Journal of Genetics, Genomics & Plant Breeding 8(3) 63-72 (July, 2024)
ISSN (Online): 2581-3293

RESEARCH ARTICLE

Genetic diversity and principal component analysis of bean quality traits in hararghe coffee (*Coffea arabica* l.) collections Eastern Ethiopia

D. Deresa [1], W. Mohammed [2], A. Sualeh [3]

1 Oromia Agricultural Research Institute, Mechara Agricultural Research Center, P.O. Box 19, Mechara, Ethiopia

2 School of Plant Sciences, Haramaya University, P O Box 219, Haramaya, Ethiopia

3 Ethiopian Institute of Agricultural Research, Jimma Agricultural Research Center, P.O. Box 192, Jimma, Ethiopia

Corresponding authors email: derejederesag@gmail.com

Manuscript received: June 13,2024; Decision on manuscript: June 30,2024; Manuscript accepted: July11, 2024

Abstract

Coffee is the most valued commodity among the stimulant crops. Hararghe coffee fetches premium prices in world market for its best quality. This research was conducted with the objectives of assessing the genetic diversity of Hararghe coffee collections for bean quality traits. In line with objectives, Sixty-one coffee collections along with three standards evaluated at Jimma Agricultural Research Center. The analysis of variance revealed a significant difference among the collections for the most of quality traits studied. Cluster analysis based on bean quality traits grouped the collection into five clusters. The Euclidean genetic distance between clusters ranged from 13.52 to 60.28 among evaluated Hararghe coffee collections for bean quality traits. The Principal component analysis for bean quality traits showed five principal components exhibited greater than one Eigenvalue accounted 85.1% of the total variation. High inter cluster distance indicated that there is a high chance for obtaining transgressed segregates and maximizing heterosis by crossing genotypes belonging to different clusters as there is a higher chance that distinct genotypes would contribute by unique desired traits. Generally, the result of the study indicates existence of significant genetic diversity among Hararghe coffee collection for bean quality traits.

Keywords: Coffee, cluster analysis, diversity, PCA, quality

Introduction

Coffee (*Coffea arabica* L.) belongs to the family Rubiaceae and the genus Coffea. Currently, around 124 coffee species belonged to Coffea (Davis *et al.,* 2011). However, *Coffea arabica* L. and *Coffea canephora* Pierre ex A. (Robusta) are the only two dominant species providing over 60% and 40% of world's production to the international coffee trade, respectively (Farah and Santos, 2014). Arabica coffee, despite its name originated in Ethiopia, it has wild populations as an understory tree in the forests of southwest and southeast Ethiopia (Tesfaye, 2006).

In Ethiopia, coffee is a backbone of the economy contributing 25-30% of foreign exchange earnings, about 25% of the total population is dependent on production, processing, distribution and exporting of coffee and It is the country's number one source of export revenues with 50% of the production consumed locally (Atalaye, 2017; CSA, 2020).

Ethiopia is the largest coffee producer, consumer in Africa and the country is globally among the top ten coffee consuming countries (ICO, 2020), however, the country has not yet fully exploited its potential as the producer due to many factors. The Hararghe coffee natural sundried and unique mocha quality coffee fetches premium prices in the world market since it has special market or consumers. The Hararghe coffee production is challenged by competition of other cash crop (Khat), diseases and pests, lack of improved variety, moisture stress and the climate change which has profoundly negative trends in the future production of Hararghe coffee (Tolera and Gerbemedin, 2015; Jima, 2020). Coffee breeding researches in Ethiopia have released 42 coffee varieties, about 12654 coffee collections are collected, and ex-situ conserved at the research centers and Institute of biodiversity conservation in field gene banks of Ethiopia (Taye, 2017). Limited systematic genetic variability studies were conducted in Hararghe coffee collections. However, the studies did not include all Hararghe coffee collections available at Mechara Research Agricultural Center. The genetic diversity and principal component study for bean quality traits have vital importance to keep the Hararghe coffee in production with its unique quality characteristic feature. Thus, this research was initiated with objectives of study genetic diversity of Hararghe coffee collections for bean quality traits.

Materials and methods

The experiment was conducted at Mechara agricultural Research Center in 2021/22 cropping season. Mechara Agricultural Research Center latitudinal and longitudinal positions are 40^{0}19'11 "North and 08^{0}35'59" East, respectively. Its altitude is 1760 m.a.s.l. with annual mean minimum and maximum air temperatures are 14 °C and 26 °C, respectively. The area receives mean annual rainfall of about 963 mm with erratic rainfall distribution. The major soil texture of the research center is sandy loam with reddish color, soil type is nitisol, and pH is 6.7-7 (McARC, 2022). Sixty-one collections along with three released varieties used for this study were established at Mechara Agricultural Research Center. The quality of raw beans was determined for the following traits: bean size, bulk density of green coffee, shape and make, color, odor, bulk density of roasted coffee, weight loss due to roast, volume change due to roast. The cupping form provides a systematic means of recording eight important quality traits. These include aromatic intensity, aromatic quality, acidity, astringency, bitterness, body, flavor, and overall standard quality as standard procedures described. The collected subjected to analysis of variance using R- software and Mean values comparison was determined by DMRT at 1% probability level.

The principal component based on a correlation matrix was calculated using XLSTAT software.

Genetic distance and cluster Analysis

Euclidean distance (ED) computed from all significant difference traits of coffee collections used for ED as;; $\mathrm{EDjk} = \sqrt{\sum_{i=1}^{n}(Xij - Xik)^2}$ (Sneath and Sokal, 1973)

Table 1: Eigenvalues, eigenvectors, and to the variability Hararghe coffee collections

Traits	PC1	PC2	PC3	PC4	PC5
Bean size	-0.032	-0.195	0.076	0.118	0.576
Shape and make	-0.079	-0.248	0.362	-0.012	0.345
Color	-0.049	-0.295	0.367	-0.211	-0.111
Total raw quality	-0.071	-0.304	0.406	-0.125	0.126
Green bean volume	0.139	-0.081	-0.228	-0.614	0.099
Bulk density green bean	-0.139	0.083	0.217	0.621	-0.090
Roasted bean volume	0.083	0.405	0.035	-0.117	0.275
Roasted bean weight	-0.121	-0.260	-0.349	0.132	0.183
Bulk density of roasted bean	-0.104	-0.408	-0.107	0.155	-0.221
Volume changed due to roast	0.025	0.419	0.143	0.108	0.227
Weight loss due to roast	0.120	0.260	0.349	-0.132	-0.183
Aroma intensity	0.268	0.020	0.176	0.000	0.051
Aroma quality	0.269	0.027	0.139	0.008	0.196
Astiregence	0.255	-0.056	0.173	0.019	-0.270
Bitterness	0.265	-0.079	0.085	-0.007	-0.338
Body	0.280	-0.073	-0.195	0.086	0.133
Acidity	0.319	-0.083	-0.105	0.102	0.016
Flavor	0.310	-0.040	-0.147	0.148	0.065
Over all standard	0.330	-0.083	-0.078	0.174	0.085
Total cup quality	0.360	-0.060	-0.009	0.090	0.002
Total coffee quality	0.323	-0.178	0.152	0.038	0.052
Eigenvalue	7.47	4.39	2.81	1.76	1.43
Variability (%)	35.58	20.90	13.37	8.37	6.81
Cumulative %	35.58	56.48	69.85	78.22	85.03

Table 2: Percentage contributions of bean quality traits of 64 Hararghe coffee collections

Traits	PC1	PC2	PC3	PC4	PC5
Bean size	0.10	3.79	0.57	1.38	33.22
Shape and make	0.63	6.18	13.09	0.01	11.93
Color	0.24	8.69	13.44	4.47	1.22
Total raw quality	0.50	9.22	16.52	1.57	1.59
Green bean volume	1.93	0.66	5.19	37.67	0.97
Bulk density green bean	1.93	0.69	4.73	38.51	0.81
Roasted bean volume	0.68	16.39	0.12	1.38	7.54
Roasted bean weight	1.45	6.78	12.21	1.74	3.35
Bulk density of roasted bean	1.09	16.63	1.14	2.39	4.88
Volume changed due to roast	0.06	17.56	2.06	1.17	5.17
Weight loss due to roast	1.45	6.79	12.21	1.73	3.35
Aroma intensity	7.19	0.04	3.09	0.00	0.26
Aroma quality	7.245	0.08	1.94	0.01	3.83
Astiregence	6.51	0.32	2.99	0.04	7.3
Bitterness	7.03	0.62	0.72	0.01	11.4
Body	7.86	0.54	3.82	0.75	1.76
Acidity	10.18	0.68	1.10	1.03	0.03
Flavor	9.6	0.16	2.15	2.19	0.43
Over all standard	10.9	0.68	0.60	3.01	0.72
Total cup quality	12.97	0.36	0.01	0.81	0.00
Total coffee quality	10.46	3.17	2.30	0.15	0.27

In the bipolt, the traits found near the origin have smaller loading and influence little, while those found far from the origin have higher loading and great influence in this classification. From finger 1, total coffee quality, total cup quality, acidity, overall standard, body, bulk density of roasted bean ,volume changed due to roast and acidity have higher loading and more influence than other traits in collections classification. When this biplot is examined (Fig. 1), there is a positive relationship between the narrow angle features of traits (traits in same quadrant), most of cup quality traits in fourth quadrant, for raw bean quality in third quadrant, weight loss due to roast and volume changed due to roast in first quadrant. Wide-angle features of traits have negative relationships (traits found opposite quadrants) with each other's for instance between bulk density of green coffee bean with total coffee quality, total cup quality, acidity and overall standard quality. Right angle features of traits indicated that they not related to each other for instance cup quality traits with volume changed due to roast such type is patterns of relationship. The biplot technique enables the determination of the relationships between the variables as well as the detailed description of a multivariate data set (Yan and Rajcan, 2002).

The biplot classify the collections with phenotypic traits explained by first two dimensions PC1 and PC2 in figure 2. In all four quadrants, collections distributed based on their similarity and difference by 21 traits. The collections found in top right quadrant were similar their volume changed due to roast, roasted bean volume and weight loss due to roast. While collections found in top left quadrant were similar by bulk density of green coffee beans. Collections grouped in right bottom have similar traits such as acidity, flavor, overall standard, body, total cup quality, and total coffee quality. Collections grouped in left bottom similar by their total raw bean quality and bulk density of roasted coffee beans, shape and make, bean size and color. Collections such as H-574/02, H-961/98, H-109/02, H-63/02, H-1561/02, and H-26/02 were most distant or diverging from the major group in the PC axes, which were concentrated near the origin. Collections, which overlapped in the principal axes, have similar traits.

Fig 1: Principal component biplot of 21 bean quality traits of the Hararghe coffee collections

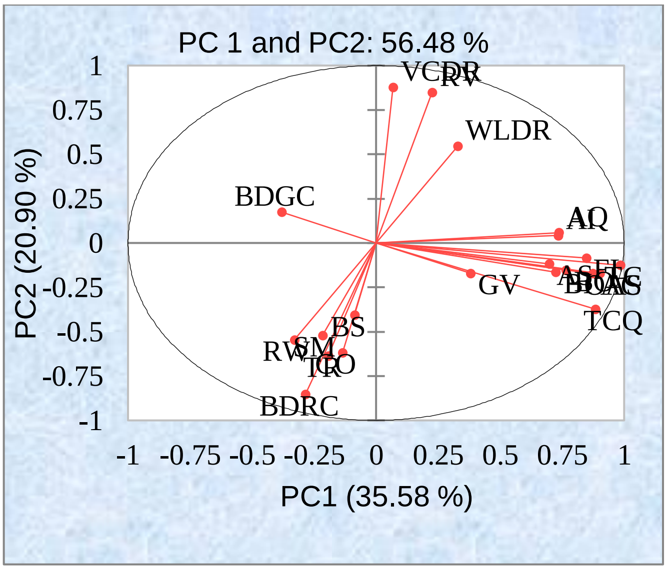
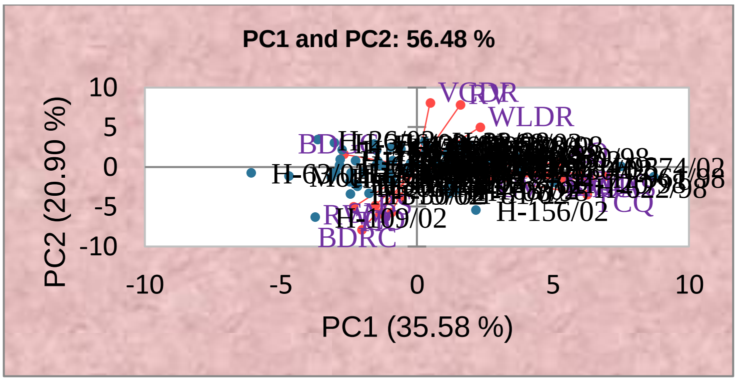


Fig 2: Principle component plots of illustrating variation among collections and their discriminating ability



Cluster analysis of Hararghe coffee collections based on 21 bean quality traits

The genetic cluster analysis of coffee collections based on bean quality traits grouped into five clusters as presented in table 3 and figure 3. The largest numbers of collections grouped in clusters I, II and V. whereas, small numbers of collections grouped in cluster III and IV. Accordingly, cluster-I was the largest and consisted of twenty-five collections (39.06%) followed by cluster II (29.69%), cluster-V consisted fourteen collections (21.88%). While, cluster IV contained small number of collections (1.56%) and cluster III contained five collections (7.81%) as presented in table 3 and figure 3. The clustering pattern of the collections revealed the existence of genetic diversity in the coffee collections of the traits considered. These findings agree with Dadi (2016), which classified 60 Hararghe coffee collections into eight clusters depending on coffee quality traits. More recently, Sorse (2018) classified 28 Hararghe coffee collections into nine clusters depending on coffee quality traits. The Clustering patterns of 49 coffee collections based on eight organoleptic traits grouped into three clusters. In additions, the clustering pattern of collections revealed the existence of moderate genetic diversity in coffee genotypes for organoleptic quality traits (Getachew *et al.*, 2015). Similarly, Yigzaw (2005) reported that cluster analysis based on coffee quality traits grouped 42 coffee collections into two main clusters and Olika *et al.*, (2011) also reported that the cluster analysis grouped 49 coffee collections into three clusters. Therefore, the present study confirmed the presence of ample genetic diversity among coffee collections collected from East and West Hararghe by cluster analysis.

Cluster mean of bean quality traits

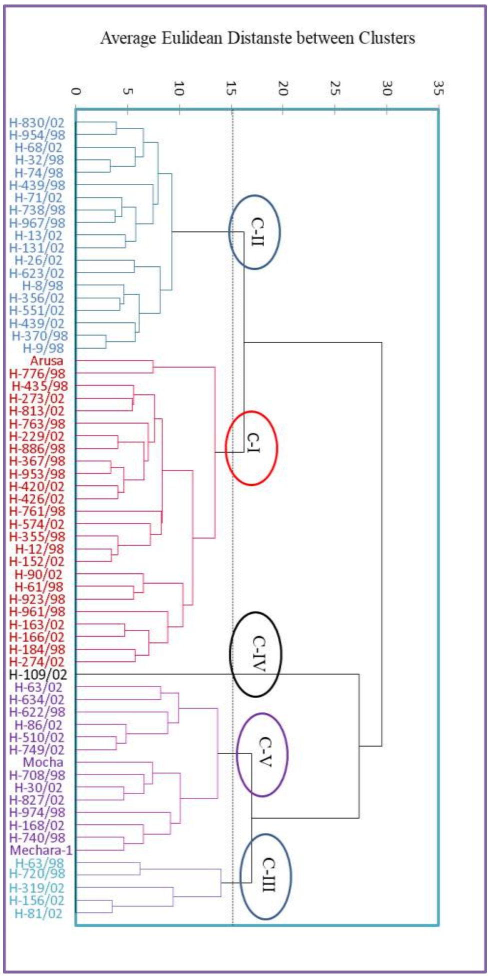
The clusters mean based on bean quality traits of 64 Hararghe coffee collections of five clusters presented in Table 4. Cluster I contained largest number of collections (39.06%). In additions, this cluster had lower bean size, and shape and make than other clusters and intermediate mean values for remains of traits as compared to other clusters. Cluster II contained 29.69% collection that distinguished for its unique characteristics of higher than other clusters for, volume of roasted bean (231.93),volume change due to roast (55.88%) and weight loss due to roast (10.49%). While, lowest mean value than other clusters for color (12.58), total raw quality (35.31),weight of roasted coffee bean (89.58), bulk density of roasted bean(0.4 g/ml), and body (7.0) and all remains traits medium mean values.

Journal of Genetics, Genomics & Plant Breeding **8(3) 63-72 (July, 2024)**
ISSN (Online): 2581-3293

Table 3: Distribution of 64 Hararghe coffee collections in to five clusters

Cluster	Number of genotypes per cluster	Name of Hararghe coffee genotypes per each cluster				
I	25(39.06%)	H-12/98	H-435/98	Arusa	H-152/02	H-163/02
		H-184/98	H-229/02	H-273/02	H-274/02	H-355/98
		H-420/02	H-426/02	H-574/02	H-61/98	H-761/98
		H-776/98	H-813/02	H-886/98	H-90/02	H-923/98
		H-961/98	H-166/02	H-367/98	H-763/98	H-953/98
II	19(29.69%)	H-32/98	H-370/98	H-8/98	H-13/02	H-131/02
		H-356/02	H-439/02	H-439/98	H-551/02	H-623/02
		H-71/02	H-738/98	H-74/98	H-830/02	H-9/98
		H-967/98	H-26/02	H-68/02	H-954/98	
III	5(7.81%)	H-63/98	H-156/02	H-319/02	H-720/98	H-81/02
IV	1(1.56)	H-109/02				
V	14(21.88%)	H-168/02	H-30/02	H-510/02	H-622/98	H-63/02
		H-708/98	H-740/98	H-749/02	H-827/02	H-86/02
		Mechara-1	Mocha	H-634/02	H-974/98	

Fig 3: Dendrogram generated based on UPGMA clustering method depicting genetic diversity among Hararghe coffee collections based on bean quality traits.



Cluster III contained 7.81% of collections characterized by higher for volume of green coffee bean (156.98%), overall standard (7.43), body (7.5), acidity (8.0), flavor (7.47), total cup quality (45.6%), and total coffee quality (81.47 %). In addition, lowest overall clusters by bulk density of green bean (0.64 g/ml) and astringency (3.73%) while other traits exhibited intermediate mean values. Cluster VI contained only one collection and distinguished for its unique characteristics of lowest mean values than other cluster for overall standard quality, body, acidity, flavor, total cup quality, weight loss due to roast, volume of roasted bean, volume change due to roast, weight loss due to roast, aroma intensity, and bitterness. In additions, it exhibited highest mean values for color, shape, and make, total raw quality, weight of roasted bean and bulk density of roasted coffee beans while medium mean values for the remained traits. Cluster V, which contained 21.88% of collections, distinguished for its traits of high mean values than other clusters for bean size (95.6) bulk density of green coffee beans (0.68) and astringency(3.85) and lower mean values than other clusters for volume of green coffee beans (146.7) and aroma quality (3.94%) while signified medium mean value for the remained traits. From previous work, (Dadi, 2016)) reported clustering of 64 coffee collections for bean quality traits, noted that dissimilar clusters has different breeding values that enable breeders to improve different traits and parental selection. The present study is more in harmony with the finding of Sorse (2018) who conducted clustering of 28 Hararghe coffee collections based on bean quality traits and revealed high variation for various traits and the cluster means revealed that the clusters different for the different traits measured.

Table 4: Cluster mean values of five clusters collections based bean quality traits

Traits	**Cluster-I**	**Cluster-II**	**III**	**IV**	**V**
Bean size	93.64	94.27	94.85	94.70	95.60
Color	12.69	12.58	13.07	13.67	12.99
Shape and make	12.73	12.73	12.80	13.33	13.00
Green bean volume	151.14	149.65	156.98	151.70	146.73
Total raw quality	35.43	35.31	35.87	37.00	35.99
Roasted beans volume	221.00	231.93	204.46	184.00	203.44
Roasted beans weight	89.91	89.58	90.93	92.29	90.27
Volume changed due to roast	47.20	55.88	30.33	19.70	38.81
Weight loss due to roast	10.09	10.42	9.07	7.71	9.73
Bulk density of green bean	0.66	0.67	0.64	0.66	0.68
Bulk density of roasted bean	0.40	0.39	0.45	0.51	0.44
Aroma intensity	4.01	3.99	4.00	3.67	3.92
Aroma quality	4.10	4.10	4.03	4.00	3.94
Astringency	3.85	3.83	3.73	3.83	3.85
Bitterness	3.89	3.81	3.87	3.83	3.80
Acidity	7.37	7.29	7.57	7.00	7.26
Body	7.45	7.27	7.50	7.00	7.29
Flavor	7.29	7.22	7.47	6.83	7.08
Over all standard	7.35	7.25	7.43	7.00	7.19
Total cup quality	45.29	44.75	45.60	43.17	44.32
Total coffee quality	80.72	80.06	81.47	80.17	80.31

Euclidean genetic distance within and between clusters based on quality traits

The Euclidean genetic distance within and between clusters of the 64, coffee collections based on bean quality traits presented in Table 5. The Euclidean genetic distance within cluster based on coffee bean quality traits ranged from 0.0 to 7.81. The highest Euclidean genetic distance within cluster exploited within cluster V followed by cluster III and I while the lowest Euclidean genetic distance observed within cluster II and IV. The Euclidean genetic distance between clusters ranged from 13.52 to 60.28 for bean quality traits. Accordingly, the highest Euclidean distance between clusters was recorded between cluster II and IV (60.28) followed by between clusters IV and I (46.34) and between cluster II and III (38.32). The minimum Euclidean distance recorded between clusters V and III (13.52) followed by between cluster I and II (14.1).The lowest inter cluster distance observed indicated that the collections in these clusters were not genetically diverse significantly or there was little genetic diversity between them. However, a high inter cluster distance indicated that there is a high chance for obtaining transgressed segregates and maximizing heterosis by crossing genotypes belonging to different clusters as there is a higher chance that distinct genotypes would contribute by unique desired traits. Hence, the maximum recombination and segregation of the progenies expected from crosses involving parents selected from cluster II and IV followed by cluster I and IV. However, the selection of parents should consider special advantages of each cluster and each genotype within a cluster depending on the specific objective of hybridization program.In the inter-cluster distance analysis, clusters that are more divergent are helpful sources of genotypes that could use in the hybridization plan to obtain an ample choice of difference between the isolation and to exploiting heterosis from genetically diverse parental lines. Crosses involving genotypes belonging to most divergent cluster distances used for hybridization program to obtain good manifestations of heterosis and wide variability (Singh and Chaudhury, 1987). In the same way, Olika *et al.,* (2011) conducted inter-cluster distance based on coffee organoleptic quality traits and reported that distance between clusters ranged from 4.66 to 38.82 and similar observations were reported by Abeyot *et al.,* (2011).

In conclusion the investigation depicts the presence of significant differences among the Hararghe coffee collections for bean quality traits. The results of PC analysis suggested the importance of using all bean quality traits to assess the variability. Acidity, flavor, overall standard quality, total cup quality, and total coffee quality contributed more toward the first principal component and most differentiated the clusters. Through the Euclidean distance studies it is being noted that the high inter cluster distance observed which indicates that there is a high chance for obtaining transgressed segregates and maximizing heterosis by crossing genotypes belonging to different clusters as there is a higher chance that distinct genotypes would contribute by unique desired traits.

Table 5: Intra (bold diagonal) and inter (off diagonal) Euclidean distance of five clusters of Hararghe coffee collections based bean quality traits

	Cluster-I	**Cluster-II**	**Cluster-III**	**Cluster-IV**	**Cluster-V**
Cluster-I	**6.96**				
Cluster-II	14.09	**5.63**			
Cluster-III	24.43	38.32	**7.19**		
Cluster-IV	46.34	60.28	23.96	**0.00**	
Cluster-V	20.10	33.40	13.52	27.93	**7.81**

References

1. Abeyot Tessema, Sentayehu Alamerew, Taye Kufa, and Weyessa Garedew 2011. Genetic diversity analysis for quality attributes of some promising (*Coffea arabica* L) germplasm collections in southwestern Ethiopia. J. Biologi. Sci., 11: 236-244.
2. Abrar Sualeh and Negussie Mekonnen 2015. The manual for coffee quality laboratory. ISBN: 978999446605410-8.
3. Abrar Sualeh, Mikru Tesfa, Negussie Mekonene, Bealu Girma, Wubishet Tamirat, Desalegn Alemayehu, Adugna Beyeta, Tamene Haile. 2022. Coffee quality profile of Jimma zone, South Western Ethiopia. Int. J. Food Sci., Biotech., 7(2):48-53.
4. Atalaye Nigussie Temesgen 2017. Time series analysis of the volatility of export price of coffee in Ethiopia. MSc Thesis submitted to graduate studies of Hawasa University, Ethiopia.
5. Brejda JJ, Karlen DL, Smith JL, Allan DL. 2000. Identification of regional soil quality factors and indicators: II. Northern Mississippi loess hills and Palouse prairie. Soil Sci. Society America J., 64: 2125–2135.
6. Central Statistical Agency (CSA). 2020. The Federal democratic republic of Ethiopia central statistical agency. Agricultural sample survey report on area and production of major crops, private peasant holdings, meher season.
7. Chahal, G.S. and Gosal, S.S. 2002. Principles and procedures of plant breeding: Biotechnology and conventional approaches. Alpha Science International, UK, 604.
8. Dadi Gudeta 2016. Genetic variability and association of coffee quality traits in Hararghe coffee (*Coffea arabica* L.) collections. M.Sc. Thesis presented to School of Graduate Studies Haramaya University, Haramaya, Ethiopia.
9. Davis, A.P., Tosh, J., Ruch, N. and Fay, M.F. 2011. Growing coffee: Psilanthus (Rubiaceae) subsumed on the basis of molecular and morphological data; implications for the size, morphology, distribution and evolutionary history of Coffea. Bot. J. Linnean Soci., 167(4):357-377.
10. Ethiopian Commodity Exchanges (ECX). 2015. Coffee contracts. Ethiopia commodity exchange, Addis Ababa, Ethiopia.
11. Farah, A, and Dos Santos, T.F. 2014. The coffee plant and beans: An Introduction. In: R. Preedy V, editor. Coffee in health and disease prevention. Elsevier; pp. 5–10.
12. Getachew WeldeMichael, Sentayehu Alamerew, and Taye Kufa 2015. Genetic diversity analysis of some Ethiopian specialty coffee (*Coffea arabica* L.) accessions for cup quality attributing traits. J. Bio, Agri., Healthcare. ISSN 2224-3208.
13. International Coffee Organization (ICO). 2020. Total production by all exporting countries. International Coffee Organization. Gray's Inn Road, London.
14. International Coffee Organization (ICO). 2021. International coffee organization coffee market report August.
15. Ministry of Agriculture, Ethiopian. 2019. Plant variety release, protection, and seed quality control directorate, Addis Ababa, Ethiopia. Crop Variety Regis, Issue No. 22.
16. Olika Kitila, Sentayehu Alamerew, Taye Kufa, and Weyessa Garedew 2011. Variability of quantitative traits in Limmu coffee (*Coffea arabica* L.) in Ethiopia. Int. J. Agriil. Res., 6: 482-493.

17. Sneath, P.H.A. and Sokal, R.R. 1973. Numerical taxonomy: The Principles and Practice of Numerical Classification. WF Freeman and Co., San Francisco, 573p.
18. Sorse Bayisa 2017. Genetic diversity and associations among quality attributes in Hararghe coffee (*Coffea arabica* L.) landraces. Msc Thesis presented to School of Graduate Studies Haramaya University, Haramaya, Ethiopia.
19. Taye Kufa 2017. What Make Ethiopian Coffees Special: A View from Coffee Research 15th African fine coffee conference and exhibition reshaping the african coffee Industry for Productivity and Investment. February 15-17/2017, Millennium Hall, Addis Ababa, Ethiopia.
20. Tesfaye Kassahun 2006. Genetic diversity of wild *Coffea arabica* L populations in Ethiopia as a contribution to conservation and use planning. Ecol. Develop, Series 44.
21. Yigzaw Dessaleng 2005. Assessment of genetic diversity of Ethiopian Arabica coffee genotypes using morphological, biochemical, and molecular markers. Ph.D. Dissertation, University of the Free State, South Africa.