798, 98798, 94298, 99998, 98598, 97998, 94398		Darolabu	1300-1900	
11	45998		Doba	-	
12	81798, 88798, 87198, 86198, 88898, 88398, 90398		Habro	1600-1800	
13	55798, 56498		Hardim	-	
14	64198, 64498, 64098, 62198, 63198, 63498, 58798, 59398, 60798, 57498		Kuni	1700-1900	
15	73098, 74298, 75678, 80998, 75898, 76598, 77598, 77898, 79098		Mesela	1500-1800	
16	41298, 42698, 40198, 44498, 44598, 42498, 38298		Tulo	-	
17	F-59		Kaffa	Bonga	1650
18	74140, 74165		Illubabor	Metu	1550-1750
19	75227		Jima	Gera	1900

Analysis of variance was computed using nested design for each quantitative character in order to see the variability among accessions for each trait. Since the experiment was treated as nested design for the purpose of analysis, hierarchical classification was used for the partitioning of the variation into different sources of variations. The ANOVA was constructed by considering the experimental units (the four coffee trees within each accession) as factor B nested within levels of factor A (the 104 coffee accessions) (Sokal and Rolf, 1969). The data on quantitative characters was standardized to a mean of zero and a variance of unity before cluster and principal

component analyses were made to avoid differences in scales used to measure different traits.

Clustering was performed by average linkage method and the number of clusters was determined by examining the pseudo F statistic and the pseudo t^2 statistic using the SAS software package (SAS Institute, 2001). Genetic diversity between clusters, as standardized Mahalanobis D^2 values between clusters and principal components based on correlation matrix, were calculated using the same software employed in cluster analysis. The D^2 values obtained for pairs of clusters were considered as the calculated values of Chi-square (X^2) and were tested for significance both at 1% and 5% probability levels against

the tabulated values of X^2 for 'P' degree of freedom, where P is the number of characters considered (P=14 in the present case) (Singh and Chaudhary, 1996). The important traits in each principal component that significantly contributed to the variation observed were identified as suggested by Johnson and Wichern (1988).

3. Results and Discussion

3.1. Analysis of Variance

Mean squares due to treatments were highly significant for all the 14 characters considered, suggesting the

presence of high variability among the accessions (Table 2). In view of this, it may be reasonable to state that there is a good chance of improving Hararge coffee accessions through selection and breeding. The prevalence of such a high variability in an autogamous species like *C. arabica* appears to be significant. This may be attributed either to evolutionary tendencies, as the species is indigenous to Ethiopia, or to the natural mutations occurring to the population of the crop (Avice and Hamric, 1997; Hedrick, 2000).

Table 2. Mean squares from the analysis of variance for 14 quantitative traits.

Characters	Mean squares (MS)		Broad sense heritability values
	Treatments (MT)	Error (ME)	
Plant height (cm)	912.903**	178.046	50.8
Internodes length of stem (cm)	0.989**	0.297	36.9
Internode length of branch (cm)	15.478**	0.418	90.0
Number of internodes of stem	21.630**	5.855	40.3
Number of internodes on the longest primary branch	18.408**	7.986	24.6
Total number of internodes per plant	22106.285**	7492.994	32.8
Canopy diameter (cm)	696.938**	131.200	51.9
Stem diameter (cm)	0.465**	0.199	25.1
Leaf area (cm)	177.327**	46.202	41.5
Number of primary branches	87.553**	23.878	40.0
Angle of primary branches from the main stem (in degrees)	27.262**	10.252	29.3
Number of secondary branches	8375.144**	2355.480	39.0
Length of the longest primary branch (cm)	299.536**	73.424	43.5
Average length of primary branches (cm)	111.770**	34.231	36.2

** Significant at 0.01 probability level

MS=mean squares, MT=mean squares of treatments and ME= mean squares of error.

Note: degrees of freedom for treatments and error for all the 14 characters were the same i.e. 103 and 312, respectively.

3.2. Cluster Analysis

The 104 coffee germplasm accessions were grouped into 6 clusters (Table 3). The size of cluster varies from 5 accessions in cluster V to 44 accessions in cluster III. Clusters I, II, and IV contained accessions mainly from the Western Hararge districts whereas clusters III and V had an almost equal number of accessions from both east as well as West Hararge districts. The five accessions in cluster VI were from the two districts of West Hararge, out of which 4 originated in Kuni and only one in Chiro District. Three of the coffee berry disease (CBD) resistant cultivars (75227, 74165 and 74140) used as checks were grouped in cluster I where middle- to high-altitude accessions from Western Hararge districts was most frequent. The fourth check, F-59, was grouped in cluster II, confirming the fact that this cultivar was distinctly different from the rest of the standard checks in morphology and geographical origin. Lin and Binns (1985) and Lin *et al.* (1986) also highlighted the advantages of hierarchical cluster analysis in identifying useful germplasm, particularly by including reference cultivars.

It was evident that the accessions from the Eastern Hararge districts showed close similarity (Table 3) with regard to their clustering patterns. For instance, the germplasm accessions from Gursum, Bedeno and Dedder Districts were found to be distributed in clusters II and III. On the other hand, accessions from Kombolcha, Girawa and Meta were scattered in clusters I, II and III where the majority of their accessions were grouped in cluster III. In general, cluster III represented 58.5 per cent of the germplasm accessions from Eastern Hararge districts. Similarly, more than 65 per cent of the germplasm accessions from Darolabu, Mesela and Tulo Districts of Western Hararge were concentrated in cluster III.

Accessions from Habro and Boke Districts appeared in the same clusters i.e. clusters I, II, and III, even though the majority of their accessions appeared in the first two clusters. The germplasm accessions of Girawa, Bedeno, Kuni, Chiro, Mesela and Habro Districts were distributed in four different clusters, which suggested that the germplasm accessions from these districts were relatively more variable. In respect to the remaining districts, the accessions were distributed in 2 or 3 clusters, probably

reflecting less variation among germplasm accessions within a particular district.

The overlapping of clustering patterns with regard to the germplasm accessions in the majority of the districts could be explained as lack of differentiation among districts, probably arising partly due to gene flow (Amsalu and Endashaw, 1999). In general, it may be possible to state that germplasm accessions from the Western Hararge districts were relatively more variable in their clustering patterns compared to those from the Eastern Hararge districts (Table 3). This pointed out that in future Hararge coffee germplasm exploration endeavors, due emphasis must be given to the Girawa, Bedeno, Kuni, Chiro, Mesela and Habro districts.

3.3. Distance Analysis

Based on Mahalanobis's D^2 statistics, highly significant inter-cluster distances were obtained. Cluster II showed the maximum and significant genetic distance (102.12) from cluster VI. Furthermore, the inter-cluster distances between clusters I and V, I and VI, II and IV, II and V, II and VI, III and VI, IV and V, and V and VI in that order were found to be highly significant (Table 4). These

distances indicated that germplasm in the above paired clusters are significantly ($p = 0.01$) divergent from each other. Since the magnitude of heterosis largely depends upon the degree of genetic diversity among the parental lines, the progenies of those germplasm accessions belonging to the pairs of distant clusters could be very useful in a hybridization program for obtaining a wide spectrum of variation among the segregates. Crossing of parental lines extracted from germplasm accessions belonging to different clusters of wide Mahalanobis distance (D^2) could maximize opportunities for transgressive segregation as there is a higher probability that unrelated genotypes would contribute unique desirable alleles at different loci (Peters and Martinelli, 1989). Therefore, it may be possible to conclude that the germplasm accessions from cluster II and cluster VI could offer relatively better potential parental lines that, when intercrossed, could produce hybrids with maximum heterotic value, even though other clusters with significant genetic distances are also good sources of parental lines.

Table 3. Distribution of the 104 coffee genotypes over six clusters based on quantitative traits.

Zone	Cluster						Total accessions
	I	II	III	IV	V	VI	
East Hararge	6	6	24	2	3	-	41
West Hararge	11	13	20	7	3	5	59
*South west Ethiopia	3	1					4
Total	20	20	44	9	6	5	104

*Represented standard checks

Note: This table was extracted from the dendrogram

Table 4. Inter-cluster distances among 104-coffee genotypes.

	Cluster I	Cluster II	Cluster III	Cluster IV	Cluster V	Cluster VI
Cluster I	-					
Cluster II	14.91634	-				
Cluster III	10.78772	16.56251	-			
Cluster IV	14.25636	47.83423**	22.36447	-		
Cluster V	38.84501**	66.12385**	18.28813	29.39752**	-	
Cluster VI	47.52314**	102.12226**	68.19061**	16.35693	59.78586**	-

** = Significant at $p < 0.01$ ($X^2 = 29.141$)

3.4. Principal Component Analysis

The first four principal components represented 78.5 per cent of the total variation (Table 5). Principal component 1 accounted for more than one third of the variation. The length of the longest primary branch, stem diameter, the average length of primary branches, the total number of internodes per plant and the total number of primary branches per plant were the most important factors contributing to the total variation of the first principal component. In the second principal component, the internode length of the stem, the leaf area, the total number of internodes per plant, the number of internodes on the stem, the number of primary branches

per plant and the average internode length of primary branches made a significant contribution.

In light of the results obtained from principal component analysis, it may be possible to deduce that the maximum variation (38.5%) of principal component 1 was based on quantitative characters such as the length of the longest primary branch, the stem diameter, the total number of internodes per plant and the total number of primary branches per plant. This perhaps emphasizes the significance of these characteristics to the appraisal of genetic diversity.

Table 5. Eigenvalues, total variance, cumulative variance and eigenvectors for the 14 quantitative traits.

Characters	PC 1	PC 2	PC 3	PC 4
Plant height	-0.283	0.160	0.483	0.080
Inernode length of stem	-0.086	0.485	0.309	-0.077
Internode length of branch	0.027	-0.281	0.209	-0.633
Number of internodes of stem	-0.288	-0.299	0.364	0.150
Number of internodes on the longest primary branch	-0.264	-0.191	-0.399	0.133
Total number of internodes per plant	-0.343	-0.310	-0.010	0.172
Canopy diameter	-0.310	0.133	-0.160	-0.322
Stem diameter	-0.358	0.097	-0.129	0.023
Leaf area	0.039	0.462	0.143	0.006
Number of primary branches	-0.313	-0.285	0.313	0.050
Angle of primary branches from the main stem	0.083	0.077	0.070	0.626
Number of secondary branches	-0.229	0.136	-0.398	-0.006
Length of the longest primary branch	-0.365	0.213	-0.086	0.019
Average length of primary branches	-0.354	0.217	-0.064	-0.136
Eigenvalues	5.383	2.669	1.8642	1.079
%Total variance	38.50	19.10	13.30	7.70
%Cumulative variance	38.50	57.50	70.80	78.50

Note: PC1, PC2, PC3 and PC4 are the first four principal components with Eigenvalues greater than unity

Table 6. Grouping of Hararge coffee accessions into different diversity classes.

Cluster	Number of Accessions	Accessions included in the cluster
I	20	56498,55098,88398,45998,13998,59398,30998,13098,9198,74198,74198,88798,20898,48998,46698,74298,75227,51198,51698,16898
II	20	1998,54398,66798,81798,90398,10598,49698,99998,88898,11898,64498,48198,80998,55798,57498,F-59,50395,21298,15398,8598
III	44	27098,17398,42498,35198,33398,23498,22898,42698,98798,41298,35398,76598,9598,8898,92798,29798,87198,9998,37598,94398,20098,71398,7098,97998,98598,22098,47398,31398,17598,25298,26798,75898,75698,73098,5398,14298,4298,67598,5498,40198,77598,18498,44498,79098
IV	9	86198,69298,94298,31698,64198,63198,69098,28898,62198
V	6	44598,38298,36798,77898,27498,16698
VI	5	63498,60798,58798,67298,64098

Table 7. Cluster means for the 14 quantitative traits of 104 coffee germplasm accessions.

Traits	Cluster					
	I	II	III	IV	V	VI
Plant height	128.00	122.85	129.36	151.53	128.58	147.25
Inernode length of stem	4.74	5.17	5.11	5.26	4.79	4.96
Internode length of branch	6.34	5.26	4.79	4.84	4.76	4.57
Number of internodes of stem	21.63	18.63	19.97	24.17	20.92	24.70
Number of internodes on the longest primary branch	20.85	18.14	20.57	21.61	23.54	25.25
Total number of internodes per plant	452.06	340.39	412.07	524.89	492.58	621.60
Canopy diameter	101.06	96.01	110.80	115.50	124.83	115.18
Stem diameter	3.22	3.04	3.48	3.71	3.64	3.80
Leaf area	47.78	54.02	50.80	49.25	46.98	47.30
Number of primary branches	39.01	32.86	36.48	44.19	38.89	44.40
Angle of primary branches from the main stem	63.81	63.37	64.65	63.28	63.58	64.50
Number of secondary branches	76.90	66.51	139.22	112.08	204.29	112.65
Length of the longest primary branch	77.36	71.21	82.19	88.58	85.08	89.75
Average length of primary branches	53.65	50.29	56.76	59.97	59.00	59.95

4. Conclusion

It must be acknowledged that Hararge coffee has limited geographical significance. In view of this, the germplasm accessions considered in the present study represented collections from the Eastern and Western areas of Hararge and these were appraised at pre-bearing stage only. It is however, necessary for different characteristics to be studied with additional accessions over several bearing years. Furthermore, other traits of interest and molecular techniques may be very useful in order to confirm the present encouraging result that indicated the presence of considerable variations within Hararge coffee populations and provides immense potential for the development of improved varieties from the local landraces in the area.

If further studies are undertaken, consideration of yield and pest reactions must receive due attention. To this end, the study pointed out that Western Hararge appeared to be the target area for the future intensive germplasm exploration endeavors of Hararge Coffee. In the meantime the evaluation of Hararge coffee germplasm collections for yield, quality and disease resistance must continue to provide improved cultivars for coffee growers in the region in the shortest time possible to minimize the risk of losing smallholder coffee orchards challenged by the severe competition with chat (*Catha edulis*). In fact, this program could be attached to the "Local Landrace Development Program for Arabica coffee" launched by Jimma Agricultural Research Center. The local landrace development approach is advantageous with regard to maintaining the typical quality of Hararge Coffee and may avoid adaptation problems faced by the released coffee berry disease-resistant cultivars of Southwest Ethiopian origin. Moreover, local cultivars are generally preferred by the smallholder local farmers over improved cultivars introduced from other areas.

5. References

- Amsalu, A. and Endashaw, B. 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.
- Avise, J.C. and Hamric, J.L. 1997. Conservation Genetics: Case histories from nature. Chapman and Hall Cop., New York.
- Bayetta, B. 1987. Coffee (*Coffea arabica* L.) Genetic erosion and germplasm collection in Hararge region. Plant Germplasm Resources Center/Ethiopia. Institute of Livestock Center for Africa (PGRC/E.I.LCA). *Germplasm Newsletter* 15: 8 - 13.
- Bayetta, B. 2001. Arabica Coffee Breeding for Yield and Resistance to Coffee Berry Disease (*Colletotrichum Kabayae* sp.nov.), Ph.D. Thesis, Imperial College at Wye University of London.
- Brown bridge, J.M. and Eyassu, G. 1968. The quality of some of the main Ethiopian mild coffees. *Turrialba* 18: 361-372.
- Catter, R. 1992. Study and structure of the phenotypic variation of *Coffea arabica* from Ethiopia. Tropical Agriculture Data Base. pp.51.
- Hedric, P.W. 2000. *Genetics of Populations*, 2nd edition. Johns and Bartlett Pub. New York.
- IBPGR (International Board for Plant Germplasm Resources). 1987. Annual Report for 1987. Rome, Italy.
- JARC (Jimma Agricultural Research Center). 1996. Recommended Production Technologies for Coffee and Associated Crops. Institute of Agricultural Research, Addis Ababa. p.18
- Jonson, R.A. and Wichern, D.W. 1998. *Applied Multivariate Statistical Analysis*. Prentice Hall, Englewood Cliffs, NJ.
- Lashermes, P., Trouslot, P., Anthony, F., Combs, M.C. and Charier, A. 1996. Genetic diversity for RAPD markers between cultivated wild accessions of *Coffea arabica*. *Euphytica* 87: 59-64.
- Lin, C.S. and Binns, M.R. 1985. Procedural approach for assessing cultivar's location data: pair wise genotype-environment interactions of test cultivars with checks. *Canadian Journal of Plant Science* 65:1065-1071.
- Lin, C.S. and M.R. Binns and L.P. Lefkovich. 1986. Stability analysis: where do we stand? *Crop Science* 26:894-900.
- Mesfin, A. 1982. Heterosis in crosses of indigenous coffee (*Coffea arabica* L) selected for yield and resistance to Coffee berry disease (CBD). At first bearing stage. *Ethiopian Journal of Agricultural Science* 4: 33-43.
- Peters, L.P. and Martinelli, J.A. 1989. Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theoretical and Applied Genetics* 78:42-48.
- SAS Institute. 2001. SAS user's guide: *Statistics*, 5th edition. SAS Inst., Cary, NC.
- Singh, R.K. and Chaudhary, B.D. 1996. *Biometrical Methods in Quantitative Genetic Analysis*. Kalyani Publishers, New Delhi. p. 318.
- Sokal, R.R. and Rohlf, F.J. 1969. *Biometry: The Principles and Statistics in Biological Research*. Freeman and Comp., San Fransisco, USA. p. 776
- Workafes, W. and Kassu, K. 2000. Coffee production system in Ethiopia. In: Proceedings of the Workshop on Control of Coffee Berry Disease in Ethiopia. 13-15th August, 1999, Addis Ababa, Ethiopia.