

Research Article

Phenotypic Diversity of Arabica Coffee (*Coffea Arabica* L.) Genotypes for Qualitative Characteristics at Awada, Ethiopia

Habtamu Gebreselassie^{1, 2, *} , Bizuayehu Tesfaye² , Andargachewu Gedebo²

¹Crop Research Directorate, Wondo Genet Center, Wondogenet, Ethiopia

²School of Plant and Horticultural Science, Hawassa University, Hawassa, Ethiopia

Abstract

Understanding the amount and distribution of genetic diversity is crucial in breeding programs. This study aimed to assess the variation in qualitative morphological traits among 17 Arabica coffee genotypes. The traits displayed a wide range of phenotypic variation, with the Shannon-Weaver diversity index ranging from 0.22 to 1.12 with a mean of 0.67. The chi-square test revealed significant differences for most of the traits, suggesting dominant phenotypic variation among the evaluated traits. Path coefficient analysis showed a positive direct effect of angle of insertion of primary branches (0.485), canopy diameter (0.264), overall appearance (0.101), and leaf apex shape (0.014) on branching habit. According to the PCA, leaf shape (-0.41) and fruit color (-0.36) from the first PCA and leaf apex shape (-0.43) from the second PCA were the important variables contributing more to the variations. The genotypes were classified into five clusters and the pairwise generalized squared distance among the clusters showed significant divergence between most of the clusters. In conclusion, the present study confirmed the existence of qualitative morphological trait variation among evaluated South Ethiopian Arabica coffee genotypes. It is recommended that the studied genotypes be properly conserved and utilized for the coffee genetic improvement program through selection and hybridization.

Keywords

Cluster Analysis, Genetic diversity, Principal Component Analysis, Qualitative Traits, Shannon Diversity Index

1. Introduction

Coffee is one of the most widely traded commodities in the world and the second most consumed non-alcoholic beverage after water [30]. Its popularity and volume of consumption are growing every year, and coffee shops are the fastest-growing part of business. Among the 124 species of the *Coffea* genus in the Rubiaceae family, only Arabica coffee (*Coffea arabica* L.) and

Robusta coffee (*Coffea canephora* P.) are the two most economically important species [7, 12], with Arabica accounting for about 60% of global coffee production and Robusta accounting for approximately 40% [20]. Coffee is produced by about 80 tropical countries and exported by 50 countries, with an annual production of about 9 million tons [31].

*Corresponding author: habteline@gmail.com (Habtamu Gebreselassie)

Received: 15 September 2025; **Accepted:** 29 September 2025; **Published:** 26 January 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Ethiopia is the birthplace of Arabica coffee (*Coffea Arabica* L.), which has the largest genetic diversity and abundance of genetic resources in the world [40]. Arabica coffee is a crucial component of the Ethiopian economy and is deeply intertwined with the social, economic, political, cultural, and spiritual aspects of the country. Ethiopia's unique climate and environment provide ideal conditions for producing high-quality coffee [10]. As of the 2022/23 fiscal year, the total area covered by coffee in Ethiopia is estimated to be 590,000 hectares, with a production and productivity of 496,000 tons and 0.84 tons per hectare, respectively [39].

Genetic diversity refers to the amount of genetic variability present among individuals within a population or variety of a species [24]. It can be observed at three different levels: diversity between species, diversity between populations within one species, and diversity between individuals within one population [29]. This diversity is the result of genetic material (DNA) recombination during the inheritance process, mutations, gene flow, and genetic drift, which lead to variations in DNA sequence, protein structure or isoenzymes, epigenetic profiles, physiological properties, and morphological properties [29].

Genetic diversity is essential for maintaining a healthy population since it enables individuals to adapt to various biotic and abiotic stresses and promotes resistance to pests, diseases, and other stress conditions [29]. Under environmental changes, different crop varieties survive due to the presence of genetic variation, which enables the varieties to adapt [29].

The wealth of diversity in a species offers ample opportunities for genetic improvement of the crop to develop varieties suitable for different agro-ecologies, cropping systems, and purposes. Genetic diversity aids breeders in maintaining crossbred varieties, resulting in the conservation of desired characteristics in the variety [36]. The more diverse the breeding materials, the higher the probability that they will contain desirable genes and the higher the probability of improving the crop for traits of interest [36]. Therefore, diverse lines are necessary for commercial variety defect correction, novel variety development, identification of diverse lines, creation of diversity, and subsequent utilization.

Information about the amount and distribution of genetic diversity among the germplasm is essential for its efficient management and effective utilization in the breeding program. Plant genetic resource genetic diversity analysis can be conducted using phenotypic or morphological traits and cytological, biochemical, and molecular assessment techniques [44]. Morphological markers were initially used for the analysis of diversity, and they are still in use. Morphological evaluation and characterization of plant traits using qualitative traits are the most direct methods of identifying genotype identity and studying genetic diversity among genotypes [35]. This technique is easy, simple, cost-effective, time-efficient, and does not require complicated

equipment or a high level of specialist knowledge for scoring [37].

Several studies have been conducted to estimate the genetic diversity of Arabica coffee using qualitative traits. Different scholars from different countries reported the presence of high variability among Arabica coffee genotypes using qualitative morphological traits [4, 5, 18, 25, 32, 41-43]. Morphological traits are still successfully used to analyze the genetic diversity required for germplasm conservation, reduce accession numbers by identifying and eliminating duplicates, and enhance crop breeding through selection.

Indigenous genotypes play a key role in sustainable production and have an inherent value as the storehouse of wild coffee genetic resources [13]. They also have high functional diversity in terms of disease resistance [11] and pest and drought tolerance [8, 9, 19]. As part of a future-proofing resource, and especially for providing genetic potential for mitigating climate change, indigenous populations are perceived as a key resource for the medium- to long-term sustainability of Arabica production [19]. However, indigenous coffee genetic resources are being lost at a rapid pace due to varied threats, such as the conversion of natural forest to agricultural land in the center of origin, deforestation, land degradation, climate change, leading to an increased incidence of pests and diseases, a higher incidence of drought, and unpredictable rainfall patterns [31].

Hence, indigenous coffee genetic resources are being lost at a rapid pace, an urgent strategy has to be sated in coffee research program to collect, conserve, and utilize indigenous coffee genetic resources. Accordingly, coffee accessions were collected in the major coffee-growing areas of southern Ethiopia (Sidama, Gedeo, Amaro, Gamo Gofa, Jinka, and Guji), and over 700 coffee accessions have been collected and maintained at the Awada field bank. Among the southern collection, 14 promising accessions were selected for variety development after wider adaptability studies at different testing locations. Although these accessions were selected for their high yield potential, coffee berry disease resistance, and coffee leaf rust resistance, they have not yet been systematically characterized for qualitative morphological traits. Therefore, this study was conducted to assess the variation in qualitative morphological diversity among these Arabica coffee genotypes for effective utilization and strategic conservation.

2. Materials and Methods

2.1. Description of Study Area

The experiment took place during the 2021 cropping season at Awada agricultural research sub-center, located in Sidama Regional State, Ethiopia. The center is situated at a latitude of 6°45'46"N and a longitude of 38°22'36"E, with an elevation of 1740 meter above sea level (a.s.l.) The agro ecology of the experimental site is characterized by a semi-bimodal rainfall pattern with double wet and dry seasons, as described by Ajema

& Nigussie [2]. The average wind speed, precipitation, root zone moisture, maximum temperature, and minimum temperature of the experimental site over the past forty-one years (01/01/1981-12/31/2021) were 1.46 m/s, 1399.20mm, 0.68, 34.110C, and 7.610C, respectively [26]. The experimental site has eutric-nitosol and chromotic-cambisol soil type, which is ideal for coffee production [23].

2.2 Planting Materials

For this study, a total of 17 Arabica coffee genotypes were used, consisting of 14 accessions from the south Ethiopian coffee gene pool and three standard reference varieties. The genotypes were originally collected from various agroecologies in south Ethiopia and have been maintained at Awada Sub Center [18]. The genotypes and their sources are listed in Table 1.

Table 1. Description of experimental materials.

Serial No.	Genotypes	Description	Source
1	AW105	Promising Selection	ARSC
2	AW1777	Promising Selection	ARSC
3	AW1995	Promising Selection	ARSC
4	AW3106	Promising Selection	ARSC
5	AW4994	Promising Selection	ARSC
6	AW5994	Promising Selection	ARSC
7	AW7494	Promising Selection	ARSC
8	AW7705	Promising Selection	ARSC
9	AW9622	Promising Selection	ARSC
10	AW9623	Promising Selection	ARSC
11	AW9628	Promising Selection	ARSC
12	AW9641	Promising Selection	ARSC
13	AW9644	Promising Selection	ARSC
14	AW9662	Promising Selection	ARSC
15	Feyate	Released Variety	ARSC
16	Angafa	Released Variety	ARSC
17	74112	Released Variety	JARC

ARSC= Awada Agricultural Research Sub-center; JARC=Jima Agricultural Research Center

2.3. Experimental Design

The experimental materials were planted in August 2015, utilizing a randomized complete block design with three blocks. Each experimental plot comprised eight coffee trees, planted in a single row for each genotype, with a spacing of 2 meters between plants and rows. The distance between replications was 4 meters.

2.4. Experimental Procedures

Seed preparation, seedling media preparation, seed sowing, seedling management, land preparation, planting, and tree

management (weeding, nutrient management, permanent and temporary shedding, soil conservation, and sucker management) were carried out according to Endale *et al.*, (2008) [17].

2.5. Data Collection

The data on qualitative traits were evaluated using the International Plant Genetic Resources Institute (IPGRI) Board descriptor for coffee [21]. The genotypes were assessed for 15 qualitative traits, which are listed in Table 2 below. The qualitative traits were assessed and coded by an experienced team of breeders' subjective judgment.

Table 2. Trait descriptions and codes used for the study.

No.	Traits	Trait Description	Code	No.	Traits	Trait Description	Code
1	PH	Very short	1	9	YSC	Green	1
		Short	3			Dark brown	2
		Tall	7			Other	3
		very tall	9			Elongated conical	1
2	BH	Very few branches (primary)	1	10	OAA	Pyramidal	2
		Many branches (primary) with few secondary	2			Bushy	3
		Many branches (primary)	3			Compact	1
		Many branches (primary) with many secondary and tertiary	4			Intermediate	2
3	AIPB	Drooping	1	11	CD	Open	3
		Horizontal	2			Yellow	1
		Semi erect	3			Yellow orange	2
		Round	1			Orange	3
4	SS	Ovate	2	12	FC	Orange-red	4
		Triangular	3			Red	5
		Deltate (equilaterally triangular)	4			Red – purple	6
		Trapeziform,	5			Purple	7
5	YLC	Other	6	13	FSH	Purple-violet	8
		Greenish	1			Violet	9
		Green	2			Black	10
		Brownish	3			Others	11
6	LSH	Reddish	4	14	PTH	Roundish	1
		Bronze	5			Obovate	2
		Others if any	6			Ovate	3
		Obovate	1			Elliptic	4
7	LASH	Ovate	2	15	SSH	Oblong	5
		Elliptic	3			Others	6
		Lanceolate	4			Thin	3
		Others if any	5			Intermediate	5
8	LPC	Round	1	15	SSH	Thick	7
		Obtuse	2			Round	1
		Acute	3			Obovate	2
		Acuminate	4			Ovate	3
8	LPC	Apiculate, spatulate	5	15	SSH	Elliptic	4
		Others	6			Oblong	5
		Green	1			Others if any	6
		Dark brown	2				

No.	Traits	Trait Description	Code	No.	Traits	Trait Description	Code
	Other		3				

Where: PH=Plant height; BH=Branching habit; AIPB=Angle of insertion of primary branch; SS=Stipule shape; YLC=Young leaf colour; LSH=Leaf shape; LASH=Leaf apex shape; LPC=Leaf petiole colour; YSC=Young shoot colour; OAA=Overall appearance; CD=Canopy Diameter; FC=Fruit colour; FSH=Fruit shape; PT=Pulp thickness; SSH=Seed shape.

2.6. Data Analysis

Descriptive statistics such as frequency distribution and percentage of various categories of the 15 phenotypic qualitative traits were conducted for all genotypes using Microsoft Excel version 2010.

A chi-square test was conducted to test the homogeneity of the populations for the 15 qualitative traits by using Minitab version 21.4 statistical software.

The Shannon diversity index was computed for each of the 15 qualitative morphological characters using the formula described by Shannon and Weaver [28, 33] as follows:

$$H' = - \sum_{i=1}^n (p_i \times \ln p_i) \quad (1)$$

Where H' = Shannon diversity index; n = total number of phenotypic classes for a character; p_i = the proportion of genotypes that have the character in the i^{th} class for a character; $\ln$ = natural logarithm.

Path coefficient analysis was conducted to study independent traits direct and indirect effects. For path coefficient analysis, coffee branching habit was taken as a dependent variable, while the rest of the variables were considered contributory factors. The direct and indirect effects of the independent characters on branching habit were studied using the formula suggested by Dewey and Lu [16]. Principal component analysis was done using SAS software to determine the relative importance of the traits responsible for variation among the coffee genotypes. Pearson's correlation coefficients were calculated among sensory attributes and chemical compositions traits to estimate their relationships by using SAS. The cluster analysis was executed, and the dendrogram was constructed based on the complete linkage method with Euclidean distance by using Minitab version 21.4 software. All 15 qualitative trait data were standardized before clustering.

Inter-cluster distances were calculated based on the Mahalanobis's D^2 statistics as:

$$D^2_{ij} = (x_i - x_j)' \text{cov}^{-1} (x_i - x_j) \quad (2)$$

Where, D^2_{ij} = the distance between genotype i and j ; x_i and x_j = vectors of the values of the variables for cases i and j ; and cov^{-1} = the pooled within groups variance-covariance matrix.

The significance of D^2 values was tested by comparing the D^2 values between any two clusters against tabulated chi-square (χ^2) values at $p-1$ degrees of freedom, where p refers to the number of qualitative characters considered. Inter-cluster distances based on the Mahalanobis's D^2 statistics were computed using Minitab.

3. Results and Discussions

3.1. Frequency Distribution of Qualitative Traits

In this study, 15 qualitative traits were used to characterize the genotypes of south Ethiopian Arabica coffee genotypes (Table 3). According to the percentage of the frequency distribution, the predominant traits among the majority of the studied genotypes were green leaf petiole color (94.12%), elliptic leaf shape (88.24%), pyramidal overall appearance (88.24%), tall plant height (82.35%), red fruit color (88.24%), bronz young leaf color (70.59%), triangular stipule shape (64.71%), intermediate pulp thickness (64.71%), obovate fruit shape (58.82%), obovate seed shape (58.82%), many primary branches (52.94%), and horizontal angle of insertion of primary branch (52.94%). The variability observed in these phenotypic characters indicates the presence of genetically diverse genotypes, which can be used as markers for future selection and hybridization programs. Previous research has also reported a wide variability in percent frequency distribution for qualitative traits among Arabica coffee genotypes [5, 8, 14, 27, 41, 43].

The number of branches is an important qualitative trait that influences crop yield. An increase in the number of branches of the coffee plant helps to increase productivity of the crop. In this study, the frequency of the studied genotypes for branching habits showed that 52.94% of genotypes had many primary branches, whereas 47.06% had many primary, secondary, and tertiary branches. Hence, the 47.06% of genotypes having many primary, secondary, and tertiary branches could be used for future yield improvement programs via selection and/or hybridization due to their branching habit.

The predominance of the open growth habit is a manifestation of the suitability of such traits for ease of management practice [43]. It permits uniform exposure and better interception of sunlight for all leaves and other

vegetative parts, creating a less favorable environment for disease development as compared to compact types [44]. In the current investigation, the canopy class of the genotypes was open (58.92%), intermediate (29.41%), and compact (11.76%). In addition to the canopy class, the angle of insertion of primary branches recorded was horizontal (52.94%) and drooping (47.06%). Accordingly, most of the studied genotypes showed an open growth habit and were suitable for ease of management practice, permitting uniform exposure and better interception of sunlight.

3.2. Chi Square

The chi-square test values for the 15 qualitative traits were evaluated among different genotypes and are presented in Table 3. The results showed that there is a significant difference among all qualitative traits, except for plant height and angle of insertion of primary branch, which exhibit dominant phenotypic variation among the evaluated traits. These significant differences among genotypes based on the chi-square test suggest that there is high variation within the south Ethiopian Arabica coffee genotypes for the studied qualitative traits. Previous studies have also reported significant chi-square test values for qualitative traits among Ethiopian Arabica coffee genotypes [1].

3.3. Shannon-Weaver Diversity Index

The estimated Shannon-Weaver diversity index (H') is

presented in Table 3. The studied traits contributed to the phenotypic diversity at various levels and the diversity index (H') ranged from 0.22 for leaf petiole color (lowest polymorphism) to 1.12 for seed shape (highest polymorphism), with an overall mean of 0.67 (Table 3). The highest diversity was found for seed shape (1.12), canopy diameter (0.924), leaf apex shape (0.88), fruit shape (0.85), stipule shape and pulp thickness (0.81), young leaf color (0.75), branching habit and angle of insertion of the primary branch (0.69), young shoot color (0.61), and plant height (0.58). The higher shannon-weaver diversity index implies the presence of adequate variability for these traits among the evaluated genotypes. Accordingly, the genotypes were highly diverse for seed shape, canopy diameter, leaf apex shape, fruit shape, stipule shape, pulp thickness, and young leaf color. On the other side, the lowest diversity was found for leaf petiole color (0.22), overall appearance (0.36), leaf shape (0.36), and fruit color (0.44), signifying the possibility of a close association between coffee genotypes for those traits. Many studies reported wider range of Shannon-Weaver diversity values among Arabica coffee qualitative traits. 0.67 for branching habit to 0.98 for leaf shape among 124 accessions [6]; 1.22 for stipule shape to 0.24 for calyx limb persistence among 137 accessions [5] (0.92 for growth habit to 0.13 for calyx limb persistence among 104 accessions [14]; 1.22 for fruit color to 0.35 stem habit [43]; 1.08 for angle of insertion of primary branches to 0.17 for screen size among 49 accessions [41].

Table 3. Frequency distribution, proportion, Shannon-waver diversity index, and chi square of 15 Qualitative traits of 17 south Ethiopian coffee genotypes.

Trait	Code	Frequency	% contribution	Contribution to Chi-Square	Chi-Sq
Plant height	3 Short	1	5.88	3.84	18.47** 0.578
	7 Tall	14	82.35	12.25	
	9 Very tall	2	11.76	2.37	
Branching habit	3 Many branches (primary)	9	52.94	0.03	0.06 ^{ns} 0.691
	4 Many branches (primary) with many secondary and tertiary	8	47.06	0.03	
Angle of insertion of primary brunch	1 Drooping	8	47.06	0.03	0.07 ^{ns} 0.691
	2 Horizontal	9	52.94	0.03	
Stipule shape	3 Triangular	11	64.71	5.02	8.94* 0.808
	4 Deltate (equilaterally triangular)	5	29.41	0.08	
	5 Trapeziform,	1	5.88	3.84	
Young leaf color	1 Greenish	1	5.88	3.84	11.41** 0.753
	2 Green	4	23.53	0.49	
	5 Bronze	12	70.59	7.08	

Trait	Code		Frequency	% contribution	Contribution to Chi-Square	Chi-Sq	
Leaf shape	3	Elliptic	15	88.24	4.97	9.94*	0.362
	4	Lanceolate	2	11.76	4.97		
Leaf apex shape	1	Round	1	5.88	3.84	5.76*	0.876
	4	Acuminate	8	47.06	0.96		
Leaf petiole color	5	Apiculate, spatulate	8	47.06	0.96	13.24**	0.224
	1	Green	16	94.12	6.62		
Young shoot color	3	Other	1	5.88	6.62	2.88 ^{ns}	0.606
	1	Green	5	29.41	1.44		
Overall appearance	3	Dark brown	12	70.59	1.44	9.94*	0.362
	1	Elongated conical	2	11.76	4.97		
Canopy Diameter	2	Pyramidal	15	88.24	4.97	5.76*	0.924
	1	Compact	2	11.76	2.37		
Fruit colour	2	Intermediate	5	29.41	0.08	23.06**	0.444
	3	Open	10	58.82	3.31		
Fruit shape	4	Orange-red	1	5.88	3.84	7.18*	0.846
	5	Red	15	88.24	15.37		
Pulp thickness	6	Red – purple	1	5.88	3.84	8.94*	0.808
	1	Roundish	6	35.29	0.02		
Seed shape	2	Obovate	10	58.82	3.31	10.53*	1.122
	3	Ovate	1	5.88	3.84		
	4	Thin	1	5.88	3.84		
	5	Intermediate	11	64.71	5.02		
	6	Thick	5	29.41	0.08		
	1	Round	2	11.76	1.19		
	2	Obovate	10	58.82	7.78		
	3	Ovate	3	17.65	0.37		
	4	Elliptic	2	11.76	1.19		

3.4. Correlation

The correlation analyses for 15 qualitative morphological characteristics were assessed, and the results are given in Table 4. Branching habit was significantly and positively correlated with the angle of insertion of the primary branch ($r = 0.417$), overall appearance ($r = 0.344$), and canopy diameter ($r = 0.548$), whereas, it was negatively correlated with leaf shape ($r = -0.344$) and leaf apex shape ($r = -0.203$) (Table 4). Pulp thickness was significantly and positively associated

with fruit shape ($r = 0.412$), seed shape ($r = 0.632$), young leaf color ($r = 0.275$), and young shoot color ($r = 0.278$), whereas it was negatively correlated with fruit color ($r = -0.577$). Seed shape was positively correlated with fruit shape ($r = 0.56$) and negatively correlated with leaf shape ($r = -0.352$) and leaf petiole color ($r = -0.393$). Canopy diameter was significantly and positively correlated with overall appearance ($r = 0.772$), branching habit ($r = 0.548$), plant height ($r = 0.427$), and angle of insertion of the primary branch ($r = 0.299$), whereas it was negatively correlated with leaf shape ($r = -0.509$), fruit color ($r = -0.493$), and stipule shape ($r = -0.323$) (Table 4).

Table 4. Correlation coefficients among 15 qualitative morphological characters.

	FC	FS	PT	SS	PH	BH	AIPB	SH	YLC	LS	LAS	LPC	YSC	OA	CD
FC	1	-0.601**	-0.314*	-0.208 ^{ns}	-0.577**	-0.144 ^{ns}	0.344*	-0.286*	-0.349*	0.532**	0.000 ^{ns}	0.000 ^{ns}	-0.376*	-0.532**	-0.493**
FS		1	0.412**	0.560**	0.174 ^{ns}	0.073 ^{ns}	-0.280*	0.010 ^{ns}	0.370*	-0.452**	0.267*	-0.309*	0.346*	0.13 ^{ns}	0.052 ^{ns}
PT			1	0.632**	0.182 ^{ns}	0.02 ^{ns}	-0.241 ^{ns}	0.063 ^{ns}	0.275*	-0.158 ^{ns}	-0.133 ^{ns}	-0.108 ^{ns}	0.278*	0.158 ^{ns}	0.173 ^{ns}
SS				1	0.120 ^{ns}	0.093 ^{ns}	0.05 ^{ns}	-0.245 ^{ns}	0.228 ^{ns}	-0.352*	0.264 ^{ns}	-0.393**	0.231 ^{ns}	0.130 ^{ns}	0.066 ^{ns}
PH					1	0.198 ^{ns}	-0.397**	0.000 ^{ns}	0.202 ^{ns}	-0.615**	0.104 ^{ns}	0.000 ^{ns}	0.217 ^{ns}	0.615**	0.427**
BH						1	0.417**	-0.058 ^{ns}	0.122 ^{ns}	-0.344*	0.203*	-0.236 ^{ns}	0.091 ^{ns}	0.344*	0.548**
AIPB							1	-0.335*	-0.362*	-0.022 ^{ns}	0.290*	-0.265*	-0.350*	0.022 ^{ns}	0.299*
SH								1	0.239 ^{ns}	0.358*	-0.622**	0.662**	0.228 ^{ns}	-0.054 ^{ns}	-0.323*
YLC									1	-0.138 ^{ns}	-0.346*	0.160 ^{ns}	0.989**	0.138 ^{ns}	0.030 ^{ns}
LS										1	-0.685**	0.685**	-0.165 ^{ns}	-0.433**	-0.509**
LAS											1	-0.862**	-0.342*	-0.079 ^{ns}	0.146 ^{ns}
LPC												1	0.161 ^{ns}	0.091 ^{ns}	-0.169 ^{ns}
YSC													1	0.165 ^{ns}	0.065 ^{ns}
OA														1	0.772**
CD															1

3.5. Path Analysis

To assess the magnitude of direct and indirect contributions of qualitative characters to Arabica coffee branching habit, path coefficient analysis was performed using five significantly associated traits with branching habit. Among the 5 significantly correlated characters with branching habit, an angle of insertion of primary branches (0.485), followed by

canopy diameter (0.264), overall appearance (0.101), and leaf apex shape (0.014), exerted a positive direct effect, whereas leaf shape (-0.003) had a negative direct influence on branching habit (Table 5). An angle of insertion of primary branches and canopy diameter had the highest positive direct effect on branching habit; they are found to be important components, and direct selection for this trait may be rewarding for the improvement of Arabica coffee branching habit.

Table 5. Path coefficient analysis (bold and diagonal value indicating direct effect; above and below the diagonal value indicating indirect effect) of 5 qualitative morphological traits on branching habit.

	AIPB	LS	LAS	OA	CD
AIPB	0.485	0.000	0.004	-0.002	0.079
LS	-0.011	-0.003	-0.009	0.044	-0.134
LAS	0.141	0.002	0.014	0.008	0.039
OA	0.011	0.001	-0.001	0.101	0.204
CD	0.145	0.001	0.002	-0.078	0.264

3.6. Principal Component Analysis

The principal component analysis (PCA) of 17 Arabica coffee genotypes for 15 qualitative morphological traits were performed to estimate the relative contribution of each attribute to the observed variability, and the results are presented in Table 6. The PCA analysis generated the fifteen eigenvalues and eigenvectors. However, factors to be retained should have more than 1 eigenvalue, at least 5% of the variance explained for each component, and/or more than 70% of the cumulative proportion of variance explained [30]. Accordingly, the first five components that explained 83.26 percent of total variation were used for displaying characters (Table 6).

The PCA analysis might be useful for deducing the nature of attributes and reducing the complexity of data collection [3]. The first PCA explained 29.26% of the total variations (Table 6). The relative weight given to the variables in each component is determined by the importance of the variables, which possess a high positive and negative weight [22]. Based on this suggestion, the most important characters contributing more to the variation were leaf shape (-0.41), fruit color (-0.36), plant height (0.31), canopy diameter (0.31), fruit shape (0.30), and overall appearance (0.3). The sign indicates the direction of the relationship between the components and the characters [38]. Accordingly, in the first principal component,

all traits listed except leaf shape and fruit color were positively associated.

The second PCA explained 23.42% of the variation, and leaf apex shape (-0.43), angle of insertion of primary branches (-0.34), stipule shape (0.39), leaf petiole color (0.39), young leaf color (0.36), and young shoot color (0.36) mainly accredited this variation. The third PCA explained 8.33% of the total variation, and this variation is mainly contributed by canopy diameter (0.42), overall appearance (0.41), seed shape (-0.38), and fruit shape (-0.35) in decreasing order (Table 6).

The fourth and fifth principal components explained 8.95% and 7.37%, respectively. High variation in the fourth PC is attributed to angle of insertion of primary branches (0.52), plant height (-0.42), branching habit (0.4), young shoot color (0.29), seed shape (0.24), while high variation in the fifth PC is attributed to pulp thickness (0.59), young leaf color (-0.41), and seed shape (0.38) (Table 6).

From the PCA result of the present study, it may be concluded that important variables in Arabica coffee genotype with respect to qualitative morphological characters were leaf shape, leaf apex shape, canopy diameter, overall appearance, angle of insertion of primary branches, plant height, branching habit, pulp thickness, and young leaf color. These variables might be taken into consideration for the effective selection of parents.

Table 6. Eigen values and Eigenvectors of the first 5 principal components (PCA) for 15 qualitative morphological characters.

Eigenvectors	PCA1	PCA2	PCA3	PCA4	PCA5
Fruit colour	-0.36	-0.2	-0.12	0.17	-0.03
Fruit shape	0.30	0.09	-0.35	-0.06	0.04
Pulp thickness	0.22	0.15	-0.25	0.16	0.59
Seed shape	0.26	-0.04	-0.38	0.24	0.38
Plant height	0.31	0.10	0.22	-0.42	-0.01
Branching habit	0.22	-0.11	0.29	0.40	-0.15
Angle of insertion of primary branches	-0.05	-0.34	0.18	0.52	0.04
Stipule shape	-0.11	0.39	0.09	-0.06	0.06
Young leaf color	0.19	0.36	-0.15	0.3	-0.41
Leaf shape	-0.41	0.18	0.02	0.2	0.21
Leaf apex shape	0.18	-0.43	-0.18	-0.19	-0.19
Leaf petiole color	-0.21	0.39	0.27	0.00	0.15
Young shoot color	0.20	0.36	-0.13	0.29	-0.4
Overall appearance	0.30	0.06	0.41	-0.02	0.20
Canopy Diameter	0.31	-0.1	0.42	0.16	0.13
EV	4.39	3.51	2.14	1.34	1.11

Eigenvectors					
	PCA1	PCA2	PCA3	PCA4	PCA5
PVE	29.26%	23.42%	14.26%	8.95%	7.37%
CPVE	29.26%	52.68%	66.94%	75.89%	83.26%

Where; EV = Eigen value, PVE = Present variation explained. and CPVE = Cumulative present variation explained

3.7. Cluster Analysis

The classification of the 15 qualitative traits using complete linkage hierarchical clustering with Euclidean distance revealed the presence of five distinct clusters. Figure 1 and Table 7 illustrate the distribution of genotypes, with 7 genotypes (41.18%) in Cluster-V, 6 genotypes (35.29%) in Cluster-III, 2 genotypes (11.76%) in Cluster-IV, and 1 genotype, 74112, in Cluster-I (5.88%) and AW9623 in Cluster-II (5.88%). Similarity was considered within a cluster, while dissimilarity was considered between different clusters. The distribution of genotypes into different clusters indicates the presence of genetic variation among coffee genotypes. Many studies reported the clustering of Arabica coffee accessions into distinct groups based on qualitative traits. Accordingly, 10 distinct clusters among 124 accessions based on 7 qualitative traits were reported by Atinafu and Mohamed [30]; 5 distinct clusters among 137 based on 12 qualitative traits were reported by Asegid [5]; 5 distinct groups among 104 accessions based on nine qualitative traits were reported

by Degefa [14]; 6 clusters among 64 accessions based on 13 qualitative traits were reported by Desalegn [15]; and 5 distinct groups among 62 accessions based on 12 qualitative traits were reported by Yirga [43].

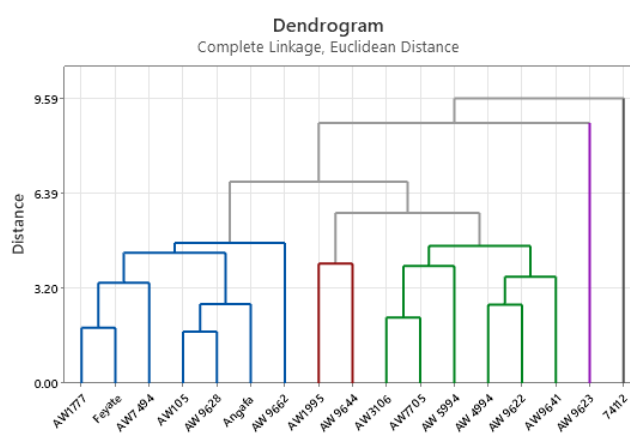


Figure 1. Cluster dendrogram describing variation among 17 genotypes of Arabica coffee for 15 qualitative phenotypic traits.

Table 7. Clustering patterns of 17 coffee genotypes based on 12 qualitative characters.

Cluster No.	Number of Genotype	Percent (%)	List of genotype
I	1	5.88	74112
II	1	5.88	AW9623
III	6	35.29	AW3106, AWA7705, AW5994, AW4994, AW9622, AW9641
IV	2	11.76	AW1995, AW9644
V	7	41.18	AW1777, Feyate, AW7494, AW105, AW9628, Angafa, AW9662

3.8. Distance Between Cluster Centroid Ward Method

The pairwise generalized squared distance (D^2) among the five clusters based on the Mahalanobis's D^2 statistics revealed significant ($P \leq 0.05$) divergence between most of the clusters

(Table 8). The maximum genetic distance was observed between clusters C-IV and C-V ($D^2=9.086^*$), followed by C-I and C-IV ($D^2=8.166^*$), C-II and C-IV ($D^2=7.768^*$), C-II and C-V ($D^2=7.570^*$), C-III and C-V ($D^2=7.501^*$), C-III and C-IV ($D^2=7.290^*$), C-I and C-V ($D^2=7.032^*$), and C-I and C-II ($D^2=4.023^*$) indicating divergence of genotypes in these clusters. Non-significant inter-cluster distances were observed

between clusters C-I and C-II ($D^2 = 3.332^{ns}$) and clusters C-II and C-III ($D^2 = 3.135^{ns}$) indicating that genotypes in these clusters share similar genetic backgrounds (Table 8). The higher inter-cluster distance value showed the existence of wider genetic variability among the tested groups of genotypes [5]. The maximum genetic recombination is expected from the hybridization of the parents selected from divergent cluster groups. Accordingly, the maximum genetic

recombination and higher heterotic F1 offspring are expected from crosses involving parents selected from clusters I and V, followed by I and IV, II and IV, II and V, III and V, III and IV, I and V, and I and II in decreasing order. However, the selection of parents should also consider the special advantages of each cluster and each genotype within a cluster depending on the specific objectives of hybridization [34].

Table 8. Inter cluster distance for 15 qualitative traits of south Ethiopian Arabica coffee genotypes.

	I	II	III	IV	V
I	0	3.332 ^{ns}	4.023*	8.166*	7.032*
II		0	3.135 ^{ns}	7.768*	7.570*
III			0	7.290*	7.501*
IV				0	9.086*
V					0

4. Conclusion

Estimates of frequency distribution, Shannon index, and cluster analysis based on 15 qualitative morphological traits revealed the existence of genetic variation among Arabica coffee genotypes. The frequency distribution of traits showed a wide range of phenotypic variation, with the maximum Shannon index (H') found for seed shape, canopy diameter, leaf apex shape, fruit shape, stipule shape, pulp thickness, and young leaf colour. Results of the chi-square test showed significant differences for most of the traits indicating dominant phenotypic variation among the evaluated traits.

Path coefficient analysis showed a positive direct effect of angle of insertion of primary branches, canopy diameter, overall appearance, and leaf apex shape on branching habit. From the PCA, important variables in Arabica coffee genotype with respect to qualitative morphological characters were leaf shape, leaf apex shape, canopy diameter, overall appearance, angle of insertion of primary branches, plant height, branching habit, pulp thickness, and young leaf color. These variables might be taken into consideration for the effective selection of parents.

The genotypes were classified into five groups based on the results of the cluster analysis. Cluster-V had the maximum number of genotypes (41.18%), followed by cluster-III (35.29%), and cluster-IV (11.76%). 74112 is the only variety found in cluster I and this genotype had a chance to develop hybrid vigor through crossing diverged parents found in different clusters. The pairwise generalized squared distance among the clusters showed significant divergence between the clusters. However, the maximum genetic recombination and

higher heterotic F1 offspring are expected from crosses involving parents selected from clusters 1 genotype 74112 and clusters 5 genotype AW1777, Feyate, AW7494, AW105, AW9628, Angafa, AW9662.

The study confirmed the existence of significant genetic variability among the Arabica coffee genotypes for most of the studied qualitative traits, which provides excellent opportunities for genetic gain through selection, hybridization, and strategic conservation. However, the morphological diversity observed in this study needs to be further confirmed using molecular techniques of characterization.

Abbreviations

AIPB	Angle of Insertion of Primary Branch
BH	Branching Habit
CD	Canopy Diameter
D^2	Pairwise Generalized Squared Distance
DNA	Deoxyribonucleic Acid
FC	Fruit Colour
FSH	Fruit Shape
H'	Shannon Diversity Index
IPGRI	International Plant Genetic Resources Institute
LASH	Leaf Apex Shape
LPC	Leaf Petiole Colour
LSH	Leaf Shape
OAA	Overall Appearance
PCA	Principal Component Analysis
PH	Plant Height
PT	Pulp Thickness
R	Correlation Coefficient

SS	Stipule Shape
SSH	Seed Shape
YLC	Young Leaf Colour
YSC	Young Shoot Colour
χ^2	Chi-square

Acknowledgments

The author wishes to express gratitude to the Awada Agricultural Research Sub Center staff members for their valuable material support, technical assistance, and cooperation in the field throughout the data collection period. This research work was financed by Ethiopian Institute of Agricultural Research and the institutional collaboration program between Hawassa University (Ethiopia) and the Norwegian University of Life Science.

Author Contributions

Habtamu Gebreselassie: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing

Bizuayehu Tesfaye: Methodology, Supervision, Writing – original draft, Writing – review & editing

Andargachewu Gedebo: Data curation, Project administration, Resources, Supervision, Writing – review & editing

Conflicts of Interest

The authors declare that they have no conflict of interest regarding the publication of this work.

References

- [1] Adem, A., (2009): Agro-morphological characterization of coffee (*Coffea arabica* L.) landrace collected from Mesela, West Harerge, Ethiopia. M.Sc. Thesis Submitted to Graduate Studies of Hawassa University, Hawassa, Ethiopia. 88 pp.
- [2] Ajema, L., & Nigussie, A. (2021): Yield and Yield Advantage of the Component Crops as Affected by Strip Intercropping of Coffee (*Coffea arabica* L.) with Pineapple (*Ananas comosus* L.). American Journal of BioScience, 9(4), 116. <https://doi.org/10.11648/j.ajbio.20210904.11>
- [3] Al-Sayed, H. M., Fateh, H. S., Fares, W. M., Attaya, A. S., (2012): Multivariate analysis of sugar yield factors in sugarcane. *Am.-Eurasian J. Sustain. Agric.* (AEJSA) 6 (1), 44–50.
- [4] Ameha, M., (1986): The significance of Ethiopian coffee genetic resources to coffee improvement. In the proceedings of an international symposium on the conservation and utilization of Ethiopian germplasm, pp. 247-253. Addis Ababa, Ethiopia, 13-16 October, 1986.
- [5] Asegid, A. (2019): Morphological Characterization of Bale and West Arsi Coffee (*Coffea arabica* L.) Collections at Gera, Southwest Ethiopia. A thesis Submitted for the Partial Fulfillment of the Degree of Masters of Science in Plant Breeding of Jimma University. PP 1-91.
- [6] Atinafu, G., & Mohammed, H. (2017): Agro-Morphological Characterization of Sidama Coffee (*Coffea Arabica* L.) Germplasm Accession under its Specialty Coffee Growing Area, Awada, Southern Ethiopia. *International Journal of Research Studies in Science, Engineering and Technology*. 4(12) 11-23.
- [7] Bicho, N. C. C., Oliveira, J. F. S., Leitão, A. E., Ramalho J. C. and Lidon, F. C. (2011): Coffee – Origins, Production, Processing and Quality Definition. Livraria Escolar Editora. P. 170, Lisboa, Portugal.
- [8] Burkhardt J., Kufa, T., Beining, A., Goldbach, H. E., Fetene, M. (2007): Drought adaptation strategies of *Coffea arabica* populations along a rainfall gradient in Ethiopia. 21st International Conference on Coffee Science, Montpellier, France, 11–15 September: 1032–1036.
- [9] Carvalho, F. G., Sera, G. H., Andreazi, E., Sera, T., de Barista Fonseca, I. C. B., Carducci, C., et al. (2017): Drought tolerance in seedlings of coffee genotypes carrying genes of different species. *Coffee Sci.* 12, 9–16.
- [10] CSA (Central Statistical Authority), (2015): Federal Republic of Ethiopia, Central Statistical Agency, Agricultural Sample Survey Report on Area and Production of Major Crops, 2014/2015, Vol. 1, Statistical Bulletin 578, May 2015, Addis Ababa, 121P.
- [11] Davis, A. P., Kiwuka, C., Faruk, A., Mulumba, J., Kalema, J. (2023): A review of the indigenous coffee resources of Uganda and their potential for coffee sector sustainability and development. *Frontiers in Plant Science*. 1-14. <https://doi.org/10.3389/fpls.2022.1057317>
- [12] Davis, A. P. (2011): *Psilanthusmannii*, the type species of *Psilanthus*, transferred to *Coffea*. *Nordic Journal of Botany*, 29: 471–472. <https://doi.org/10.1111/j.1756-1051.2011.01113.x>
- [13] Davis, A. P., Chadburn, H., Moat, J., O'Sullivan, R., Hargreaves, S., and Nic Lughadha, E. (2019): High extinction risk for wild coffee species and implications for coffee sector sustainability. *Sci. Adv.* 5, eaav3473. <https://doi.org/10.1126/sciadv.aav3473>
- [14] Degefa, M. (2021): Phenotypic Classification of Coffee (*Coffea arabica* L.) Germplasm in Southern Ethiopia. *American Journal of Agriculture and Forestry*, 9(6), 358. <https://doi.org/10.11648/j.ajaf.20210906.15>
- [15] Desalegn, A. (2018): Genetic Variability and Character Association of Amaro Coffee (*Coffea arabica* L.) Accessions at Awada, Southern Ethiopia. Thesis Submitted to Graduate Studies of Jimma University, Jimma, Ethiopia. P1-38.
- [16] Dewey, D. R., Lu, K. H. (1959): A correlation and path coefficient analysis of components of crested wheat grass seed production. *Agro. J.* 51: 515-518.

- [17] Endale, T., Kufa, T., Nestre, A., Shimber, T., Yilma, A., Ayano, T. (2008): Research on coffee field management. Proceedings of the workshop on four decades of coffee research and development in Ethiopia: A National workshop, Addis Ababa, Ethiopia. pp. 187-195.
- [18] Gebreselassie, H., Tesfaye, B., Gedebo, A. (2024): Genetic diversity of Arabica coffee genotypes in south Ethiopia using quantitative agro-morphological traits. *Genet Resour Crop Evol* 71, 3485–3506. <https://doi.org/10.1007/s10722-023-01853-9>
- [19] Hein, L., Gatzweiler, F. (2006): The economic value of coffee (*Coffea arabica*) genetic resources. *Ecol Econ* 60: 176–185.
- [20] ICO (International Coffee Organization) (2018): Coffee market report. Available at <http://www.ico.org/>
- [21] IPGRI (International Plant Genetic Resources Institute) (1996): Descriptions of *coffea* spp. and *Psilanthus* spp. International Plant Genetics Resource Institute, Rome.
- [22] Jeffers, J. N. (1967): Two case studies in the application of principal component analysis. *Appl. Stat.* 16: 225-36.
- [23] Kebede, M. and Bellachew, B. (2008). Phenotypic Diversity in the Hararge Coffee (*Coffea arabica* L.) Germplasm for Quantitative Traits. *East African Journal of Sciences*, 2, 13-18. <https://doi.org/10.4314/eajsci.v2i1.40358>
- [24] Minter, M., Nielsen, E. S., Blyth, C., Bertola, L. D., Kantar, M. B., Morales, H. E., Orland, C., Segelbacher, G., Leigh, D. M. (2021): What Is Genetic Diversity and Why Does it Matter? *Frontiers for Young Minds*, 9: 656168 <https://doi.org/10.3389/frym.2021.656168>
- [25] Montagnon, C. and Bouharmont, P. (1996): Multivariate analysis of phenotypic diversity of *Coffea arabica*. *Genetic. Resour. Crop. Evol.* 43: 221 – 227.
- [26] NASA (National Aeronautics and Space Administration) (2023): ArcGIS World Geocoding Service Power data access viewer. Retrieved from <https://power.larc.nasa.gov/data-access-viewer/> Accessed 13 June 2023
- [27] Ndikumana, J. (2022): Agro-morphological Characterization of Arabica Coffee Cultivars in Burundi. A Thesis Submitted in Partial Fulfilment of the Requirements for the degree of Master of Science in Plant Breeding of the Jomo Kenyatta University of Agriculture and Technology. P1-135.
- [28] Rain, R.. (2023): "Shannon Diversity Index Calculator". Available at: <https://www.omnicalculator.com/ecology/shannon-index> Accessed: 10 May 2023.
- [29] Salgotra, R. K., & Chauhan, B. S. (2023): Genetic Diversity, Conservation, and Utilization of Plant Genetic Resources. *Genes*, 14(1). <https://doi.org/10.3390/genes14010174>
- [30] Samper, L, Giovannucci, D., Marques-Vieira, L. (2017): The Powerful Role of Intangibles in the Coffee Value Chain. WIPO Economic Research Working Paper No. 39. Geneva: WIPO.
- [31] Sarada K. (2017): Sustainable Coffee Production. Oxford Research Encyclopedia of Environmental Science. 1-3 <https://doi.org/10.1093/acrefore/9780199389414.013.224>
- [32] Selvakumar, M. and Sreenivasan, M. S. (1989): Studies on morphology and quality of Ethiopian arabica coffee. *Journal of Plantation Crops*, 16, pp. 321-324.
- [33] Shannon, C. E., & Weaver, W. (1949): The Mathematical Theory of Communication. Urbana, IL: The University of Illinois Press, 1-117.
- [34] Singh, B. D. (2001): Plant breeding: Principles and Methods. Kalyani Publishers, New Delhi. 896p.
- [35] Tatari, M., & Abdollahi, H. (2021): Evaluation of Vegetative and Reproductive Characteristics of Some Quince (*Cydonia Oblonga* Mill.) Genotypes from Central Regions of Iran. *International Journal of Fruit Science*, 21(1), 945–954. <https://doi.org/10.1080/15538362.2021.1948377>
- [36] Temesgen, B. (2021): Role and economic importance of crop genetic diversity in food security. *International Journal of Agricultural Science and Food Technology*, 7, 164–169. <https://doi.org/10.17352/2455-815x.000104>
- [37] Tesfa, M. (2022): Characterization and Estimation of Diversity of Sugarcane (*Saccharum officinarum*) Genotypes Based on Qualitative Morphological Traits. 1–17.
- [38] Tsehay, Y., Kebebew F. (2002): Morphological diversity and geographic distribution of adaptive traits in finger millet (*Eleusine coracana* (L.) Gaertn. Subsp. *coracana* [Poaceae]) populations from Ethiopia. Addis Ababa, Ethiopia.
- [39] USDA (The United States Department of Agriculture), (2023): Ethiopia Coffee Annual Report, GAIN Report Number: ET2023-0014, 7p. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Coffee%20Annual_Addis%20Ababa_Ethiopia_ET2023-0014.pdf
- [40] Vavilov, N. I. (1951): The origin, variation, immunity and breeding of cultivated plants. (Translated by S. K. Chestitee). *Chronica Botonica*, 13, 1-366.
- [41] Weldemichael, G. (2021): Phenotypic Diversity of Ethiopian Coffee (*Coffea arabica* L.) Accessions Collected from Gomma District in Jimma Zone for Qualitative Traits. *International Journal of Research Studies in Science, Engineering and Technology* 7(8): 08-14. <https://www.researchgate.net/publication/349916670>
- [42] Yigzaw D. (2005): Assessment of cup quality, morphological, biochemical and molecular diversity of *Coffea arabica* L. genotypes of Ethiopia. PhD. Thesis, University Free State. 197 pp.
- [43] Yirga, M. (2021): Phenotypic Characterization of Coffee (*Coffea Arabica* L.) Germplasm, in Ethiopia. *American Journal of BioScience*, 9(2), 34. <https://doi.org/10.11648/j.ajbio.20210902.11>
- [44] Zhu, Y., Liang, D., Song, Z., Tan, Y., Guo, X., & Wang, D. (2022): Genetic Diversity Analysis and Core Germplasm Collection Construction of *Camellia oleifera* Based on Fruit Phenotype and SSR Data. *Genes*, 13(12). <https://doi.org/10.3390/genes13122351>