VARIABILITY ANALYSIS OF WILD GUATEMALAN AVOCADO GERMPLASM BASED ON AGRO-MORPHOLOGICAL TRAITS †

[ANÁLISIS DE VARIABILIDAD DEL GERMOPLASMA DE AGUACATE SILVESTRE GUATEMALTECO BASADO EN CARACTERÍSTICAS AGROMORFOLÓGICAS]

José Alejandro Ruiz-Chután[1,2], Julio Ernesto Berdúo-Sandoval [2], Anna Maňourová[1,3], Marie Kalousová[1], Carlos Enrique Villanueva-González[1,4], Eloy Fernández[1], Jana Žiarovská[5], Amílcar Sánchez-Pérez[2], and Bohdan Lojka[1*]

*[1]Department of Crop Sciences and Agroforestry, Czech University of Life Sciences Prague, Kamýcká 129 165 00 Praha – Suchdol, Czech Republic. Emails: ruiz_chutan@ftz.czu.cz, manourova@ftz.czu.cz, kalousovam@ftz.czu.cz, villanueva@ftz.czu.cz, eloy@ftz.czu.cz, *lojka@ftz.czu.cz*

[2] Facultad de Agronomía, Universidad de San Carlos de Guatemala, Ciudad Universitaria Zona 12, Guatemala. Emails: josealejandro.ruiz@icloud.com, berduoj@hotmail.com, gramisp@hotmail.com

[3]Department of Plant Protection Biology, Swedish University of Agricultural Sciences, SE-230 53 Alnarp, Sweden. Email: anna.manourova@slu.se

[4]Facultad de Ciencias Ambientales y Agrícolas, Universidad Rafael Landívar, Campus San Pedro Claver, S.J. de la Verapaz, km 218.5, carretera a San Juan Chamelco, Alta Verapaz, Guatemala. Email: krevilla01@gmail.com.

[5]Department of Genetics and Plant Breeding, Research Centre AgroBioTech, Slovak University of Agriculture in Nitra, Slovak Republic. Email: jana.ziarovska@uniag.sk.

**Corresponding author*

SUMMARY

Background. For the Guatemalan culture, the avocado (*Persea americana* Mill) represents more than a crop, since it has religious, mythical, economical, and medicinal significance typical of the center of domestication of the species. Although there is speculation about the wide variability of wild avocado germplasm, there are no studies describing avocado populations. **Objective.** To explore wild germplasm based on agro-morphological characteristics using the standardized descriptors for avocado (*Persea* spp.). **Methodology.** The distribution of characteristics of 189 avocado trees across eight populations was evaluated using cross-tabulation and Chi-square testing. In order to properly handle qualitative and quantitative data, it was employed the factor analysis of mixed data (FAMD) and hierarchical clustering on principal components (HCPC) to analyze the relationship and variation of features within and between populations. **Results.** The samples showed various morphological characteristics that indicate the presence of Mexican, Guatemalan, and West Indian avocado races in Guatemala. The FAMD revealed that the fruit shape, skin color and flesh texture, and anise odor in the leaves are ideal traits for distinguishing between individual trees. At the population level, the HCPC showed considerable variance amongst the studied trees but no distinct geographic groups of the samples. **Implications.** Although not all wild avocados grown in Guatemala are of the highest quality, several trees of extraordinary excellence are strewn over the republic's highlands. Thus, due to the danger posed by the introduction of commercial varieties, it is advisable preserving the germplasm. **Conclusion.** The clustering methods revealed that the 189 avocado trees could be regrouped into three main clusters. The morphological

---

[†] *Submitted December 5, 2022 – Accepted March 17, 2023. http://doi.org/10.56369/tsaes.4663*

ISSN: 1870-0462.

ORCID = J.A. Ruiz-Chután: http://orcid.org/0000-0003-3897-855X; **J.E. Berdúo-Sandoval**: http://orcid.org/0000-0001-6043-9152; **A. Maňourová**: http://orcid.org/0000-0002-5668-9850; **M. Kalousová**: http://orcid.org/0000-0001-7078-4064; **C.A. Villanueva-González:** http://orcid.org/0000-0003-4620-1441; **E. Fernández**: http://orcid.org/0000-0002-5059-1975; **J. Žiarovská**: http://orcid.org/0000-0002-0005-9729; **A. Sánchez-Pérez**: http://orcid.org/0000-0002-9938-6315; **B. Lojka**: http://orcid.org/0000-0002-5059-1975

descriptors proved to be useful for grouping trees according to known races and can therefore be used in the characterization of cultivars and wild trees whose ancestry is unknown.

Key words: *Persea americana*; diversity; phenotypic variation; factorial analysis of mixed data; wild germplasm; conservation.

RESUMEN

Antecedentes. Para la cultura guatemalteca, el aguacate (*Persea americana* Mill.) representa más que un cultivo ya que tiene un significado religioso, mítico, económico y medicinal propio de un centro de domesticación de la especie. Aunque se especula sobre la amplia variabilidad del germoplasma nativo de aguacate, no existen estudios que describan las poblaciones de aguacate. **Objetivo.** Explorar el germoplasma nativo a partir de características agromorfológicas utilizando los descriptores estandarizados para aguacate (Persea spp.). **Metodología**. Se evaluó la distribución de las características de 189 árboles de aguacate en ocho poblaciones mediante tabulación cruzada y pruebas de chi-cuadrado. Para manejar adecuadamente los datos cualitativos y cuantitativos, se empleó el análisis factorial de datos mixtos (FAMD) y el agrupamiento jerárquico de componentes principales (HCPC) para analizar la relación y la variación de características dentro y entre poblaciones. **Resultados**. Las muestras presentaron diversas características morfológicas que indican la presencia de las razas de aguacate mexicano, guatemalteco y antillano en Guatemala. El FAMD reveló que la forma, color de la superficie y textura de la pulpa del fruto y el olor a anis en las hojas son rasgos ideales para distinguir entre árboles individuales. A nivel de población, el HCPC mostró una variación considerable entre los árboles estudiados pero no grupos geográficos distintos de las muestras. **Implicaciones**. Aunque no todos los aguacates silvestres que crecen en Guatemala son de la mejor calidad, varios árboles de extraordinaria excelencia están diseminados por las tierras altas de la república. Por ello, debido al peligro que supone la introducción de variedades comerciales, es aconsejable preservar el germoplasma. **Conclusión.** Los métodos de agrupación revelaron que los 189 árboles de aguacate pueden reagruparse en tres conglomerados principales. Los descriptores morfológicos demostraron ser útiles para agrupar árboles según las razas conocidas y, por tanto, pueden utilizarse en la caracterización de cultivares y árboles silvestres cuya ascendencia se desconoce.

Palabras clave: *Persea americana*, diversidad, variación fenotípica, análisis factorial de datos mixtos, germoplasma silvestre, conservación.

INTRODUCTION

The avocado (*Persea americana* Mill.) belongs to the Lauraceae family, which includes about 50 genera and is considered one of the oldest angiosperm families (Renner, 1999, 2004). The avocado is a subtropical perennial species; nevertheless, it has been adapted to diverse climates (Galindo-Tovar and Arzate-Fernández, 2010). Therefore, its production has spread practically throughout the American continent, Asia, Africa continents, Middle East, and Europe (Guzmán *et al*., 2017). It is one of the most economically and culturally important fruit trees in Mesoamerican region its use in the region dates back more than 8,000 years, as shown by archaeological evidence in the same area (Smith, 1966), where the native people have maintained a generalized culture of knowledge, conservation, and use (Galindo-Tovar *et al*., 2008, 2013).

The current classification of avocado recognizes three horticultural races: Mexican (*P. americana* var. *drymifolia*), Guatemalan (*P. americana* var. *guatemalensis*) and West Indian (*P. americana* var. *americana*) (Bergh, 1992). Each of races maintains specific characteristics in their morphology, phenology, ecology and adaptation, and is part of the diversity of the genus (Storey *et al*., 1986). Only West Indian race is well suited to genuinely tropical climes, while the Guatemalan and Mexican accessions typically exhibit poor fruit sets (Bergh, 1992). The Guatemalan race is well adapted from about 1,000 to 2,000 m, while Mexican accessions remain in the colder avocado areas above 1,500 until 3,000 m. The three avocado races are compatible, and trees of various races can hybridize when cultivated close to one another and adapt to a wide range of environments (Storey *et al*., 1986). As one of the three main centers of domestication of the species, the Guatemalan avocado populations exhibit a high diversity with great potential for its use (Galindo-Tovar *et al*., 2013).

The avocado variability is manifested in various traits, such as the shape, color, size of fruit and leaves, flavor of fruits, and their phenology (Acosta-Díaz *et al*., 2013). The avocado crop is of high socio-economic importance in Guatemala's farming systems (Ledesma and Carter, 2020). At the same time, Guatemala is a country particularly susceptible to the effects of climate change (Lopez-Lalinde and Maierhofer, 2021). So it is crucial to protect and properly use the native diversity to develop new genotypes able to adapt to actual and coming adverse environmental conditions brought by climate change (Grüter *et al*., 2022). Nonetheless the wild Guatemalan avocado populations are increasingly displaced and threatened by various factors, such as increase in human population and consequent deforestation (Bullock *et al*., 2020), but mainly by the increase in cultivated areas with introduced uniform cultivars (Rincón-Hernández *et al*., 2011). Thus, developing new strategies for conservation and sustainable use of Guatemalan avocado genetic resources is necessary.

A crucial stage in the description and characterization of germplasm is the evaluation of variability. The morphological markers are suitable because they provide a practical method for determining the degree of variation. These readily visible morphological traits are helpful for conservation and sustainable use strategies (Ranjitha *et al*., 2021). The assessment of the genetic diversity of avocados using morphological characterization has been reported around the world, such as in Mexico (Rincón-Hernández *et al.*, 2011), Colombia (López-Galé *et al*., 2022), Tanzania (Juma *et al*., 2020), Ghana (Abraham *et al*., 2018), proving its validity and usefulness. In Guatemala, despite being a center of domestication, morphological characterization of populations not subjected to cultivation has not been carried out. These populations present a free genetic exchange unlike populations under cultivation that are usually inbred. Accordingly, the information limitations about Guatemala's avocado can lead to the loss of genes that can be useful for adaptability and resilience to climate change, as well as genes of interest for improving pest and disease resistance and fruit quality.

This study aims to explore and characterize the variability of wild avocado germplasm from Guatemala using 13 qualitative and 7 quantitative morphological traits. The outcomes of this investigation will support future planning for breeding strategies and germplasm management and conservation.

MATERIALS AND METHODS

Study site and sampling

Based on information from the Guatemalan atlas of wild relatives of cultivated plants (Azurdia *et al*., 2011) and with Rafael Landívar University Herbarium staff and local people familiar with wild avocados a reconnaissance of area was carried out to identify wild individuals. A total of 189 unique avocado trees were sampled these were organized into eight geographic populations, covering three different physiographic regions. The populations were in Sacatepéquez, Chimaltenango, Sololá, Totonicapán, Quiché, Huehuetenango, Alta Verapaz, and Baja Verapaz departments; located in central, western, and northern physiographic regions (Figure 1). Table 1 displays the ecological traits of each department examined for this study. The number of individuals per population varied from 8 to 36, due to the accessibility and availability of avocado trees in each population. When an avocado tree was located, the site's latitude, longitude, and elevation were recorded.

Table 1. Ecological characteristics of the studied locations.

Location	Code	N	Latitude	Longitude	Region	maT °C	arR (mm yr^{-1})	ar masl
Chimaltenango	Chi	23	90.683	14.595	Central	9.2	1,779 to 2,573	2,097 to 3962
Sacatepéquez-Chimaltenango	Sac-Chi	32	90.810	14.594				
Sacatepéquez	Sac	36	91.063	14.801				
Sololá	Sol	8	91.156	14.832				
Totonicapán-Quiché	To-Qui	36	91.416	15.024	Western	10.4	1,141 to 2,056	1,801 to 2,990
Huehuetenango	Hue	23	91.128	15.822				
Baja Verapaz	BV	11	89.943	15.204	Northern	15.8	1,850 to 3,410	784 to 1,877
Alta Verapaz	AV	20	90.509	15.423				

N: number of sampled trees, maT: mean annual temperature, arR: annual rainfall range, ar masl: altitudinal range in meters above sea level.

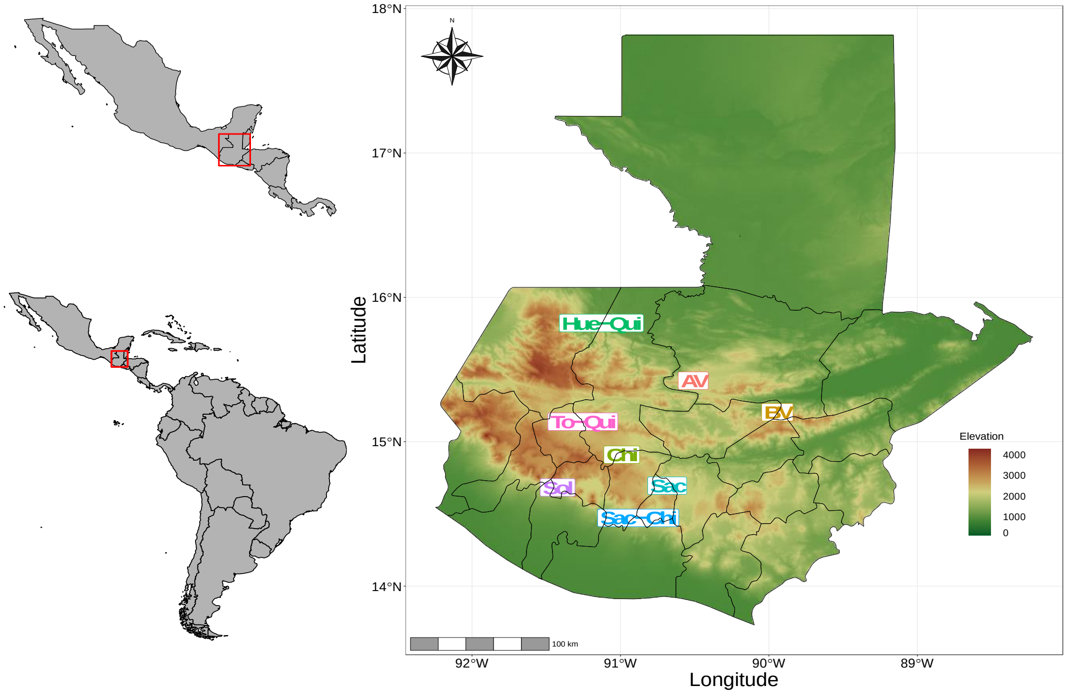

Figure 1. Map of Guatemala, displaying the geographical location of sampled avocado populations. Source: own elaboration.

Measurement of qualitative and quantitative morphological traits

The International Plant Genetic Resources Institute (IPGRI) field guide for avocado crops was used to characterize the tree trunk, young twigs, leaves of comparable age, flowers, ripe fruits, and seeds for each selected tree through 21 plant descriptors (IPGRI, 1995). Using the methodology described by Juma *et al*. (2020) each tree was examined for five twigs, three to five leaves, two to four fruits, two to three flowers, and three seeds. These descriptions make differentiating between phenotypes simple and rapid they are often quite observable, highly heritable, and equally exhibited in all environments (IPGRI, 1995). Highly discriminating descriptors available during the study period were used. The 21 descriptors registered and their possible variants are presented in Table 2.

The flesh texture was determined by tasting the ripped fruit, trunk circumference was determined with measuring tape and vernier caliper was used to record leaf length, width, and sepal length. The fruit and seed weight were measured using a portable semi-analytical balance. Bare hands were used to feel the smoothness or roughness of trunk's surface. Samples' leaf, fruit, pedicel, and seed shapes were compared to corresponding pictures in field guide. Cotyledon surface smoothness or roughness was determined by touching the seed with bare hands. Anise odor was identified by crushing and smelling the leaves. Color of a young twig, a mature leaf, petal pubescent, and mature fruit skin color were recorded based on visual observation.

Data analysis

The compareGroups package (Subirana *et al*., 2014) in R software v.4.2.0 (R Core Team, 2022) was employed to assess statistical significance of the factor location on measured morphological values. We used a one-way analysis of variance (ANOVA) and Tukey's tests at a significance threshold = 0.05 for each quantitative morphological value. The coefficient of variation was estimated to determine the degree of variability in each quantitative attribute across populations. GGally package (Schloerke *et al*., 2021) was used to estimate a correlation matrix a *p*-value of < 0.*05* was considered statistically significant. Data on qualitative morphological attributes were treated to a cross-tabulation statistical approach to ascertain the frequency distribution of the features amongst populations. The Pearson Chi-square (χ2) test was used to ascertain if the cross-tabulation variables were related. The the ggstatsplot package (Patil, 2021) implemented in R was used to carry out the cross-tabulation and Chi-square tests.

Table 2. List of quantitative and qualitative traits assessed and their alternative variants.

Tree part	Quantitative character	Abbreviation	Measurement unit
Overall tree	Trunk circumference	TC	cm
	Leaf length	LL	mm
	Leaf width	LW	mm
Flower	Sepal length	SL	mm
Fruit	Fruit weight	FW	g
	Fruit length	FL	cm
	Seed weight	SW	g
	Qualitative character		**Level**
Overall tree	Trunk surface	TS	even, rugged, very rugged
	Color young twig	CYT	yellow, green, coopery, maroon, red
	Color mature leaf	CML	green, dark green
	Leaf shape	LS	ovate, narrowly, obovate, oval, roundish, cordiform, lanceolate, oblong, oblong-lanceolate
	Leaf anise smell	AS	absent, intense
Flower	Petal pubescent	PP	scarce, intermediate, dense
Fruit	Pedicel shape	PS	cylindrical, conical, rounded
	Fruit skin surface	FSS	even, intermediate, rugged
	Mature fruit skin color	MFSC	clear green, green, dark green, yellow, red, purple, black
	Fruit shape	FS	oblate, spheroid, high spheroid, ellipsoid, narrowly obovate, obovate, pyriform, clavate, rhomboidal
	Flesh texture	FT	watery, buttery, doughy, granular
Seed	Seed shape	SS	oblate, spheroid, ellipsoid, ovate, broadly ovate, cordiform, base flattened apex rounded, base flattened apex conical
	Cotyledon surface	CS	smooth, intermediate, rough

The factor analysis of mixed data (FAMD) was performed to explore the morphological relationships of the 189 avocado trees and the association between all quantitative and qualitative variables. The FAMD method may be thought of as combining principal component analysis (PCA) and multiple correspondence analysis (MCA). FAMD is dedicated to analyzing a data set containing quantitative and qualitative variables to balance their influence in the analysis (Pages, 2004). Since variables measured at different scales do not contribute equally to the FAMD analysis, the variables were standardized prior to analysis (Kenkel, 2006), thus making it possible to optimize the variance explained in each of the dimensions. FactoMineR (Lê *et al.*, 2008) and factoextra packages (Kassambara and Mundt, 2020) was used to extract the FAMD results.

Hierarchical Clustering on Principal Components (HCPC) based on FAMD analysis was applied to identify the clustering and population structure of the sampled trees. HCPC is a robust method combining principal component methods, hierarchical clustering, and partitioning clustering, particularly the k-means method, which are the three standard methods used in multivariate data analyses (Husson *et al.*, 2010). HCPC results was visualized through a dendrogram and factor map using the factoextra package. For the purpose of customizing and visualizing the dendrogram was exported in Newick tree format to the Interactive Tree Of Life v6 (iTOL) (Letunic and Bork, 2019). The test values (Morineau, 1984) were computed and used to rank the most distinctive quantitative and qualitative categories within each cluster found by HCPC.

RESULTS AND DISCUSSION

Quantitative traits

The description of quantitative traits of each of eight geographic avocado populations is displayed in Table 3. About overall tree traits among all populations, highest trunk circumference TC was

118.33 cm from Sac-Chi and Chi population exhibited lowest TC value with 82.59 cm. The coefficient of variation for TC ranged of 10.55 % (Sac-Chi) to 33.4 % (BV) and mean of 22.2 %. The data reported in this study are lower than those reported in the characterization conducted in wild avocados from Ghana, with mean TC values of 133.04 cm (Abraham *et al*., 2018).

Fruit weight and length ranged from 164.8 g (AV) to 393.82 g (Hue-Qui) and 9.85 cm (To-Qui) to 13.13 cm (BV), respectively. For both fruits traits, the values are consistent with those reported from Mexico (López-Guzmán *et al*., 2015) and Colombia (López-Galé *et al.*, 2022).

For seed traits, the weight varied from 75.04 g (AV) to 96.02 g (Sac), with an overall mean of 86.56 g and the highest and lowest CV of 22.98% (Chi) to 14.26 % (Sac-Chi). The high variation in each qualitative trait was also reported in the characterization of Mexican avocado germplasm (López-Guzmán *et al*., 2015; Rincón-Hernández *et al., 2011)*. With the analysis of variance, excluding leaf width and petal length, all quantitative traits exhibited statistically significant ($p < 0.05$) differences when compared between populations. Based on Tukey's test, Sac and To-Qui locations had the highest average values for fruit weight and length, while the lowest values were found in AV (Table 3). The quantitative morphological traits assessed revealed significant diversity in all eight populations. There was CV with more than 20 % for 87.5 % of the descriptors taken into consideration. A larger proportion for a property may suggest greater variability (Hidalgo, 2003). The fruit weight (FW) was the trait that varied the most between the populations. This variability might

Table 3. Description of quantitative traits among geographical populations. Mean values, standard deviation (SD) and coefficient of variation (CV) complemented by Tukey post-hoc test.

L	A	E	TC (cm)	LL (cm)	LW (cm)	SL (mm)	FW (g)	FL (cm)	PL (mm)	SW (g)
Sac	1,725	mean	118.22^{a}	20.61^{b}	12.82	3.51^{ab}	364.83^{a}	11.29^{ab}	3.55	96.02^{a}
		SD	13.042	5.49	3.873	0.636	85.133	2.745	0.281	18.5
		CV	11.03	26.66	30.21	18.13	23.34	24.32	7.93	19.27
Chi	2,054	mean	118.33^{a}	20.34^{b}	12.06	3.57^{ab}	260.04^{b}	13.04^{a}	3.5	82.68^{ab}
		SD	12.478	4.479	3.161	0.701	69.417	1.95	0.276	11.791
		CV	10.55	22.02	26.21	19.64	26.69	14.96	7.91	14.26
Sac-Chi	2,305	mean	82.59^{c}	24.68^{ab}	13.8	3.43^{ab}	305.09^{b}	10.64^{ab}	3.55	84.05^{ab}
		SD	25.555	6.283	3.39	0.838	75.323	2.933	0.332	19.32
		CV	30.94	25.46	24.57	24.44	24.69	27.57	9.35	22.98
Sol	3,315	mean	117.99^{ab}	20.31^{b}	13.36	3.07^{b}	349.51^{ab}	11.13^{ab}	3.59	88.24^{ab}
		SD	13.071	7.203	3.048	0.659	57.75	2.345	0.267	13.283
		CV	11.08	35.47	22.82	21.46	16.52	21.06	7.45	15.05
Hue-Qui	1,469	mean	93.2^{bc}	22.51^{ab}	13.68	4.03^{a}	295.04^{b}	9.85^{b}	3.37	79.69^{b}
		SD	23.295	7.566	2.532	0.708	54.382	2.948	0.42	17.684
		CV	25	33.6	18.5	17.7	18.4	29.9	12.3	22.2
To-Qui	2,820	mean	104.64^{ab}	24.83^{a}	12.45	3.26^{b}	393.82^{a}	12.08^{a}	3.46	95.28^{a}
		SD	18.145	5.387	4.043	0.93	74.263	2.877	0.29	15.161
		CV	17.34	21.7	32.47	28.53	18.86	23.82	8.34	15.91
BV	810	mean	103.21^{abc}	26.91^{a}	10.11	3.17^{b}	295.53^{b}	13.13^{a}	3.31	76.5^{b}
		SD	34.498	5.548	3.289	1.095	46.472	2.377	0.426	12.77
		CV	33.4	20.6	32.5	34.5	15.7	18.1	12.9	16.7
AV	1294	mean	105.79^{ab}	20.66^{ab}	12.9	3.88^{ab}	164.8^{c}	12.05^{ab}	3.41	75.04^{b}
		SD	13.731	4.885	3.74	0.636	91.705	1.501	0.39	16.99
		CV	13	23.6	29	16.4	55.6	12.5	11.4	22.6
Total	--	mean	104.37	22.66	12.8	3.52	313.14	11.55	3.48	86.56
		SD	23.14	6.11	3.57	0.81	99.39	2.75	0.33	18.05
		CV	22.2	27	27.9	23.14	31.7	23.8	9.59	20.9

L: location, A: altitude in meters above sea level, E: estimator, N: number of sampled trees, TC: trunk circumference, LL: leaf length, LW: leaf width, SL: sepal length, FW: fruit weight, FL: fruit length, PL: pedicel length, SW: seed weight. Different letters indicate significant differences ($p< 0.05$)

be accounted for by the different genetic basis of each population as well by the existence of the microenvironments and various agro-ecological circumstances between the populations that directly influence the fruit size and weight (Chen *et al*., 2007; Henao-Rojas *et al*., 2019; Cañas-Gutiérrez *et al*., 2022).

Qualitative traits

The studied trees showed three trunk surface variations, with even, rugged, and very rugged appearing in 54.5, 24.9, and 20.6 % of all samples, respectively. There was a statistically supported association between trunk surface and populations (χ^2 = 29.21, df = 14, p < 0.001) (Figure 2A). Rough and very rough trunk surfaces were predominant, which is in accordance with the study area because most of sampled trees were located in medium or highlands and just a few in lowlands. It has been reported that tree bark Guatemalan and Mexican avocado races, mainly established over 1,500 masl, is less rough. In contrast, those of West Indian race are rougher and typically well-suited to lowlands (Bergh, 1992).

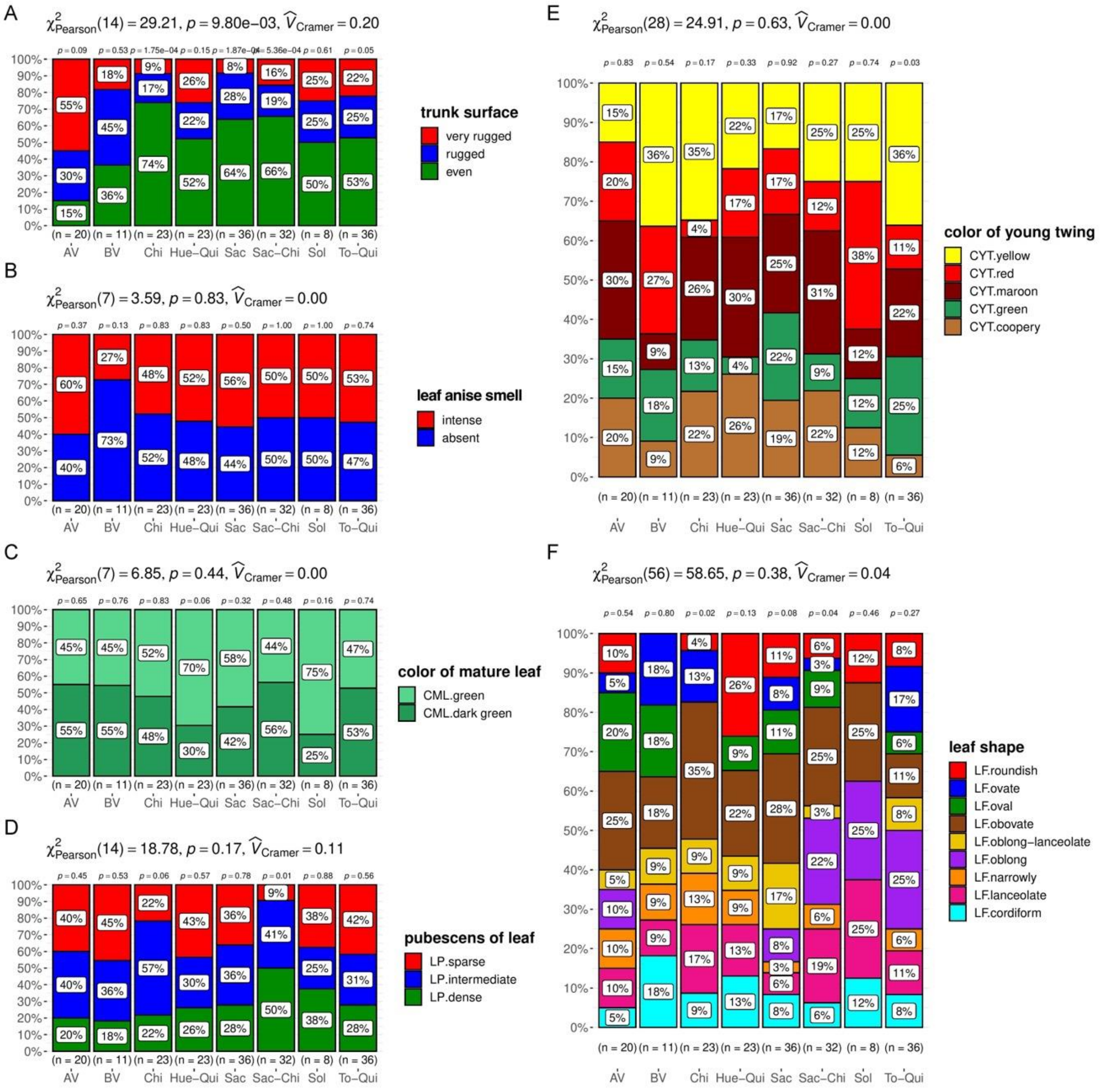

Figure 2. Description of frequency and chi-squared test of the qualitative traits of the overall tree. A panel: trunk surface. B panel: leaf anise smell. C panel: color of the mature leaf. D panel: pubescence of leaf. E panel: color of the young twig. F panel: leaf shape

Regarding anise aroma in leaves, 48 % of sampled trees did not show this aroma. In comparison, in remaining 52 %, it was present, and statistical analysis did not find an association with populations ($p > 0.05$) (Figure 2B). Anise-scented smell is linked to abundance of estragole, an organic compound detected only in Mexican avocado cultivars (Pino *et al*., 2006; Pereira *et al*., 2013).

Concerning pubescence leaves, it showed sparse, intermediate, and dense with a distribution of 32.80, 37.56 and 29.62 %, respectively. There was not association between leaf pubescence and populations ($p > 0.05$) (Figure 2D). The presence of leaf pubescence can lessen the photosynthetic light plants absorb. Through growth season, pubescence rises, changing energy balance of leaf and significantly slowing down photosynthetic activity (Ehleringer *et al*., 1976). Nevertheless, it has been reported that leaf pubescence can increase water use efficiency by promoting condensation (Konrad *et al*., 2015), which is an essential trait for adaptation to drier environments, especially in areas vulnerable to drastic environmental variations due to climate change.

Nine different leaf shapes were identified in avocado characterization from Guatemala, and there was no statistical association between this trait and geographical populations ($p > 0.05$) (Figure 2F). The leaf form is a crucial characteristic since it can indicate how much leaf area there is and, therefore, how much light is reflected, which can directly affect plant productivity (Nkansah *et al*., 2013). Some of these shapes have been previously reported in avocado germplasm characterized from Colombia (López-Galé *et al*., 2022), Mexico (Acosta-Díaz *et al*., 2020), Tanzania (Juma *et al.*, 2020) and Ghana (Abraham *et al*., 2018). The higher number of leaf shapes (9) described in present study is comparable with the reported in Tanzania (Juma *et al*., 2020). It could be due to a larger sample size because both studies were focused on description of wild populations and not only core collections.

The avocado fruit displayed three and nine pedicel and fruit shapes, respectively (Figure 3A, 3B). Avocado fruit shape is a commercially important morphological trait that appeals to market consumers. Having great diversity in fruit shape and mature skin color offers a chance to diverse interest customers. Most of shapes recorded among sampled tree were previously published by Juma *et al*. (2020). They also provided details on relationship between various shapes and cultivars developed from three horticultural races of avocados. The existence of these fruit morphologies in the studied trees shows presence of genetic material from all three avocado races in Guatemala.

For fruit flesh texture, classes buttery (34.86%), watery (21.16%), pastose (26.45%), and granular (17.51%) were described (Figure 3E). Since most Guatemalan avocado buyers like it for its flavor, buttery texture has significant economic importance. Many Guatemalan avocado farmers have chosen varieties with a buttery feel. Mexican and certain Guatemalan avocados are typically said to have a buttery flesh texture (Bost *et al*., 2013). Similarly, some studies noted that Mexican and Guatemalan avocado cultivars had moderate to high oil contents, which may be related to buttery and pastose flesh textures (Pereira *et al*., 2013; Espinosa-Alonso *et al*., 2017). In current study, the presence of buttery and pastose flesh textures suggests the presence of avocados from Mexican and Guatemalan races. In contrast, presence of watery flesh textures indicates presence of avocados from West Indian race. Interestingly, during fruit collection, some producers indicated that some avocado trees produced fruits with buttery flesh in dry seasons and watery texture in rainy seasons. A similar phenomenon was reported by Juma *et al*. (2020), who suggest that environmental factors may influence changes in texture, so it is essential to consider that fruit quality may be affected in future by climate change, especially in vulnerable regions such, as Guatemala.

Avocado fruits displayed eight seed shapes. Data on distribution of seed shapes encountered are reported in Figure 4A. The Chi-squared test was not significant ($p < 0.05$) except for BV, which exhibited an association with oblate and cordiform seed shape. For the seed cotyledon surface smooth, rough, and intermediate texture were recorded with a distribution of 36.51, 35.97 and 27.51%, respectively, among the eight populations. Statistically, a significant association was not detected between the cotyledon surface and the populations (Figure 4B).

The wide variety of seed forms identified in this study agrees with the report based on the morphological characterization of Tanzanian avocados, where about 17 different seed forms are described (Juma *et al*., 2020). On the other hand, the Indian and Colombian avocado characterization reports six and three different seed forms, respectively (López-Galé *et al*., 2022; Ranjitha *et al*., 2021). The wide difference between this study and the Indian and Colombian reports may be due to the larger sample size and the greater genetic diversity expressed through the various seed shapes. Popenoe (1974) linked the spheroid, obovate and oblong-conic seed shapes with Guatemalan, West Indian, and Mexican avocado races. The current study's observation of these seed morphologies suggests that Guatemala's wild avocado germplasm is derived from all three avocado races.

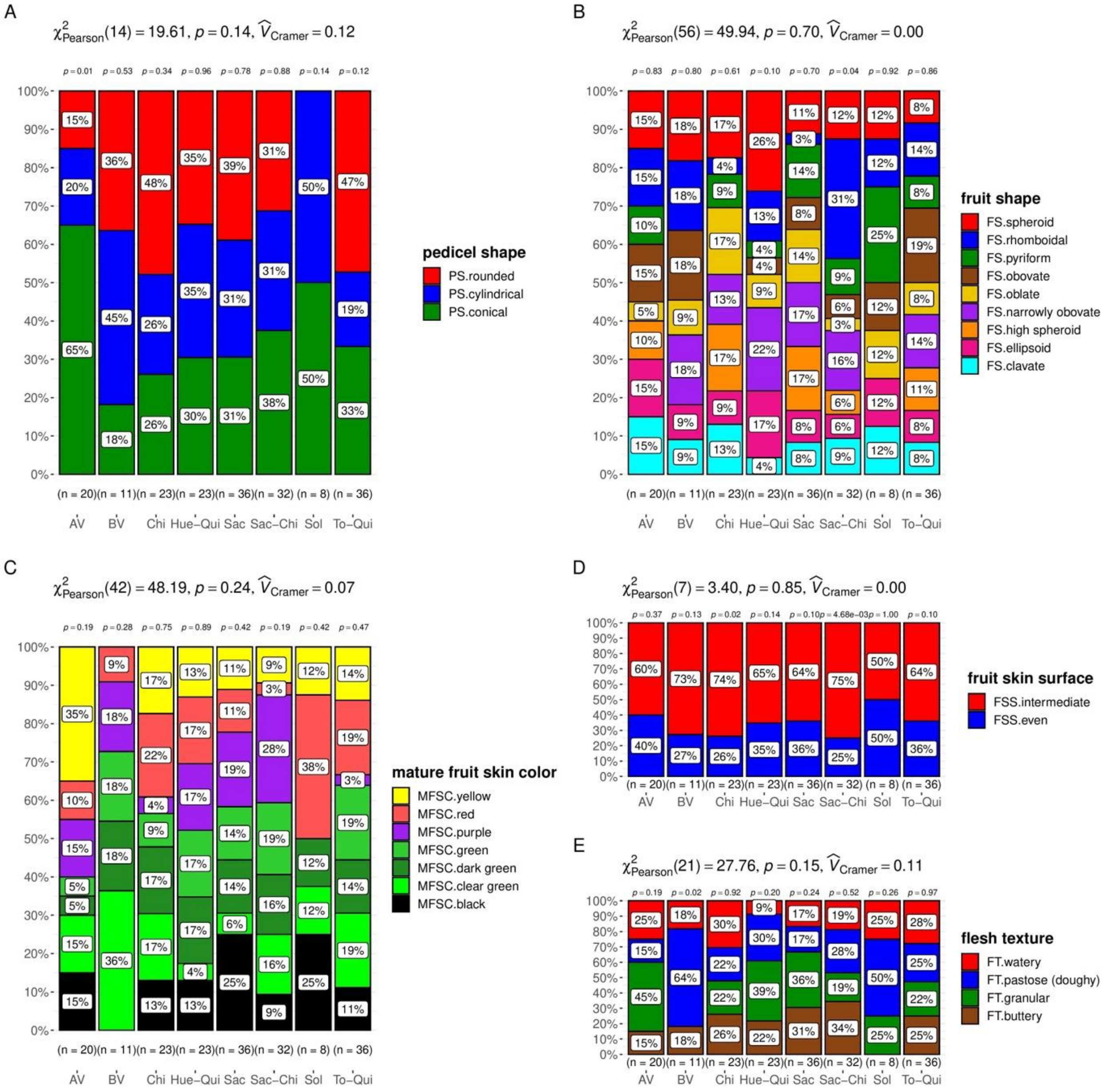

Figure 3. Description of frequency and chi-squared test of the fruit's qualitative traits. A panel: pedicel shape. B panel: fruit shape. C panel: mature fruit skin color. D panel: fruit skin surface. E panel: flesh texture.

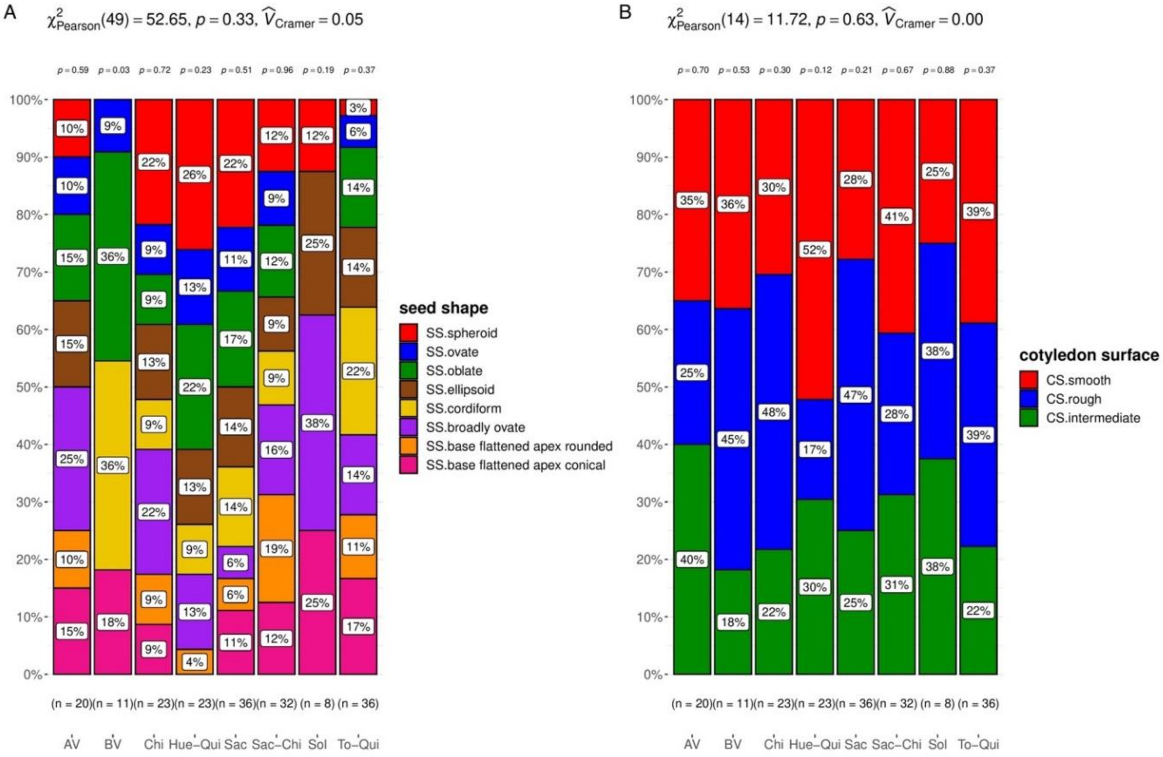

Figure 4. Description of frequency and chi-squared test of the avocado seed qualitative traits. A panel: seed shape. B panel: cotyledon surface.

The analyzed avocado plants exhibited a wide variety of the traits investigated. Some of these characteristics, such as fruit form, the color of the ripe fruit's skin, and flesh texture, have economic value and may be utilized by farmers as selection criteria for future production and by breeders to create superior cultivars. According to Barrett *et al.* (2010) the externally brilliant color, shine, and fruit form entice buyers and cause them to make impulsive purchases. When consumers eat the fruit, the texture, newness, and other flavor characteristics are pivotal to the consumers' enjoyment (Barrett *et al*., 2010). Consumers utilize visual cues to judge the freshness and flavor quality of food at the moment of sale, albeit these cues can occasionally be deceptive (Shewfelt, 2000; Barrett *et al*., 2010).

FAMD analysis

The FAMD analysis revealed that dimensions 1, 2, and 3 retained 21.1, 15.6, and 14.9 % of the total variance, respectively. At the same time, 75 % of the cumulative variance was achieved with 5 dimensions. Regarding the correlation between quantitative and qualitative variables, the fruit shape and skin color, seed shape, flesh texture and anise odor in the leaves contributed 13.85, 13.72, 13.56, 12.62 and 5.03 % to the two first dimensions (Figure 5B). These characteristics have been used to discriminate between avocado horticultural races. For example, anise odor in the leaves, fruits with rough surfaces and rough cotyledon surfaces are traits linked to Mexican, Guatemalan and West Indian races, respectively (Bergh, 1992; Janick, 2005; Popenoe, 1935).

We next used coordinates, squared cosine ($\cos^2$) and contribution to assess a correlation between variables and the primary two dimensions. The angle between variable point and the axis, or $\cos^2$, revealed the quality of the factor map's depiction of variables. Additionally, the contribution was calculated using the separation between variable's perpendicularly projected point and associated dimension axis. According to whether a variable has a positive or negative contribution, it is placed in a different quadrant of the two dimensions. Several variables were highlighted in a redder color, such as fruit weight and seed weight in the quantitative plot (Figure S1), as well as FT. buttery, FSS.even, and SS.spheroid in qualitative plot (Figure S2).

The FAMD revealed an association between all quantitative and qualitative variables of the avocado germplasm analyzed and usefulness for profiling among sampled trees (Figure S3). In this study, the FAMD was very useful in demonstrating that the analyzed trees, which compose wild Guatemalan avocado germplasm, exhibit a mixture of different levels of each of analyzed qualitative characteristics. It may be due to the development of hybrids between three horticultural races due to absence of sterility barriers and floral biology, that favors cross-pollination (Alcaraz and Hormaza, 2011; Gross-German and Viruel, 2013). However, the FAMD demonstrated that the features that can be beneficial for distinguishing between individuals are those that most significantly contributed to dimensions 1 and 2, such as fruit shape and skin color, seed shape, flesh texture and anise odor in the leaves.

Separate explorations of connections between quantitative and qualitative factors were conducted. Consequently, the angles in correlation circle for each pair of quantitative variables represented their correlations (Figure S1). The orthogonal angle revealed a weak association, whereas the acute and obtuse angles suggested positive and negative correlations. After the correlation analysis (Table 4) the variables of fruit weight and seed weight demonstrated a positive correlation ($r = .49$, $p < 0.0001$). For conducting pre-harvest yield estimation and tracking shifts in avocado production, the relationship between fruit weight and length is ideal (Mokria *et al*., 2022). The remaining variables were not correlated.

Population structure

Through FAMD analysis, individual data points were visualized in the new feature space created by the first three, and thus the most informative, dimensions (Figure 6A). There was no prominent grouping of samples in the FAMD individuals plotting the avocado's qualitative and quantitative morphological attributes, which showed significant variance among the examined trees at the population level (Figure 6A). It was difficult to see the differences across the populations since there was a high phenotypic divergence within each area, revealing an absence of differentiation among populations. The considerable variety seen may have been influenced by the fact that avocado is a highly heterozygous cross-fertilizing species that produces fruits with monoembryonic seeds, so its progeny can be highly variable (Alberti *et al*., 2018). In this sense, it is to be expected that the evaluated populations, where repopulation processes happen naturally with little selection interference from producers, will exhibit high morphological variability, particularly in fruit and seed characteristics, which are those with the most significant natural variability (López-Galé *et al*., 2022).

Cluster analysis

The output of the hierarchical clustering algorithm on principal components (HCPC) resulted in a factor map and dendrogram. The HCPC suggested the presence of three clusters across 189 avocado trees. The visualization of cluster on factor map (Figure 6B) revealed that each one exhibited the presence of trees from the eight populations. Like factor map, the dendrogram from a hierarchical cluster analysis revealed three distinct groupings, each including samples from different populations, suggesting a weak population's genetic structure. The largest group was depicted by green branches, followed by blue branches, and finally, the smallest group was painted red, all containing samples from all populations (Figure 6C).

The clustering of avocado populations from various geographical locations points to a genetic mixing that may have resulted via seed exchanges between these places. Farmers in various locations frequently trade seeds with friends and family. Additionally, avocado fruits from one population can be swapped for fruits from another, which may then be used to sow the seeds. In addition, the avocado's religious, mythical, economic, and therapeutic significance in Guatemala—one of the three domestication centers— (Galindo-Tovar *et al*., 2008; Landon, 2009) could explain the genetic mingling seen among populations of avocados. This condition still exists now and has taken the lead role in influencing genetic diversity and population structure.

From HCPC analysis all quantitative characters had statistically significant links with the clusters' partition ($p < 0.0001$) with the largest explained variance for FW, SW and FL (FW: $eta^2 = 0.16$, $p < 0.0001$; SW: $eta^2 = 0.14$, $p < 0.0001$; TC: $eta^2 = 0.11$, $p < 0.001$). Chi-square test for qualitative characters exhibited statistically significant links with the clusters ($p < 0.0001$). In first cluster 90 and 61.19 % of individuals showed a rugged fruit skin surface and smooth cotyledon surface, respectively. The presence of anise odor in the leaves was not highly represented in this cluster (v.test -3.51). The lowest SW (v.test = -3.30) was found in this cluster (Table 5).

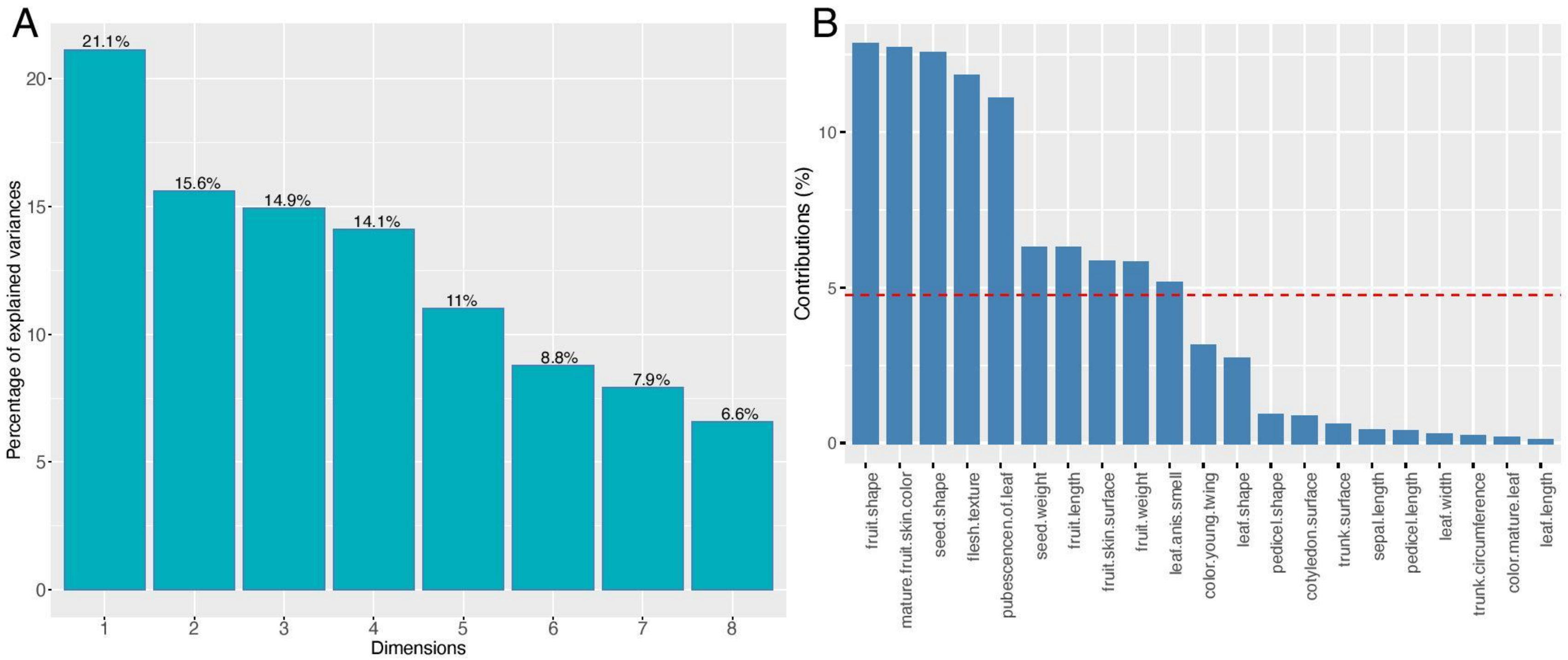

Figure 5. Scree plot showing the percentage of variance explained by each of the first 8 dimensions (A). Contribution of avocado variables to Dimension 1 and 2 (B). The red dashed line on the graph above indicates the expected average value, if the contributions were uniform.

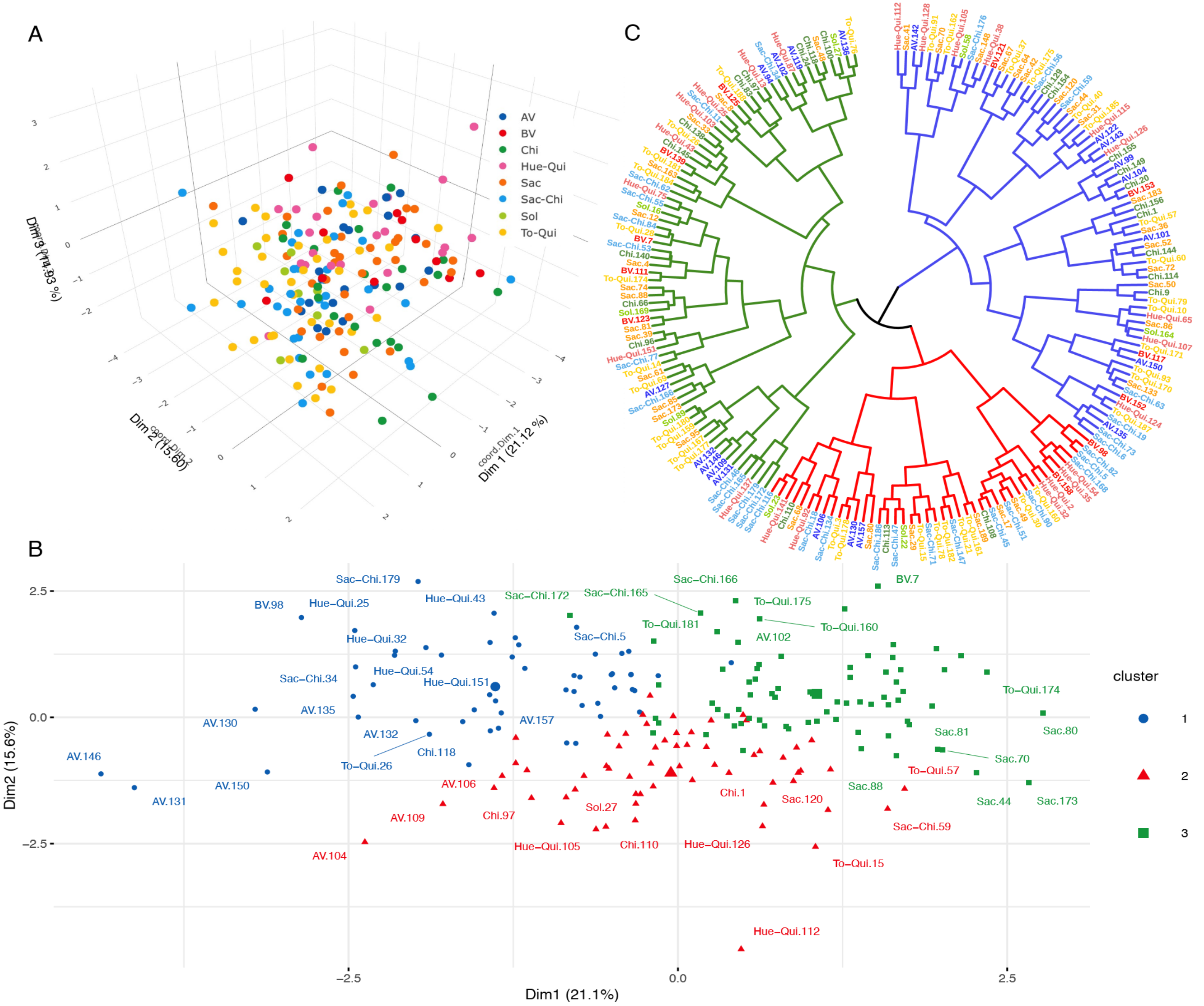

Figure 6. Avocado individuals' visualization-based A) the first three dimensions based on FAMD, B) a factor map constructed from the HCPC algorithm and C) a dendrogram from the hierarchical cluster analysis. In figure C, the same font color is used to indicate samples from the same population.

Table 4. Correlation matrix showing Pearson's correlation coefficient between each quantitative variable.

	Fruit weight	Seed weight	Fruit length	Pedicel length	Leaf length	Leaf width	Sepal length	Trunk circumference
Fruit weight	---							
Seed weight	.49****							
Fruit length	0.12**	0.1						
Pedicel length	0.09	0.11	-0.1					
Leaf length	0.06	0	-0.1	-0.14				
Leaf width	-0.04	-0.06	0	0.06	-0.03			
Sepal length	-0.19**	-0.14	-0.29***	-0.03	0.03	-0.03		
Trunk circumference	0.24**	0.08	0.11	0.01	-0.25***	-0.05	0.01	--

$p < 0.0001$ '****'; $p < 0.001$ '***', $p < 0.01$ '**', $p < 0.05$ '*'

Table 5. Description of each cluster by the quantitative and qualitative character categories based on Hierarchical Clustering on Principal Components (HCPC).

Quantitative trait	Mean in category	Overall mean	v.test	p value	Qualitative trait	Cla/ Mod	Mod/ Cla	v.test	p value
Cluster 01									
FW	303.14	289.62	2.58	> 0.0001	FSS rugged	90	34.18	5.9	> 0.0001
FL	12.59	11.54	2.23	> 0.0001	MFSC clear green	88.89	20.25	4.23	> 0.0001
SW	81.09	86.55	-3.30	> 0.0001	CS smooth	61.19	51.9	3.96	> 0.0001
					FS pyriform	87.5	17.72	3.82	> 0.0001
					AS present	9.9	39.73	-3.51	0.001
Cluster 02									
SW	90.56	86.55	2.58	> 0.0001	FT buttery	91.67	30.14	5.71	> 0.0001
FW	282.62	313.14	-2.98	> 0.0001	AS intense	68.97	27.4	3.54	> 0.0001
FL	22.66	25.65	-5.33	> 0.0001	FSS even	60.17	68.49	2.72	> 0.001
					SS oblate	47.83	60.27	2.51	0.012
					LS lanceolate	14.81	5.48	-2.53	0.012
Cluster 03									
FW	313.05	289.62	3.78	> 0.0001	PS rounded	77.92	32.16	5.33	> 0.0001
FL	13.11	11.54	2.57	> 0.0001	PP scarse	70.59	32.43	4.76	> 0.0001
LL	20.84	22.66	3.24	> 0.0001	FSS intermediate	37.31	67.57	4.41	> 0.0001
					CS rough	62.54	27.03	3.89	> 0.0001
					CYT maroon	7.46	13.51	-3.21	> 0.001

TC: trunk circumference, FW: fruit weight, FL: fruit length, LL: leaf length, SW: seed weight. The sign of the v.test indicates if the mean of the cluster is under or over-expressed for the category. FSS: fruit skin surface, MFSC: mature fruit skin color, CS: cotyledon surface, FS: fruit shape, AS: anise smell, CYT: color young twig, PP: petal pubescent, PS: pedicel shape, FT: flesh texture, SS: seed shape, LS: leaf shape, TS: trunk surface, Cla/Mod: proportion (expressed as percentages) of individuals with specific qualitative character category in the cluster; Mod/Cla: proportion (expressed as percentages) of individuals within the cluster with the specific qualitative character category.

These traits fit with description of Guatemalan race, which is characterized by absence of an anise-like fragrance in the leaves, fruits with rough surfaces usually, and seed that is generally not large in proportion to the fruit size (Bergh, 1992; Janick, 2005; Popenoe, 1935). This similarity of morphological characteristics suggests that the genetic pool of the first cluster comes from the Guatemalan race.

In the second cluster 91.67, 68.97 and 60.17 % of the individuals exhibited buttery flesh texture, anise aroma in the leaves and even fruit skin surface, respectively. Agreeing, the individuals showed the lowest FW (v.test -2.98) and significantly higher SW (v.test = 2.58) (Table 5). The qualitative and quantitative characteristics of second cluster are in accordance with the botanical description of the Mexican race, which is characterized by fruits with high oil content, giving a buttery texture, smooth surface and usually under one-half pound in weight, as well as the typical anise-like odor in leaves (Ayala-Silva and Ledesma, 2014; Bergh, 1992). This finding hints that the genetic basis of the individuals in the second group is largely derived from the Mexican race.

In third cluster 77.92, 70.59 and 62.54 % of individuals showed rounded fruit shape, scares petal pubescent and rough cotyledon surface, respectively. The highest values of FW and FL were found in this cluster (Table 5). Both qualitative and quantitative characters seem to be quite close to the description of the West Indian race, which is recognized because the leaves have no anise scent, fruit and seed size ranges from large to very large, the fruits are usually pear-shaped or round and the cotyledons are often rough on the surface (Ayala-Silva and Ledesma, 2014; Popenoe, 1935). The similarities in morphological characteristics point to the third cluster exhibiting a genetic basis from the West Indian race.

The data supplied can also be utilized to identify which individuals own genotypes most similar to avocado races mentioned. This information can be used to create germplasm collection to maintain the natural genetic variability linked to each Guatemalan avocado race, which may then be exploited in further breeding programs.

As demonstrated, the genetic basis of each cluster can be linked to one of the avocado races. However, the clusters also showed traits that are not typical to the related race, which points to a high degree of hybridization, probably due to the lack of sterility barriers between the avocado horticultural races (Alcaraz and Hormaza, 2011; Gross-German and Viruel, 2013). Thus, inter-breeding is possible, resulting in highly diverse genotypes adapted to the country's climatic conditions. The high genetic diversity was also reported in Guatemala using AFLP and SSR molecular markers. Besides, the lack of population structure found in this study is similar to the one revealed by the molecular markers (Ruiz-Chután *et al*., 2020; Ruiz-Chután *et al*., 2022).

CONCLUSION

This study represents the first characterization of wild Guatemalan avocado trees based on morphological markers. The FAMD revealed that the fruit and seed weight, seed and fruit shapes, and fruit length are ideal traits for distinguishing between individual trees and clusters. In addition, the clustering methods revealed that the 189 avocado trees could be regrouped into three main clusters. The morphological descriptors proved to be useful for grouping trees according to known races and can therefore be used in the characterization of cultivars and wild trees whose ancestry is unknown.

Acknowledgments
The authors would like to thank San Carlos University and Czech University of Life Sciences Prague, for financing this research.

Funding: This research had financial support from the budget assigned to Dirección General de Investigación (Digi) through the 4.8.63.4.41 project and The Internal Grant Agency of CULS Prague (No. 20223103).

Conflict of interests: The authors declare that there is no conflict of interest related to this publication.

Compliance with ethical standards: No ethical approval was required as any animal or feed was not used in this study.

Data availability. Data are available with J.A. Ruiz-Chután (josealejandro.ruiz@icloud.com), upon reasonable request

Author contribution statement (CRedit). J.A. Ruiz-Chután – Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing –review & editing., **J.E. Berdúo-Sandoval –** Investigation, Methodology, Writing – original draft., **A. Maňourová –** Writing – review & editing, Formal analysis., **M. Kalousová –** Writing – review & editing, Methodology., **C.A. Villanueva-González –** Investigation, Methodology., **E. Fernández –** Data curation, Formal analysis., **J. Žiarovská –** Data curation, Formal analysis., **A. Sánchez-Pérez –** Investigation, Methodology., **B. Lojka –** Funding acquisition**,** Methodology, Writing – review & editing.

REFERENCES

Abraham, J.D., Abraham, J. and Takrama, J.F., 2018. Morphological characteristics of avocado (*Persea americana* Mill.) in Ghana. *African Journal of Plant Science*, 12, pp. 88–97. https://doi.org/10.5897/ajps2017.1625

Acosta-Díaz, E., Almeyda-León, I.H. and Hernández-Torres, I., 2013. Evaluación de aguacates criollos en Nuevo León, México: región norte. *Revista Mexicana de Ciencias Agrícolas,* 4, pp. 531–542. https://doi.org/10.29312/remexca.v4i4.1184

Acosta-Díaz, E., Álvarez-Ojeda, M.G., Rodríguez-Guerra, R. and Almeyda-León, I.H., 2020. Variabilidad morfológica y genética del germoplasma de aguacate criollo en Nuevo León, México. *Biotecnología y Sustentabilidad*, 5, pp. 91–108. Retrieved November 13, 2022. https://revistas.uaz.edu.mx/index.php/biotecnologiaysustent/article/view/929/

Alberti, M.F., Brogio, B. do A., Silva, S. da R., Cantuarias-Avilés, T. and Fassio, C., 2018. Avances en la propagación del aguacate. *Revista Brasileira de Fruticultura*, *40*, pp. 1–18. https://doi.org/10.1590/0100-29452018782

Alcaraz, M.L. and Hormaza, J.I., 2011. Influence of physical distance between cultivars on yield, outcrossing rate and selective fruit drop in avocado (*Persea americana*, Lauraceae). *Annals of Applied Biology*, 158, pp. 354–361. https://doi.org/10.1111/j.1744-7348.2011.00469.x

Ayala-Silva, T. and Ledesma, N. 2014. Avocado history, biodiversity and production. *Sustainable Horticultural Systems*, 2, pp. 157–205. https://doi.org/10.1007/978-3-319-06904-3_8

Azurdia, C., Williams, K., Williams, D., Van Damme, V., Jarvis, A. and Castaño, S., 2011. Guatemalan atlas of crop wild relatives. Retrieved October 01, 2022. http://www.ars.usda.gov/Services/docs.html?docid=22225

Barrett, D.M., Beaulieu, J.C. and Shewfelt, R., 2010. Color, flavor, texture, and nutritional quality of fresh-cut fruits and vegetables: Desirable levels, instrumental and sensory measurement, and the effects of processing. *Critical Reviews in Food Science and Nutrition*, 50, pp. 369–389. https://doi.org/10.1080/10408391003626322

Bergh, B., 1992. The origin, nature, and genetic improvement of the avocado. *California Avocado Society*, 76, pp. 61–75. Retrieved November 02, 2022. http://www.avocadosource.com/cas_yearbooks/cas_76_1992/cas_1992_pg_061.pdf

Bost, J.B., Smith, N.J. and Crane, J., 2013. History, distribution and uses. In B. Schaffer, B. Wolstenholme, and A. W. Whiley (Eds.), The avocado: Botany, production and uses (2nd ed., pp. 10–30). CAB International.

Bullock, E.L., Nolte, C., Segovia, A.R. and Woodcock, C.E., 2020. Ongoing forest disturbance in Guatemala's protected areas. *Remote Sensing in Ecology and Conservation*, 6, pp. 141–152. https://doi.org/10.1002/rse2.130

Cañas-Gutiérrez, G.P., Sepulveda-Ortega, S., López-Hernández, F., Navas-Arboleda, A.A. and Cortés, A.J., 2022. Inheritance of yield components and morphological traits in avocado cv. Hass from "Criollo" "Elite Trees" via half-sib seedling rootstocks. *Frontiers in Plant Science*, 13, pp. 843099 https://doi.org/10.3389/fpls.2022.843099

Chen, H., Ashworth, V.E.T.M., Xu, S. and Clegg, M.T., 2007. Quantitative genetic analysis of growth rate in avocado. *Journal of the American Society for Horticultural Science*, 132, pp. 691–696. https://doi.org/10.21273/jashs.132.5.691

Ehleringer, J., Björkman, O. and Mooney, H.A., 1976. Leaf pubescence: effects on absorptance and photosynthesis in a desert shrub. *Science*, 192, pp. 376–377. https://doi.org/10.1126/science.192.4237.376

Espinosa-Alonso, L.G., Paredes-López, O., Valdez-Morales, M. and Oomah, B.D., 2017. Avocado oil characteristics of Mexican creole genotypes. *European Journal of Lipid Science and Technology*, 119, pp. 1600406. https://doi.org/https://doi.org/10.1002/ejlt.201600406

Galindo-Tovar, M. and Arzate-Fernández, A.,

2010. Consideraciones sobre el origen y primera dispersión del aguacate Lauraceae. *Cuadernos de Biodiversidad*, 33, pp. 11–15. Retrieved November 03, 2022. https://www.uv.mx/personal/megalindo/files/2010/07/Origen.pdf

Galindo-Tovar, M.E., Lee-Espinoza, H.E., Murgía-González, J., Leyva-Ovalle, O.R. and Landero-Torres, I., 2013. Domesticación y distribución geográfica de *Persea americana* Mill. en la época precolombina. *Revista de Geografía Agrícola*, 50–51, pp. 65–70. Retrieved October 03, 2022. https://www.redalyc.org/articulo.oa?id=75749283006

Galindo-Tovar, M.E., Ogata-Aguilar, N. and Arzate-Fernández, A., 2008. Some aspects of avocado (*Persea americana* Mill.) diversity and domestication in Mesoamerica. *Genetic Resources and Crop Evolution*, 55, pp. 441–450. https://doi.org/10.1007/s10722-007-9250-5

Gross-German, E. and Viruel, M.A., 2013. Molecular characterization of avocado germplasm with a new set of SSR and EST-SSR markers: Genetic diversity, population structure, and identification of race-specific markers in a group of cultivated genotypes. *Tree Genetics and Genomes*, 9, pp. 539–555. https://doi.org/10.1007/s11295-012-0577-5

Grüter, R., Trachsel, T., Laube, P. and Jaisli, I., 2022. Expected global suitability of coffee, cashew and avocado due to climate change. *PLoS ONE*, 17, pp. e0261976. https://doi.org/10.1371/journal.pone.0261976

Guzmán, L.F., Machida-Hirano, R., Borrayo, E., Cortés-Cruz, M., Espíndola-Barquera, M. del C. and Heredia García, E., 2017. Genetic structure and selection of a core collection for long term conservation of avocado in Mexico. *Frontiers in Plant Science*, 8, pp. 1–10. https://doi.org/10.3389/fpls.2017.00243

Henao-Rojas, J.C., Lopez, J.H., Osorio, N.W. and Ramírez-Gil, J.G., 2019. Fruit quality in Hass avocado and its relationships with different growing areas under tropical zones. *Revista Ceres*, 66, pp. 341–350. https://doi.org/10.1590/0034-737x201966050003

Hidalgo, R., 2003. Variabilidad genética y caracterización de especies vegetales. In T. L. Franco and R. Hidalgo (Eds.), Análisis estadístico de datos de caracterización morfológica de recursos fitogenéticos (pp. 2-27). Instituto Internacional de Recursos Fitogenéticos (IPGRI).

Husson, F., Josse, J. and Pages, J., 2010. Principal component methods - hierarchical clustering - partitional clustering: why would we need to choose for visualizing data? *Applied Mathematics Department*, pp. 1–17. Retrieved October 03, 2022. https://www.semanticscholar.org/paper/Principal-component-methods-hierarchical-clustering-Josse/04335d99d840ac3370f5aeb262828cf127d3ff1c

IPGRI., 1995. Descriptor for avocado (*Persea* spp.). Institute for Plant Genetic Resources. Retrieved October 03, 2022. http://avocadosource.com/papers/misc/ipgri_1995.pdf

Janick, J., 2005. The origins of fruits, fruit growing, and fruit breeding. *Plant Breeding Reviews*, pp. 255–321. https://doi.org/10.1002/9780470650301.ch8

Juma, I., Nyomora, A., Hovmalm, H.P., Fatih, M., Geleta, M., Carlsson, A.S. and Ortiz, R.O., 2020. Characterization of Tanzanian avocado using morphological traits. *Diversity*, 12, pp. 64. https://doi.org/10.3390/d12020064

Kassambara, A. and Mundt, F., 2020. factoextra: Extract and visualize the results of multivariate data analyses. R package version 1.0.7, https://cran.r-project.org/web/packages/factoextra/index.html

Kenkel, N.C., 2006. On selecting an appropriate multivariate analysis. *Canadian Journal of Plant Science*, 86, pp. 663–676. https://doi.org/10.4141/p05-164

Konrad, W., Burkhardt, J., Ebner, M. and Roth-Nebelsick, A., 2015. Leaf pubescence as a possibility to increase water use efficiency by promoting condensation. *Ecohydrology*, 8, pp. 480–492. https://doi.org/10.1002/eco.1518

Landon, A.J., 2009. Domestication and significance of *Persea americana*, the avocado, in Mesoamerica. Nebraska *Anthropologist*, 24, pp. 62–79. Retrieved October 02, 2022. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1046&context=nebanthro

Lê, S., Josse, J. and Husson, F., 2008. FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, 25, pp. 1-18. https://doi.org/10.18637/jss.v025.i01

Ledesma, N. and Carter, D., 2020. The avocados: Guatemalan food. *Acta Horticulturae*, 1299, pp. 169–172. https://doi.org/10.17660/ActaHortic.2020.1299.25

Letunic, I. and Bork, P., 2019. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research*, 47, pp. 256–259. https://doi.org/10.1093/nar/gkz239

López-Galé, Y., Murcia-Riaño, N., Romero-Barrera, Y. and Martínez, M.F., 2022. Morphological characterization of seed-donor Creole avocado trees from three areas in Colombia. *Revista Chapingo Serie Horticultura*, 28, pp. 93–108. https://doi.org/10.5154/r.rchsh.2021.06.010

López-Guzmán, G., Medina-Torres, R., Guillén-Andrade, H., Ramírez-Guerrero, L.G., Juárez-López, P. and Ruelas-Hernández, P., 2015. Caracterización morfológica en genotipos nativos de aguacate (*Persea americana* Mill.) de clima tropical en Nayarit, México. *Revista Mexicana de Ciencias Agrícolas*, 11, pp. 2157–2163. https://doi.org/10.29312/remexca.v0i11.785

Lopez-Lalinde, L. and Maierhofer, C., 2021. Creating a culture of shared responsibility for climate action in Guatemala through education. In F.M. Reimers (Ed.), Education and climate change: The role of universities (pp. 85–112). Springer International Publishing. https://doi.org/10.1007/978-3-030-57927-2_3

Mokria, M., Gebrekirstos, A., Said, H., Hadgu, K., Hagazi, N., Dubale, W. and Bräuning, A., 2022. Fruit weight and yield estimation models for five avocado cultivars in Ethiopia. *Environmental Research Communications*, 4, pp. 075013. https://doi.org/10.1088/2515-7620/ac81a4

Morineau, A., 1984. Note sur la caractérisation statistique d'une classe et les valeurs tests. *Bulletin Technique du Centre de Statistique et d'Informatique Appliqués* 2, pp. 20-27.

Nkansah, G.O., Ofosu-Budu, K.G. and Ayarna, A.W., 2013. Genetic diversity among local and introduced avocado germplasm based on morpho-agronomic traits. *International Journal of Plant Breeding and Genetics*, 7, pp. 76–91. https://doi.org/10.3923/ijpbg.2013.76.91

Pages, J., 2004. Analyse factorielle de données mixtes. *Revue de Statistique Appliquée*, 52, pp. 93–111. Retrieved October 20, 2022. http://www.numdam.org/article/RSA_2004__52_4_93_0.pdf

Patil, I., 2021. Visualizations with statistical details: The "ggstatsplot" approach. *Journal of Open Source Software,* 6, pp. 3167. https://doi.org/10.21105/joss.03167

Pereira, M.E.C., Tieman, D.M., Sargent, S.A., Klee, H.J. and Huber, D.J., 2013. Volatile profiles of ripening West Indian and Guatemalan-West Indian avocado cultivars as affected by aqueous 1-methylcyclopropene. *Postharvest Biology and Technology*, 80, pp. 37–46. https://doi.org/https://doi.org/10.1016/j.postharvbio.2013.01.011

Pino, J.A., Marbot, R. and Martí, M.P., 2006. Leaf oil of *Persea americana* Mill. var. *drymifolia* cv. Duke grown in Cuba. *Journal of Essential Oil Research*, 18, pp. 440–442. https://doi.org/10.1080/10412905.2006.9699135

Popenoe, W., 1935. Origin of the cultivated races of avocados. *California Avocado Association Yearbook*, 20, pp. 184–194. Retrieved October 15, 2022. http://www.avocadosource.com/CAS_Yearbooks/CAS_20_1935/CAS_1935_PG_184-194.pdf

Popenoe, W., 1974. Manual of tropical and subtropical fruits. Hafner Press: New York, NY, USA.

R Core Team., 2022. R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ranjitha, V., Chaitanya, H., Ravi, C., Shivakumar, B. and Naveen, N., 2021. Morphological characterization of avocado (*Persea americana* Mill.) accessions explored from hill zone taluks of Chikkamagaluru district, Karnataka state. *Journal of Pharmacognosy and Phytochemistry*, 10, pp. 310–318. Retrieved October 15, 2022. https://www.phytojournal.com/archives?year=2021&vol=10&issue=2&ArticleId=13825

Renner, S.S., 1999. Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data. *American Journal of Botany*, 86, pp. 1301–1315. Retrieved October 11, 2022. https://pubmed.ncbi.nlm.nih.gov/10487818/

Renner, S.S., 2004. Variation in diversity among Laurales, early cretaceous to present. *Biologiske Skrifter*, 55, pp. 441–458. Retrieved September 10, 2022. https://proyectoorue.wordpress.com/2009/04/02/la-investigacion-basica-y-la-investigacion-aplicada/

Rincón-Hernández, C.A., Sánchez-Pérez, J.D.L.L. and Espinosa-García, F.J., 2011. Caracterización química foliar de los árboles de aguacate criollo (*Persea americana* var. *drymifolia*) en los bancos de germoplasma de Michoacán, México. *Revista Mexicana de Biodiversidad*, 82, pp. 395–412. https://doi.org/10.22201/ib.20078706e.2011.2.474

Ruiz-Chután, J.A., Berdúo-Sandoval, J.E., Kalousová, M., Fernández, E., Žiarovská, J., Sánchéz-Pérez, A. and Lojka, B., 2022. SSRs markers reveal high genetic diversity and limited differentiation among populations of native Guatemalan avocado. *Journal of Microbiology, Biotechnology and Food Sciences*, 12, pp. e6134. https://doi.org/10.55251/jmbfs.6134

Ruiz-Chután, J.A., Berdúo-Sandoval, J.E., Kalousova, M., Lojka, B., Fernández, E., Žiarovská, J., and Sánchez-Pérez, A., 2020. Diversidad genética de materiales nativos de aguacate guatemalteco a través del marcador molecular AFLP. *Ciencia, Tecnología y Salud*, 7, pp. 155–169. https://doi.org/10.36829/63cts.v7i2.746

Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A. and Crowley, J., 2021. Ggally: Extension to 'ggplot2' R package version 2.1.2. https://CRAN.R-project.org/package=GGally

Shewfelt, R.L., 2000. Fruit and vegetable quality. In R.L. Shewfelt and B. Bruckner (Eds.), Fruit and vegetable quality: an integrated view; (pp. 160–173). CRC Press.

Smith, C., 1966. Archeological evidence for selection in avocado. *Economic Botany*, 20, pp. 169–175. https://doi.org/10.1007/bf02904012

Storey, W., Bergh, B. and Zentmyer, G., 1986. The origin, indigenous range and dissemination of the avocado. *California Avocado Society*, 70, pp. 127–133. Retrieved September 10, 2022. http://www.avocadosource.com/CAS_Yearbooks/CAS_70_1986/CAS_1986_PG_127-133.pdf

Subirana, I., Sanz, H. and Vila, J., 2014. Building bivariate tables: The compareGroups package for R. *Journal of Statistical Software*, 57, pp. 1–16. https://doi.org/10.18637/jss.v057.i12

Supplementary material

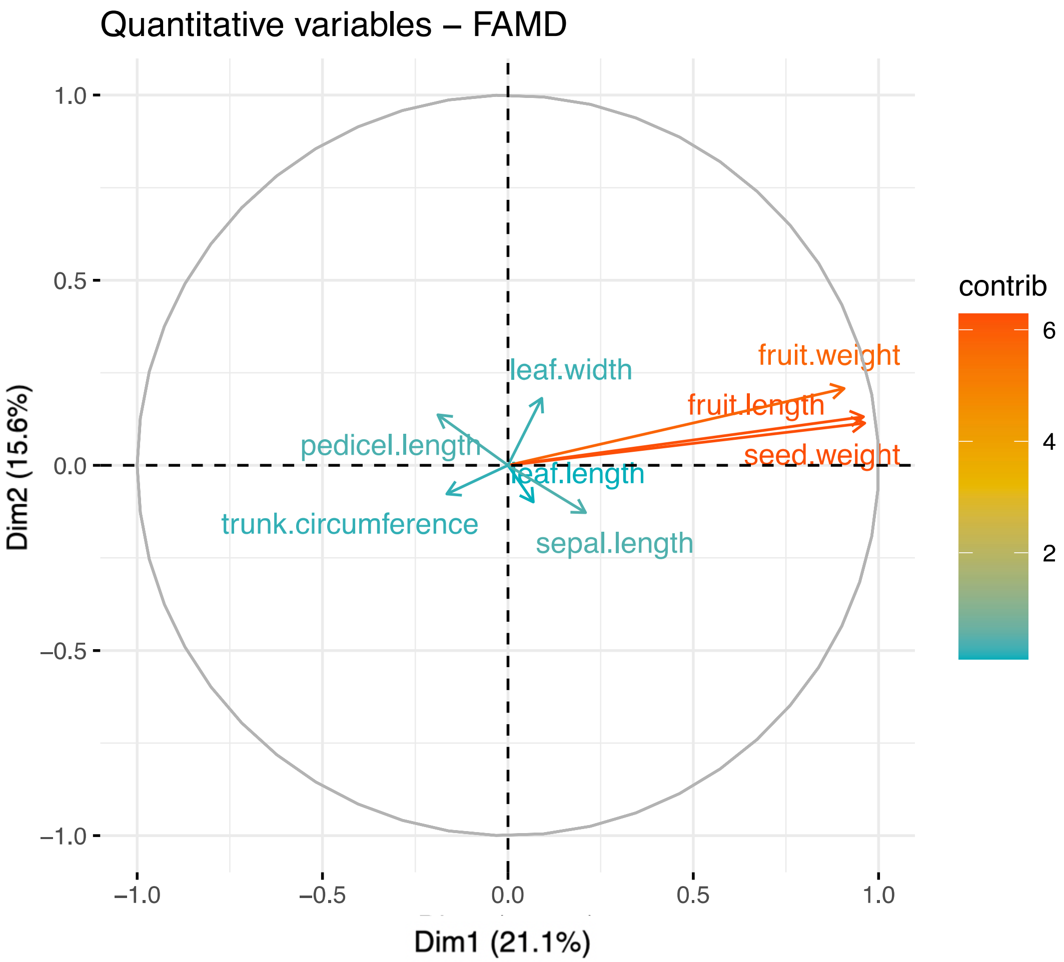

Figure S1. Correlation between quantitative variables from the FAMD analysis

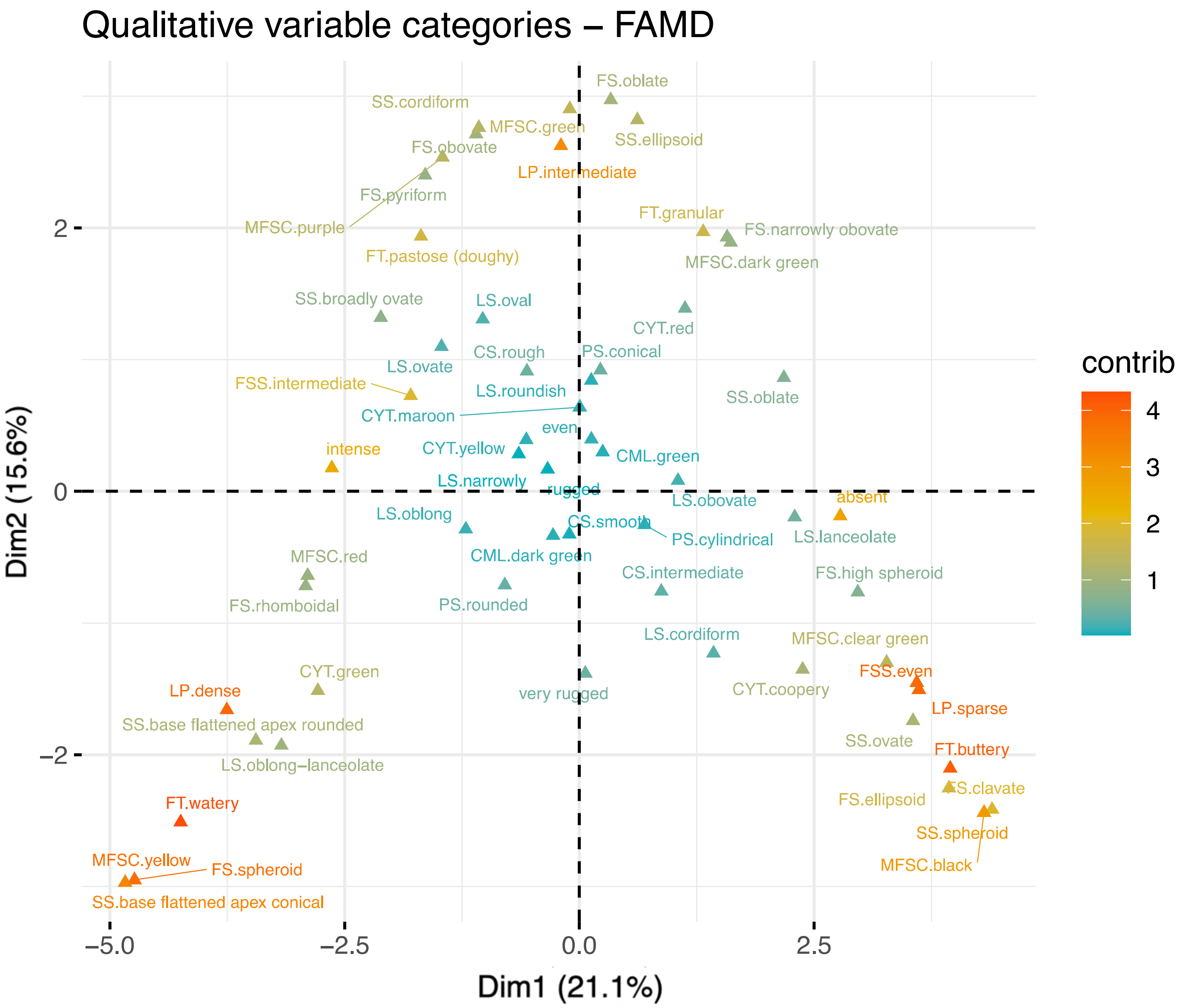

Figure S2. Correlation between qualitative variables from the FAMD analysis

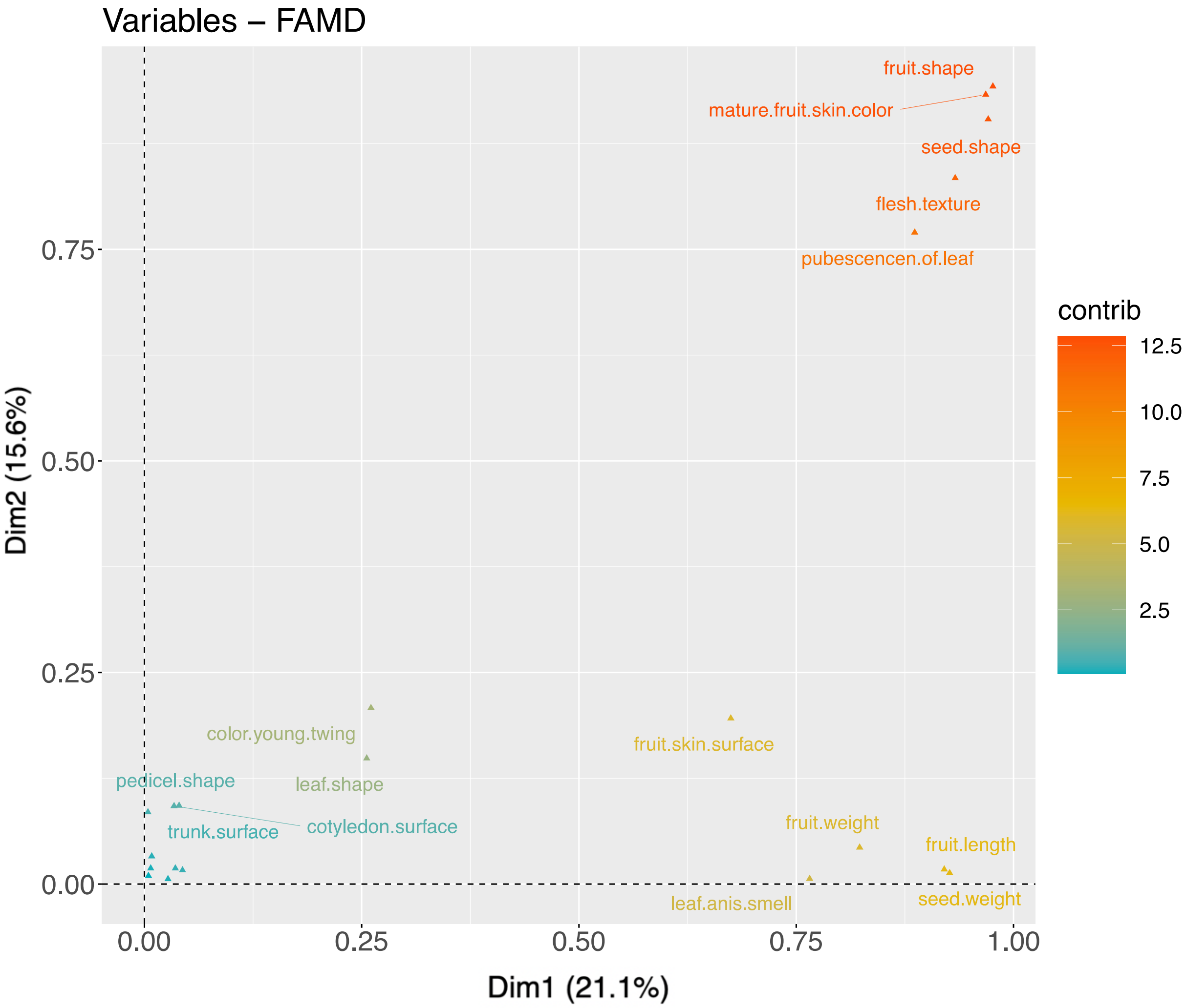

Figure S3. Association between all quantitative and qualitative variables from the FAMD analysis.