Chapter

Landscape Genetics and Phytogeography of Criollo Avocadoes *Persea americana* from Northeast Colombia

Clara Inés Saldamando-Benjumea, Gloria Patricia Cañas-Gutiérrez, Jorge Muñoz and Rafael Arango Isaza

Abstract

Avocado (*Persea americana*) Mill represents one of the most consumed fruits around the world. This species has been differentiated into three main races Guatemalan, Mexican and West Indian according to several molecular markers. However, the interaction between genotypic and phenotypic traits of this crop is still unknown. For this reason, a landscape genetics analysis was made in 90 criollo trees from Northeast Colombia (Antioquia) with 14 microsatellites, sequencing of 3 nuclear loci, endo-1-4-D-glucanase (Cell), Chalcone synthase (CHS) and serine-threonine-kinase (STK) and 28 morphological traits. High genetic diversity was found suggesting a hybrid origin of criollo trees. Morphological variation showed intermixed racial features. FST = 0.03, p =0.001 (measured with microsatellites) suggested low genetic differentiation. According to STRUCTURE, K = 2 for both microsatellites and concatenated nuclear sequences. Criollo trees were assigned together with the Guatemalan and Mexican races. Pearson correlation was significant between expected heterozygocities and elevation. Mantel test was low (r2 = 0.0097, p = 0.015) but significant demonstrating isolation by distance. Grafting is suitable between criollo trees and Hass variety is possible since both avocados are produced within the same altitudes.

Keywords: avocado, races, grafting, morphology, microsatellites, nuclear genes

1. Introduction

Avocado (*Persea americana* Mill.) is one of the most important subtropical crops of the Lauraceae family. The species is native to central Mexico [1–3] and belongs to the sub-genus *Persea* with two other species: *P. schiedeana* (Nees) and *P. parviflora* (Williams) [4]. *P. americana* is an evergreen tree that is heterogeneously branched from 40 to 80 ft. (12.9 to 24. 4 m) tall with elliptic leaves that are 3 to 10 inches (7.62 to 25.4 cm) long and has a dichogamous breeding system [5, 6]. First avocado movements from central Mexico occurred through big mammals (sloths and mammals of the family Gomphotheriidae) migrations to Mesoamerica [2]. Later on, *P. americana* was cultivated and domesticated by the first Mesoamerican cultures (The Mokayas) who transmitted their cultural practices to further civilizations such as the Mayas

and Olmecs [7]. These two civilizations further originated three main avocado races recognized as Mexican [*P. americana* var. drymifolia (Schltdl. and Cham.) S.F.Blake], Guatemalan [*P. americana* var. guatemalensis L. Wms.], and West Indian [*P. americana* var. americana Mill.]. These races are distinguishable at morphological, physiological, and horticultural levels [8, 7] and adapted to different conditions. However, the first two are adapted to cooler (Mediterranean and subtropical) climates and medium elevations, while the third requires warmer (tropical) conditions and lowland humid tropics [2–11]. In Colombia, most commercial cultivars are interracial hybrids developed from chance seedlings [12] where propagation originated native or criollo trees and selected cultivars that are asexually reproduced and vary in flavor and nutritive traits [13] with unknown genetic origin.

Several studies based on molecular characterizations have been made in avocado, mainly with microsatellites [1–16]. They are advantageous since they are co-dominant, bi-parentally inherited, and easily standardized [17]. Additionally, sequencing of nuclear genes endo-1-4-D-glucanase (Cell), Chalcone synthase (CHS) and serine–threonine-kinase (STK) have also been made as they have been useful for studying the genetic origin of avocados [7, 18]. DNA sequencing represents the best method to infer gene genealogies, as historical events (coalescence) are better followed through chloroplast and nuclear DNA in plants and mitochondrial DNA in animals [17, 19]. On the other hand, landscape genetics provides relevant information on the interaction between landscape features and microevolutionary processes, such as gene flow, genetic drift and is a very useful tool for resolving genetic differentiation within a species across different geographical scales at fine taxonomic levels [20].

The chapter aimed to carry out a landscape genetic analysis together with a phylogeographic study on 90 criollo trees sampled in Antioquia (Northeast Colombia) based on molecular (microsatellites and nuclear genes sequencing) and morphological characterizations. Results obtained here are relevant for criollo avocado certification and for future grafting purposes of the species in this department of Colombia where importation of this fruit is highly demanded, particularly the variety Hass.

2. Study of landscape genetics and phytogeography of criollo avocadoes (*P. americana*) from Northeast Colombia

2.1 Materials and methods

2.1.1 Plant material and genomic DNA extraction

Sampling was performed in the Antioquia department (Northern Colombia) during 2008 and 2009. Young leaves were collected from 90 criollo avocado trees. Trees were chosen according to differences in Holdrige life-zones and climatic conditions. Collections included two municipalities from the western sub-region páramo [Sonsón (N = 17) and Abejorral (N = 9) at 2.475 and 2.275 m.a.s.l. respectively], five municipalities from the Western sub-region altiplano between 2.080 and 2.200 m.a.s.l. [El Retiro (N = 5), Marinilla (N = 6), Rio negro (N = 8), La Ceja (N = 6) and San Vicente (N = 6)], and 4 municipalities from the Southwestern sub-region [Caramanta (N = 13), Santa Bárbara (N = 2), Valparaiso (N = 3) and Montebello (N = 12) with elevations ranging from 1.375 in Valparaiso to 2.350 in Montebello]. Total genomic DNA was obtained from leaves based on Cañas-Gutierrez et al. (2015) from a study on an AFLP (Amplified fragment length polymorphism) characterization of the species made in 111 avocado trees.

2.1.2 Microsatellites analysis

Fourteen microsatellites were selected from those used by Alcaraz and Hormaza (2007) based on their high polymorphism. PCR amplifications were performed in 15 μL vol containing 16 mM (NH4)2SO4, 67 mM Tris–HCl pH 8.8, 0.01% Tween20, 2 mM $MgCl_2$, 0.1 mM each dNTP, 0.4 μM of each primer, 25 ng genomic DNA and 0.5 units of BioTaqTM DNA polymerase (Bioline, London, UK). Forward primers were labeled with WellRed fluorescent dyes on the 5′ end (Proligo, France). Reactions were performed in an I-cycler (Bio-Rad Laboratories, Hercules, CA, USA) thermocycler using the following temperature profile: an initial step of 1 min at 94°C, 35 cycles of 30 s at 94°C, 30 s at 50°C and 1 min at 72°C, and a final step of 5 min at 72°C. The PCR products were analyzed by capillary electrophoresis in a CEQTM 8000 capillary DNA analysis system (Beckman Coulter, Fullerton, CA, USA). Each reaction was repeated twice to minimize run-to-run variation.

2.1.3 Sequencing analysis

Amplification primers for the loci Cellulase (endo-1,4-D-Glucanase) (Cell), Chalcone synthase (CHS) and Serine–threonine-kinase (STK) were designed according to avocado sequences reported in the GenBank (Chen et al. 2008, 2009), using Primer 3 Plus software (http://primer3plus.com/). Amplification reactions contained 1 × buffer (20 mM Tris–HCl, pH 8.4, 50 mM KCl), 1.5 mM of magnesium chloride, 0.2 mM of each dNTP (Thermo Scientific, Waltham, MA), 2 μM of each primer, 1.5 U of Taq DNA polymerase (Thermo Scientific), 50 ng of DNA and water to complete a final reaction volume of 25 μL. Temperature cycles were made in a T100™ Thermal Cycler (Bio-Rad, Hercules, CA) with an initial temperature of 95°C preheating for 3 min, then 95°C for 45 s, annealing temperatures of 57°C for Cell primers and 58° for CHS and STK primers for 1 min, extension temperature was 72°C. Amplification included 35 cycles with a final extension of 72°C for 10 min. The amplification products (3 μl) were separated by electrophoresis in 1.0% agarose gels with GelRedTM and visualized in a UV transilluminator.

Polymerase chain reaction (PCR) conditions were as follows: 95°C preheating for 3 min, then 95°C for 45 s, annealing temperatures of 57°C for Cell primers and 58°C for CHS and STK primers for 1 min, extension temperature was 72°C. Amplification included 35 cycles with a final extension of 72°C for 10 min. Reactions were made in an I-cycler (Bio-Rad Laboratories, Hercules, CA, USA) thermocycler [7, 18]. Allele-specific PCR (AS-PCR) was not employed here [7, 18]. Instead, presence of heterozygous and homozygous genotypes was performed from individual chromatograms following the procedure given by 21 in Rhizophora genus (mangrove).

2.1.4 Morphological descriptors

Morphological characterization of 90 criollo trees was made according to 28 morphological descriptors that are specific for avocado [21]. These traits included: leaves measurements of length and width, leaves pubescence, branch and leaves colors, leaves margins, the number of primary and secondary veins found in leaves, leaves shape, branch insertions, trunk surface, and tree shape, amongst others.

2.1.5 Landscape variables

Elevation (altitude) was considered in this study according to the altitude above the sea level measured in meters. For the landscape analysis both topography and georeferentiation measurements were also considered as they were taken on each

sampling site. Topographic analysis was associated to elevation profiles of the surface land, following FAO (1990) protocol as follows: 1) flat: elevation from 0 to 0.5%, 2) almost flat: elevation from 0.6 to 2.9%, 3) little wavy: elevation from 3 to 5.9%, 4) wavy: elevation from 6 to 10.9%, 5) broken: elevation from 11 to 15.9%, 6) hillside: elevation from 16 to 30%, 7) Strong undermined: > 30%, moderate variation of elevations, 8) mountainous: > 30%, large variations of high elevation range (˃300 m), 9) Other: specify other features.

2.1.6 Population genetics analysis

Microsatellite locus (N) diversity estimators were calculated with GENALEX 6.501 [22] for 90 criollo trees. Genepop 4.2 package [23] was used for exploring Hardy–Weinberg equilibrium and linkage disequilibrium per population. AMOVA test was performed to determine the genetic population differentiation amongst 11 sampling sites within Antioquia with the program GENALEX 6.501 [22]. The model-based clustering analysis STRUCTURE 2.3.4 was 1) first applied to 14 microsatellites and then 2) used with 3 concatenated nuclear loci in 90 individuals [24]. These two separated analyzes were carried out 1) to assess the most probable avocado cluster membership from Antioquia and 2) to determine the genetic assignment of criollo trees within the three avocado races according to GenBank data bases for these reasons, each molecular marker was analyzed separately. This program was run for 150.000 Markov chain Monte Carlo steps after a burn-in period of 15.000 interactions from K = 1–15 considering: a) no admixture model and b) model-independent allele frequency for each marker following the suggestions given by Chen et al. (2008) and Chen et al. (2009). Each K was calculated from 15 independent runs and 10 iterations. For DNA sequences, each nucleotide was numerically coded as follows: A = 1, T = 2, C = 3 and G = 4 and missing data as −9. A filter was applied to the three loci such that SNPSs were recognized from the total number of segregating sites (S) detected by the program DNAsp V5 [25]. A similar procedure was performed by Chen et al. (2009). The ad hoc estimated likelihood of K (delta K) [26] was obtained with STRUCTURE HARVESTER [27]. K mean values and their estimators were calculated with CLUMPP [28] and graphs were produced with Distruct 1.1. Finally, the Mantel test was estimated (correlation between genetic distances (from microsatellites) vs. geographic distances of collecting sites) with GENALEX 6.501.

2.1.7 Heterozygote detection in sequences and phylogeographic analysis

For sequencing analysis, DNA amplicons were purified and sequenced in Macrogen Inc., (South Korea with an ABI 3730XL sequencer (Perkin Elmer/Applied Biosystem, Foster City, CA). Sequences (forward and reverse) were edited by hand with Bioedit (Hall 1999). Chromatographs were examined to detect heterozygous (with double peaks in a polymorphic site) vs. homozygous genotypes (with single peaks in a polymorphic site) [29]. Heterozygous segregate into several haplotypes, for example: ATG/CGC/TA segregates into ATGGCAC, ATCGCTAC, ATGGTAC, etc. [30]. This variation was detected by Chen et al. (2008) from sylvester and domesticated avocados with the program POLIPHRED (Ewing and Green 1998). An alternative to this method is the use of chromatographs [29]. In this study, the highest peak for each heterozygous genotype (with two peaks per nucleotide position) in all sequence was selected per individual sample (collected specimen) to produce only one haplotype per avocado (**Figure 1**) and to simplify further population genetic analysis as Chen et al. (2008) and Chen et al. (2009)

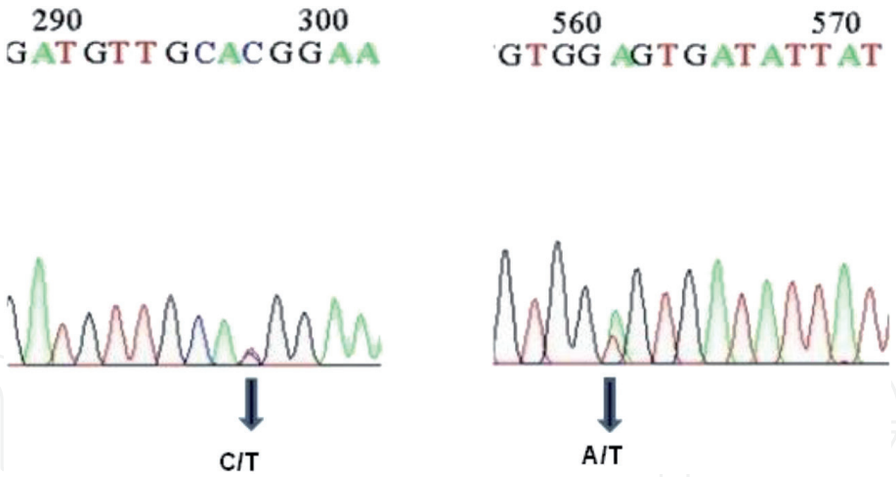


Figure 1.
Heterozygous genotypes were detected in the nuclear loci were considered in this study. Two double peaks at genotype AT.

mostly reported one haplotype per avocado (either domesticated or sylvester) in the GenBank.

Clustal W was used for all alignments in Mega 7.0 (Kumar et al. 2016). The variation in DNA sequence (extension of DNA polymorphism in a DNA sequence) for each locus was measured with the following parameters: 1) nucleotide polymorphism: q = s/a1, a1 = S 1/I and s = a number of polymorphic markers, 2) nucleotide divergence: $\pi = \Sigma$ dij/c, dij = number of nucleotide differences (substitutions) per site between the "ith" and "jth" alleles and c = a total number of sequences studied, 3) segregating sites: (S) where S = a nucleotide site with more than one nucleotide variation in "m" sequences comparisons, 4) a number of haplotypes (H), and 5) haplotype diversity = Hd [19]. All of them were computed on DNAsp V5 [25]. Sequenced loci were considered as independent genes based on the results of the linkage disequilibrium made by Chen et al. (2008). For a phylogenetic Bayesian analysis, jmodeltest [31, 32] was used to determine the model of evolution concatenated data set (Antioquian and Genebank sequences) obtaining the model General Time Reversal (GTR). The platform Beast 2.0 [33] was used to obtain the phylogeny. Data was analyzed in Beauti using an MCMC with 100 million generations. Each consensus tree obtained for the COI gene and ITS region was maintained every 1000 generations and preBurning was established for the first 10 million generations for the MCMC. The posterior probability was corroborated with Tracer v1.6 [34], the consensus trees were summarized with Treeannotator and graphed with Figtree [34]. No out-group was used in this phylogeny as no other reported sequences for Cell, CHS and STK are reported in the GenBank in other species.

2.1.8 Morphological data analysis

The program Past 1.2 [35] was used to carry out statistical analysis on morphological data. This analysis included a PCA (Principal Component Analysis) of 28 morphological descriptors used for the species together with the 14 microsatellites information obtained on each criollo tree. Finally, the average expected heterozygocity obtained for each criollo tree was correlated (Person correlation) to two environmental variables: topography and elevation according to 36 procedures. This analysis was made to determine if genetic diversity in criollo avocado varied in altitude and/or in topography [36].

3. Results

3.1 Population genetics results

Genetic diversity estimated with 14 microsatellites in 90 criollo trees showed that the total number of amplified alleles ranged from 248 to 28. The mean allele number was from 5.7 to 2, the number of effective alleles varied from 3.33 to 1.84. The Shannon Weber index ranged from 1.34 to 0.52. Maximum and minimum estimators obtained here were detected at the municipalities of Sonsón and Santa Bárbara. Observed heterozygocities were between 0.48 and 0.107 and expected heterozygocities were from 0.66 to 0.38. All microsatellites were in Hardy – Weinberg equilibria (**Table 1**). Linkage disequilibria were not significant after Bonferroni corrections (data not shown). Analysis of molecular variance (AMOVA) with 14 microsatellites generated FST = 0.05439 ($p < 0.0001$) suggesting population structure amongst 11 municipalities where avocado samples were taken. This test also showed that most genetic variation was within (94.5%) than between populations (5.44%). STRUCTURE HARVESTER estimated K = 2 according to mean Ln (P|D) = −1445.81;

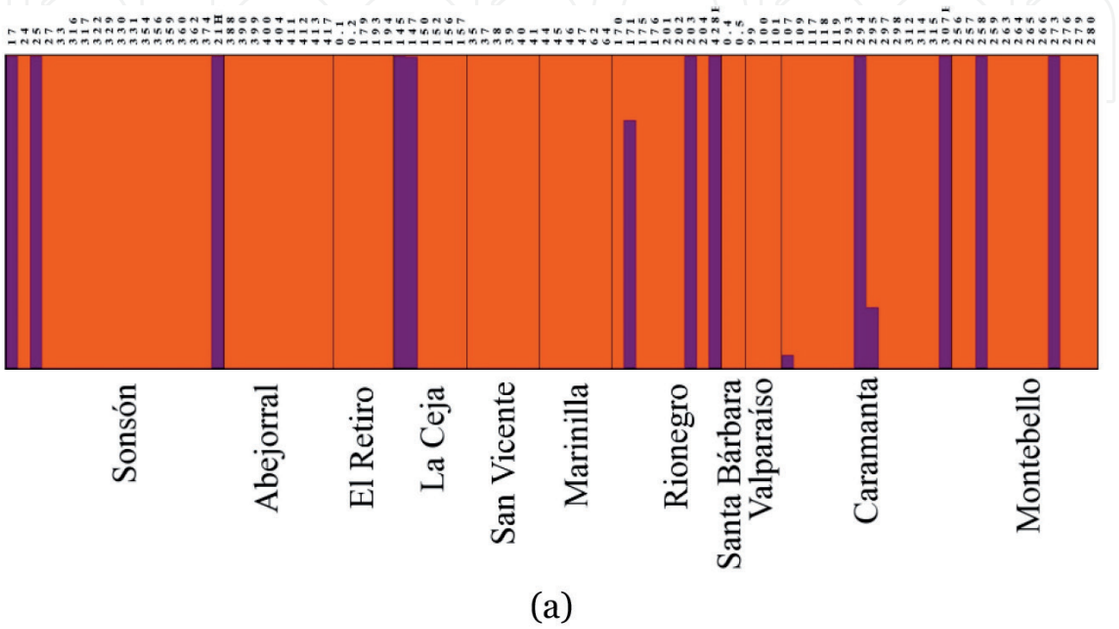
Collecting site	Ni	Nt		N	Na	Ne	I	Ho	He	F	HWE (P)
Sonsón	18	248	Mean	17.710	5.780	3.336	1.344	0.486	0.663	0.258	1.000
			SDE	0.160	0.576	0.341	0.099	0.043	0.031	0.061	
Abejorral	9	125	Mean	8.920	4.143	2.820	1.120	0.523	0.603	0.120	1.000
			SDE	0.071	0.533	0.279	0.107	0.051	0.035	0.083	
El Retiro	5	67	Mean	4.780	3.143	2.683	0.984	0.560	0.570	−0.011	0.983
			SDE	0.155	0.275	0.257	0.105	0.092	0.052	0.146	
La Ceja	6	84	Mean	6.000	4.786	3.717	1.373	0.536	0.702	0.221	1.000
			SDE	0.000	0.422	0.345	0.087	0.398	0.025	0.081	
San Vicente	6	82	Mean	5.857	3.857	2.775	1.007	0.398	0.574	0.316	1.000
			SDE	0.097	0.443	0.294	0.122	0.073	0.054	0.095	
Marinilla	6	83	Mean	5.929	3.643	2.626	1.010	0.425	0.548	0.213	1.000
			SDE	0.071	0.341	0.307	0.109	0.081	0.051	0.118	
RioNegro	9	123	Mean	8.786	5.643	4.008	1.485	0.539	0.724	0.233	0.985
			SDE	0.155	0.401	0.340	0.084	0.051	0.026	0.084	
Santa Bárbara	2	28	Mean	2.000	2.000	1.814	0.582	0.393	0.384	−0.030	
			SDE	0.000	0.182	0.157	0.097	0.107	0.061	0.193	0.985
Valparaiso	3	42	Mean	3.000	3.214	2.655	1.021	0.571	0.587	0.028	
			SDE	0.000	0.214	0.223	0.077	0.089	0.035	0.137	1.000
Caramanta	14	191	Mean	13.643	6.429	3.739	1.485	0.435	0.703	0.365	
			SDE	0.133	0.510	0.345	0.086	0.052	0.028	0.078	1.000
Montebello	12	162	Mean	11.571	5.429	3.117	1.305	0.521	0.212		
			SDE	0.137	0.402	0.241	0.088	0.055	0.073		

(Ni = number of individuals sampled per population, Na = number of different alleles, Nt = total number of alleles, Ne = effective number of alleles, N = mean allele number, I = Shanon Index SDE = standard deviation), Ho = observed heterozygocity, He = expected heterozygocity, F = Wright F index, HWE = Hardy Weinberg equilibrium test (P = p value)).

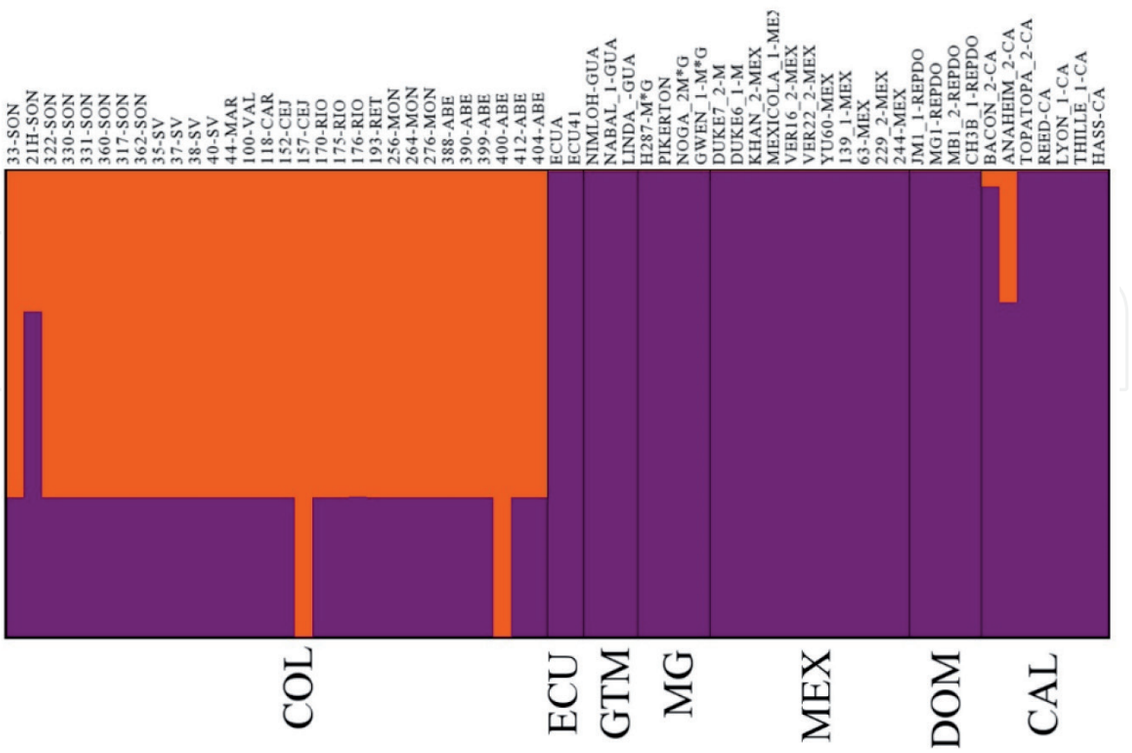
Table 1.
Genetic diversity estimated in 90 criollo avocados.

Stdev = 18.08 and ΔK = 7.79 (**Figure 2a**). This result suggests that samples collected in 11 municipalities were assigned to two sub-populations where most criollo avocados were assigned to cluster 1 (orange) and a few individuals to cluster 2 (purple).

Pearson correlation showed no association between HE (expected heterozygocity) and topography as r = −0.2893 (p = 0.3886) (**Figure 3**) but a significant association between HE and elevation (altitude) as r = 0.7112 (p = 0.014029) (**Figure 4**). Additionally, Mantel test was significant (r = 0.0097, p = 0.015), suggesting isolation by distance (**Figure 5**).



(a)



(b)

Figure 2.
Number of K subpopulations estimated by STRUCTURE HARVESTER with a) 14 microsatellites in 90 criollo trees sampled in Antioquia b) concatenated sequences.

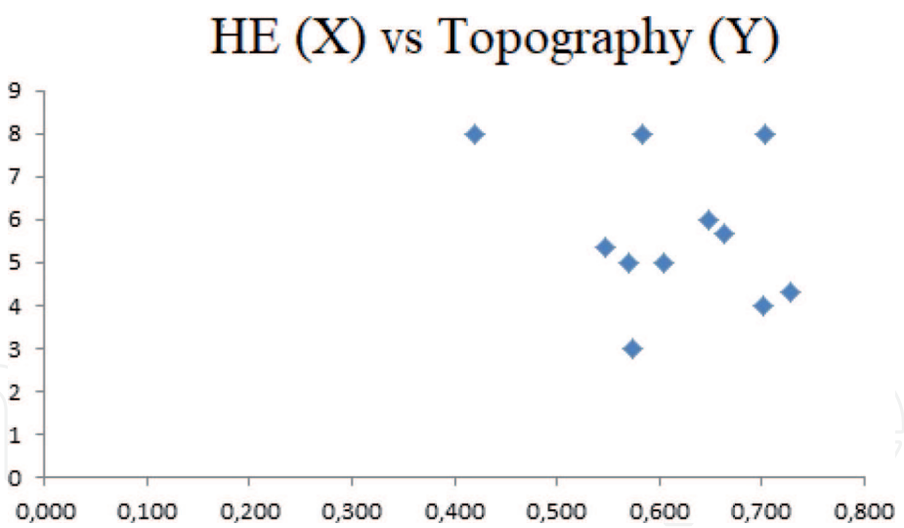


Figure 3.
Correlation between expected heterocigosity and topography.

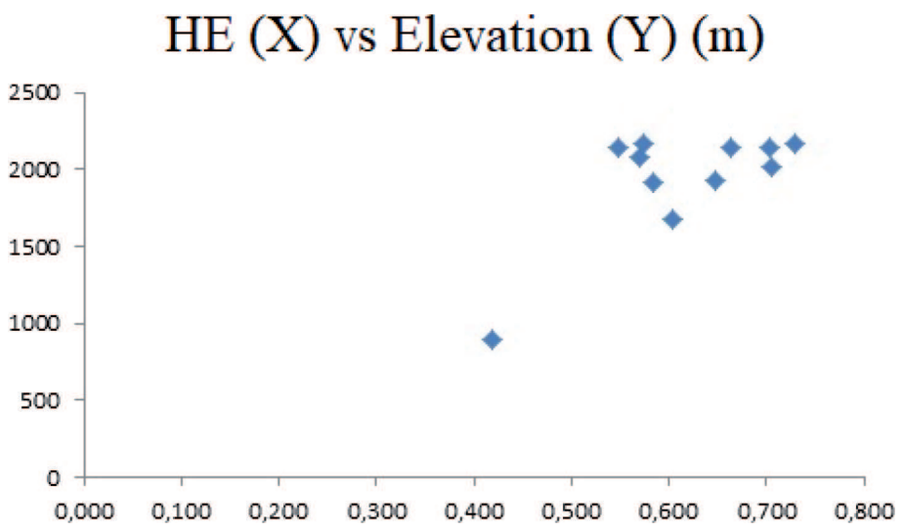


Figure 4.
Correlation between expected heterocigosity and elevation (m).

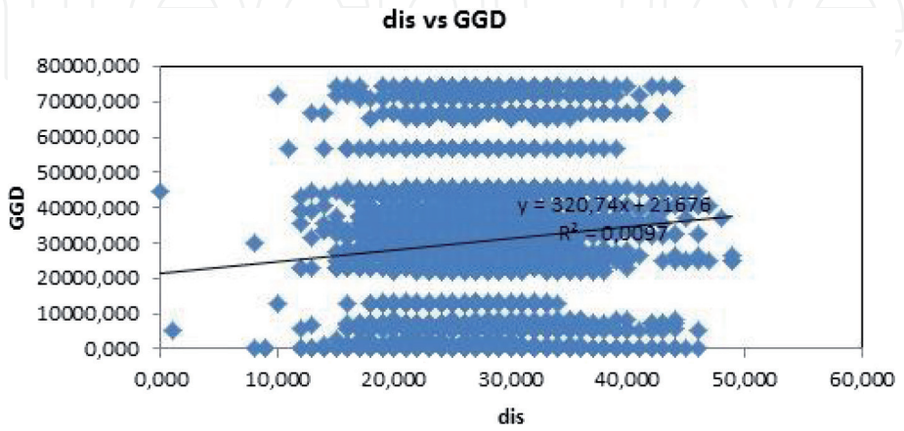


Figure 5.
Mantel test. Geographic vs. genetic distances measured in avocado.

3.2 Phylogeographic results

Amplification of the three nuclear genes was obtained from 87/90 criollo cultivars. Cell gene was sequenced in 49/87 cultivars, CHS gene in 80/87 and STK gene in 57/87. Amplification and final edition for Cell gene produced a final fragment of 997 bp, for the gene CHS of 827 bp and the gene STK of 1170 bp. These fragments were shorter than those obtained by Chen et al. (2008) and Chen et al. (2009). Chromatograph visualizations showed that locus Cell exhibited 10/49 cultivars with heterozygous genotypes in 3 positions: 68 (A/T), 75 (C/T) and 249 (C/G). Allele variation in position 75 was the most frequent. Locus CHS had 31/80 cultivars that were heterozygous in 14 positions: 9 (T/C), 19 (T/C), 58 (T/G), 75 (G/C), 96 (C/T), 121 (G/A), 126 (G/C), 201 (A/T), 214 (T/C), 218 (G/C), 257 (T/C), 258 (A/G), 302 (G/C) and 365 (T/C). The most abundant heterozygous variants were found in positions 121 and 201. Locus STK exhibited 14/57 cultivars with heterozygous genotypes in 25 positions: 17 (C/A), 88 (C/A), 89 (C/A), 104 (G/A), 193 (T/C), 194 (C/T), 209 (C/T), 215 (A/G), 229 (G/A), 297 (G/A), 307 (G/T), 314 (A/G), 379 (C/T), 399 (A/G), 414 (A/T), 418 (T/C), 428 (C/T), 468 (A/G), 472 (A/T), 485 (A/G), 631 (T/C), 648 (C/T) and 737 (A/G). The most frequent heterozygous were 88, 229, 379, and 399.

Genetic diversity obtained in the three loci showed that Cell locus exhibited high values for the criollo trees, and the high values between criollo trees vs. cultivars reported by Chen et al. (2009) from Costa Rica, Dominican Republic, Ecuador and México and the interracial hybrids between Guatemala x Mexico races (**Table 2**). D Tajima test was $D = -2.54$, $p < 0.001$ in Antioquia suggesting purifying selection or population avocado expansion [19]. Locus STK showed the highest gene diversity for all estimators in criollo avocados (even higher than Cell) and also between criollo trees vs. avocado cultivars reported by Chen et al. (2009). However, D Tajma test ($D = 0.237$, $p > 0.01$) was not significant suggesting no population expansion (positive or purifying selection pressure) or reduction (bottleneck effect) [19] in Antioquia. On the contrary, locus CHS presented the lowest gene diversities estimations within Antioquian avocados and between Antioquia vs. other cultivars [18]. Also, D Tajma test ($D = 0.796$, $p > 0.01$) was not significant. Genetic diversities increased when criollo sequences were combined with GenBank avocado sequences as more haplotypes were included in the analysis.

In locus Cell, the haplotype number was 20 (**Tables 2** and **3**). Haplotype 1 was composed by 31 sequences from Antioquia solely [Sonsón (9), Abejorral (5), Montebello (5), Marinilla (3), San Vicente (3), El Retiro (2), Rionegro (1), Santa Bárbara (1) Caramanta (1) and Valparaiso (1)]. Haplotype 2 was integrated by 30 cultivars: 7 from Antioquia [Sonsón (3), La Ceja (2), Marinilla (1) and Caramanta (1)], 12 from Guatemala, 5 to Mexico and 5 were hybrids between Guatemala x Mexico (7, 17). Haplotype 13 was integrated by 13 cultivars: 7 native to Mexico, 2 native to Ecuador and the varieties: Topa-Topa, Khan, Mexicola and Puebla. All of them were classified as Mexican races by Chen et al. (2009) except for Puebla (that is a hybrid between the three avocado races M × G × WI with assignment percentages of 6%, 82%, 12% respectively). Haplotype 16 included 5 cultivars: 4 of them from the Dominican Republic and one haplotype from the variety Arue. This haplotype was identified as the West Indian race [18]. Haplotype 15 was composed by sequences from the varieties Thomas and Duke 6 and the Mexican cultivar 63. All of them were assigned to Mexican races by Chen et al. (2009). Haplotype 14 included the commercial varieties: Zutano, Thille, Gwen, Esther, Bacon and Anaheim. All of them were assigned by Chen et al. (2008) and (2009) as Guatemalan races except

Gen	Aligned length	Population Studied	N	S	H	Θ × 10–3	π × 10–3	D Tajima	P
Cell	1540	Total	51	33	15	4.14	2.93	−0.89	>0.10
		Wild	20	30	12	4.6	3.46	−0.85	>0.10
		Cultivars	31	15	10	2.22	2.24	0.035	>0.10
	997	Total	101	89	20	4.29	4.27	2.52	<0.001
		Antioquia	49	80	11	5.89	5.43	2.54	<0.001
		Genebank	52	21	10	2.8	2.79	1.29	>0.10
CHS	1210	Total	43	35	27	6.09	4.92	−0.61	>0.10
		Wild	16	27	15	5.66	4.36	−0.81	>0.10
		Cultivars	27	29	19	5.67	5.02	−0.38	>0.10
	827	Total	115	26	26	3.6	3.58	−1.300	>0.10
		Antioquia	80	9	13	1.52	1.52	−0.79	>0.10
		Genebank	35	25	15	5.47	5.43	−1.09	>0.10
STK	1398	Total	53	48	22	6.56	5.51	−5.02	<0.001
		Wild	20	43	16	7.23	6.18	−5.51	<0.001
		Cultivars	33	27	11	4.06	4.98	−0.72	>0.10
	1170	Total	113	41	33	7.09	6.42	−0.24	>0.10
		Antioquia	57	19	19	3.68	3.66	−0.032	>0.10
		Genebank	56	31	15	4.87	4.89	−0.66	>0.10

Table 2.
Genetic diversity was estimated from three nuclear loci in criollo avocado trees sampled at the department of Antioquia. N = total number of genotyped individuals, S = segregant sites, Θ = nucleotide polymorphism, π = nucleotide diversity, D = Tajima test, P = p-value.

Haplotype	Cultivar	Haplotype	Cultivar
1	47-MAR –46-MAR –44-MAR –412-ABE –411-ABE –40-SV - 400-ABE –390-ABE –38-SV - 388-ABE - 362-SON -360-SON -35-SV -33-SON - 330-SON -322-SON -317-SON -29-H-SON - 265-MON -264-MON -263-MON -258-MON - 256-MON -24-SON -194-RET –193-RET –17-SON -170-RIO -118-CAR –100-VAL –04-SB	2	Yu60 - Reed- Pinkerton_2- Noga_2 - Nimlioh- -Nabal- Lyon- Linda- LeavenHass- - HX48 -Hass- -H670- -H287- Fuerte_2- Duke7_2 - Daily11- COSTRI - Ch35_2- 65_2 19-R-SON -145-CEJ –12-R-CEJ21-H-SON – MEX - 46_2-MEX - 445-R-MAR –307-H-CAR –25-SON -229_2- MEX - 229_1- MEX-
3	37-SV - 257-MON	4	147-CEJ
5	152-CEJ	6	157-CEJ
7	175-RIO -171-RIO	8	176-RIO
9	276-MON	10	404-ABE
11	428-H-RIO	12	446-F-MAR
13	Ver3-MEX -Ver22-MEX -TopaTopa-QRO1-MEX -Puebla_2- -Mexicola_2- Khan- -ECU - 65_1-MEX -63_1-MEX -46_1-MEX –41-ECU -139-MEX	14	Zutano_1 -Thille_1-Gwen_1- Esther_1-Bacon_1- Anaheim_1
15	Thomas_1 -Duke6_1–63_2_-MEX	16	MG1-REPD - MC1_1-REPD - MB1-REPD JM1_1-REPD Arue_1-SI
17	MC1_2-REPD -JM1_2_-REPD – Ch3B-REPD	18	244-MEX
19	Ver16_2-MEX	20	Ch35_1-MEX

Table 3.
Haplotype list of avocado cultivars produced by locus cell in 107 specimens.

for Zutano that is a hybrid between the Guatemalan x Mexican races. Haplotype 17 was composed of cultivars from the Dominican Republic that were assigned as West Indian race by Chen et al. (2008) and (2009). Finally, the rest of the haplotypes were composed of one or two sequences from Mexico or Antioquia.

For the gene CHS, the number of estimated haplotypes was 26 (**Tables 2** and **4**). Haplotype 5 (H5) included 48 avocado cultivars from Antioquia together with 2

Haplotype	Cultivar	Haplotype	Cultivar
1	CH-MB1-REPD CH-Ch3B-REPD CH-MG1_2-_ REPD CH-JM1_2-REPD	2	CH-244-MEX CH-139-MEX CH-Ver16_2-MEX CH-KHAN-MEX CH-LYON_1-CA CH-BACON_2-CA
3	CH-Yu60_2-MEX	4	CH-Ver22_1-MEX CH-63-MEX CH-ZUTANO_1-CA
5	CH-1-ECU CH-41_1-ECU CH-262-MONT CH-256-MONT CH-193-RET CH-176-RIO CH-170-RIO CH-176-RIO(2) CH-179-RET CH-193-RET(2) CH-194-RET CH-201-RIO CH-202-RIO CH-259-MONT CH-276-MONT CH-279-MONT CH-280-MONT CH-293-CARA CH-331-SON CH-356-SON CH-359-SON CH-445-R-MARI CH-446-F-MARI CH-38-SV CH-45-MARI(2) CH-99-VAL CH-109-CARA CH-145-CEJ CH-156-CEJ CH-263-MONT CH-322-SON CH-362-SON CH-46-MARI CH-100-VAL CH-117-CARA CH-329-SON CH-374-SON CH-21-H-SON(2) CH-47-MARI CH-101-VAL CH-118-CARA CH-150-CEJ CH-390-ABEJ CH-05-SB CH-37-SV CH-44-MARI CH-64-MARI CH-152-CEJ CH-317-SON CH-399-ABEJ(2)	6	CH-2-COSTRI CH-NIMLIOH-GUAT
7	CH-Ch35_1-MEX CH-H287-MEXxGUAT CH-GWEN-MEXxGUAT CH-11-R-CEJ	8	CH-229_2-MEX CH-REED-CA CH-THILLE_1-CA CH-DUKE7-MEX
9	CH-THOMAS-MEX CH-ANHEIM_2-CA	10	CH-PIKERTON-MEXxGUAT CH-NOGA_2-MEXxGUAT CH-HASS_2-CA
11	CH-NABAL-GUAT	12	CH-DUKE6-MEX
13	CH-TOPATOPA-CA	14	CH-MEXICOLA-MEX
15	CH-LINDA-GUAT	16	CH-354-SON CH-354-SON(2) CH-27-SON CH-33-SON CH-40-SV(2) CH-256-MONT(2) CH-360-SON CH-400-ABEJ
17	CH-147-CEJ	18	CH-119-CARA CH-45-MARI CH-40-SV CH-35-SV CH-O4-SB CH-157-CEJ CH-175-RIO CH-388-ABEJ CH-399-ABEJ CH – 404-ABEJ CH-412-ABEJ CH-264-MONT CH-330-SON
19	CH-12-R-CEJ	20	CH-19-SON
21	CH-21-H-SON	22	CH-25-SON
23	CH-29-H-SON CH-147-CEJ(2)	24	CH-204-RIO
25	CH-307-H-CARA	26	CH-258-MONT

Table 4.
Haplotype list of avocado cultivars produced by locus CHS in 115 specimens.

sequences from Ecuador. These last two sequences were assigned to the Mexican race by Chen et al. (2008). H5 Antioquian cultivars were distributed as follows: Sonsón (9), Marinilla (6), Montebello (8), Rionegro (5), Caramanta (4), La Ceja (4), El Retiro (4), Valparaiso (3), Abejorral (2), San Vicente (2) and Santa Bárbara (1). H8 (N = 13) was also composed by several Antioquian samples distributed as follows: Abejorral (4), San Vicente (2), Santa Bárbara (1), Marinilla (1), Sonsón (1), Rionegro (1), La Ceja (1), Caramanta (1) and Montebello (1). H16 was composed by 8 haplotypes from Antioquia and the municipalities: Sonsón (5), San Vicente (1), Abejorral (1) and Montebello (1). Haplotype 2 was composed by 6 cultivars: Ver 16, Khan-CA, Lyon-CA, Bacon-CA, and two Mexican races, all these sequences were assigned to the Mexican race by Chen et al. (2008). H7 was integrated by 4 haplotypes: two from Mexico, the haplotypes H27 and Gwen that are hybrids between Guatemala x Mexico races, and one specimen from La Ceja (Antioquia). H8 was also integrated by 4 haplotypes from Mexico: one Mexican sequence, and the varieties Reed (California), Thille (California) and Duke 7 (Mexico). H3 was composed of 3 haplotypes: 2 Noga sequences (a Hass variety) and one cultivar named Pikerton (that represents a hybrid between a Guatemalan race x a Mexican race). Finally, the other haplotypes found for CHS were composed of one sequence from the varieties: Nabal (Guatemala), Duke 6 (Mexico) and Topa-Topa (California). Topa-Topa was classified as a Mexican race by Chen et al. (2008, 2009).

Amongst three loci, gene STK presented the highest number of haplotypes with N = 33 (**Tables 2** and **5**). This gene was analyzed in 113 cultivars (57 from Antioquia and 56 from the GenBank from native cultivars and commercial avocados). H12 represented most of the haplotypes identified for this locus with 24 cultivars, that included: 12 haplotypes of the Guatemalan race, 7 haplotypes that are hybrids between Guatemala x Mexico, a Mexican race, the cultivar 65 from Mexico according to Chen et al. (2008, 2009) and the Antioquian specimens: 19RSON, 12R-CEJ, 11R-CEJ. H3 was composed by 17 cultivars from Antioquia, distributed as follows: Sonsón (5), Abejorral (3), Montebello (3), Marinilla (1), San Vicente (1), El Retiro (1), Rionegro (1), La Ceja (1) and Valparaiso (1). H2 was integrated by 16 cultivars from Antioquia composed by individuals from: Abejorral (4), Sonsón (3), Rionegro (3), Montebello (2), San Vicente (1), Caramanta (1), El Retiro (1) y La Ceja (1). H24 included 9 cultivars, 4 from the Mexican race (commercial avocados), one specimen that is a hybrid between Guatemala x Mexico races and 4 native specimens collected in Mexico and assigned within the Mexican race [7]. H22 was composed of 8 cultivars: 2 native trees from Mexico, 3 from the Dominican Republic assigned within the West Indian race [18], the varieties Thomas and Khan assigned as Mexican races and a specimen from the variety Arue that was assigned within the West Indian race [18]. H33 was composed of haplotypes of the varieties Nabal, Bacon and Anaheim all assigned to the Guatemalan race by Chen et al. (2009). H23 was integrated by the sylvester specimens MG1 and Ch3B-1 and both are native to the Dominican Republic and cultivar COSTRI, from Costa Rica, all of them were assigned within the West Indian race. H4 was composed by Antioquian cultivars from Caramanta (2) and Sonsón (1). Other haplotypes were composed of one or two haplotypes from Mexico and some commercial varieties of avocado.

STRUCTURE HARVESTER produced K = 2 for the three concatenated loci according to the values: mean Ln P (D) = −1445.81, Stdev = 18.08 and Δ K = 7.79 estimated by the Evanno test implemented in STRUCTURE HARVESTER (**Figure 2b**). The first cluster was composed of most criollo samples. These criollo avocadoes were assigned together with Mexican and Guatemalan races whereas the second cluster was composed of only two specimens, one from Sonsón and the other from La Ceja.

Phylogeny obtained from concatenated sequences diverged in two main clusters, the first cluster was constituted by criollo avocados and the second by GenBank

Haplotype	Cultivar	Haplotype	Cultivar
1	33-SON	2	-404-ABE –399-ABE –388-ABE –360-SON -356-SON -331-SON -293-CARA –276-MONT –256-MONT –202-RIO -201-RIO -193-RET –175-RIO -152-CEJ
3	44-MARI -412-ABE –400-ABE –390-ABE –362-SON -35-SV -354-SON -322-SON -280-MONT –27-SON -279-MONT	4	359-SON - 118-CARA –117-CARA
5	329-SON - 259-MONT	6	445-R-MARI –21H-SON
7	05-SB	8	330-SON -119-CARA
9	37-SV	10	317-SON Zutano_2- Whitsell_1- Thille_1 -Teague_2-Reed -Pinkerton_2- Noga_2- Nimlioh -Nabal_1-
11	38-SV	12	Linda_1 -LHass -HX48 - Hass-H670 - H287-Gwen_1- - Esther- -Duke7_2 -Daily11_1- Andes3-65_2 -19RSON –
13	25-SON	14	157-CEJ
15	176-RIO	16	204-RIO
17	266-MONT	18	428H-RIO 411-ABE –40-SV
19	446F-MARI	20	Yu60
21	Ver3	22	Ver16_2- -Thomas_2- QRO1- MC1-MB1_2-Khan_2-JM1- Arue-SI
23	MG1 – COSTRI - Ch3B_1	24	Noga_1- Mexicola_1-Lyon_1 - Ganter-Fuerte_1 - 63-46-244 - 139_1
25	Ver22_2	26	Ver16_1
27	ECUA - Duke6 - 41_1	27	ECUA - Duke6 - 41_1
28	184_1	29	TopaTopa_2
30	Puebla_1	31	Ver22_1
32	229_2	33	Nabal_2 - Bacon_2 Anaheim_2

Table 5.
Haplotype list of avocado cultivars produced by locus STK in 113 specimens.

accessions (**Figure 6**). The closest varieties to criollo avocadoes were Bacon and Anaheim. These varieties are both commercial and classified as Guatemalan races by Chen et al. (2009).

3.3 Morphological results

Descriptive analysis of morphological data (N = 28) (**Table 6**) showed high Coefficient of Variation (CV) values for most traits demonstrating that morphological variation amongst criollo avocadoes is high. For this reason, these avocadoes were not differentiated amongst the three botanical races suggesting that morphological traits are not useful to distinguish criollo trees unless molecular data are included in the analysis. Since PCA analysis showed that the most important landscape variable was elevation (altitude) as 95% of the variation was explained by the first component and 5% by the second. Concatenation of this variable together with morphological and molecular data (microsatellites) produced three clusters distributed as follows: from 800 to 1900, from 2000 to 2100 and from 2200 to 2300 m. a. s. l. The first cluster was composed of avocadoes collected in the municipalities

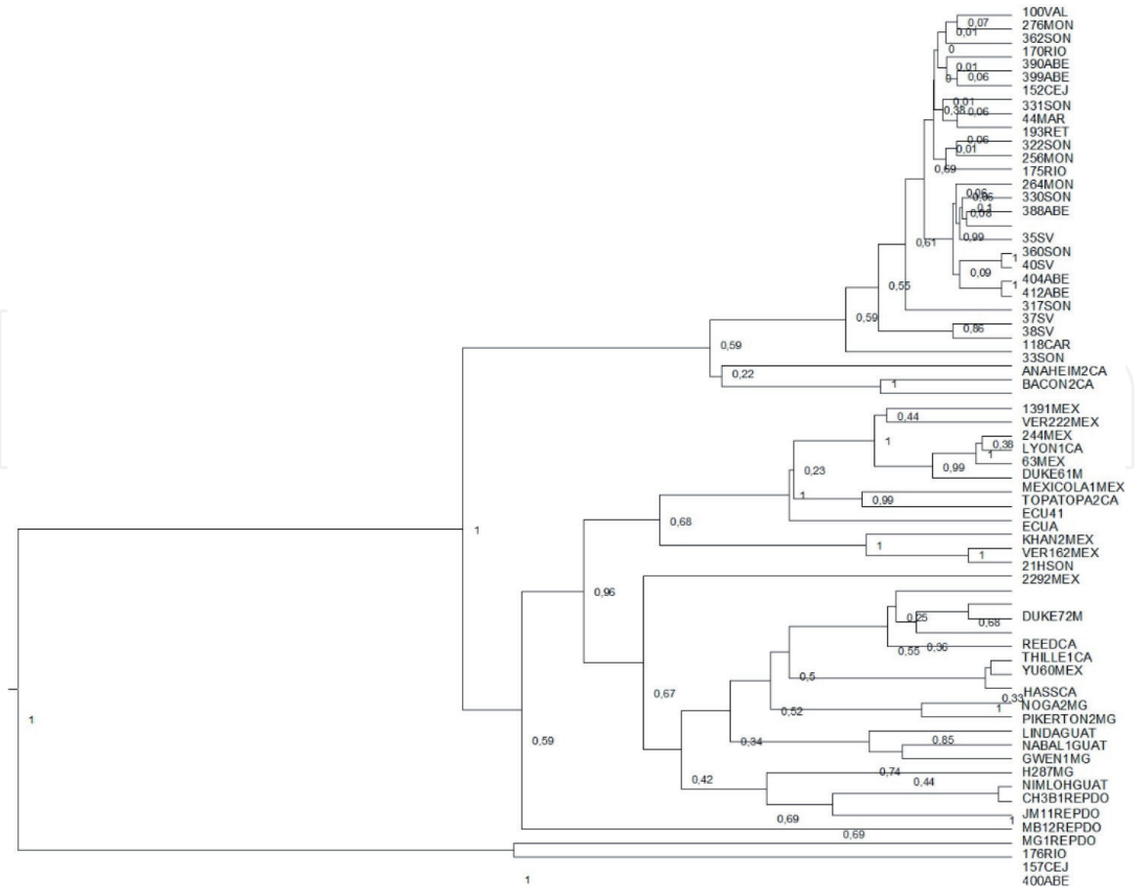


Figure 6.
Phylogeny of avocado cultivars from Colombia and world bank accessions.

Variable	Mean	Minimum	Maximum	SD	CV
Altitude (m)	2015.11	864	2281	242.79	12.05
Tree height (M)	13.05	1.8	25.7	4.66	35.72
Canopy with N-S	8.58	1	20.2	3.96	46.17
Canopy width N, W, E	8.66	1	22	4.18	48.29
Trunk circumference (cm)	117.79	19	234	45.45	38.59
Between stem lengths (cm)	0.62	0.234	1.428	0.25	39.62
Stems diameters (cm)	0.48	0.29	1.04	0.1	21.01
Leave length (cm)	18.58	9.8	30.8	4.12	22.15
Leave width (max) (cm)	8.78	4.7	14.5	2.08	23.64
Number of veins/ leave (Pairs)	7.44	1	16	3.68	49.42
Divergence between veins/leave (grades)	41.52	30	60	6.05	14.57

Table 6.
The descriptive analysis made on continuous traits measures in 90 criollo avocados.

of Abejorral, Montebello and Santa Bárbara found from 800 to 1900. The second avocadoes from Caramanta, Sonsón, El Retiro and Rionegro found from 2.000 to 2.200 and the third group by avocadoes from Valparaíso, La Ceja and Sonsón found from 2.300 to 2.400 m. a. s. l. (**Figure** 7).

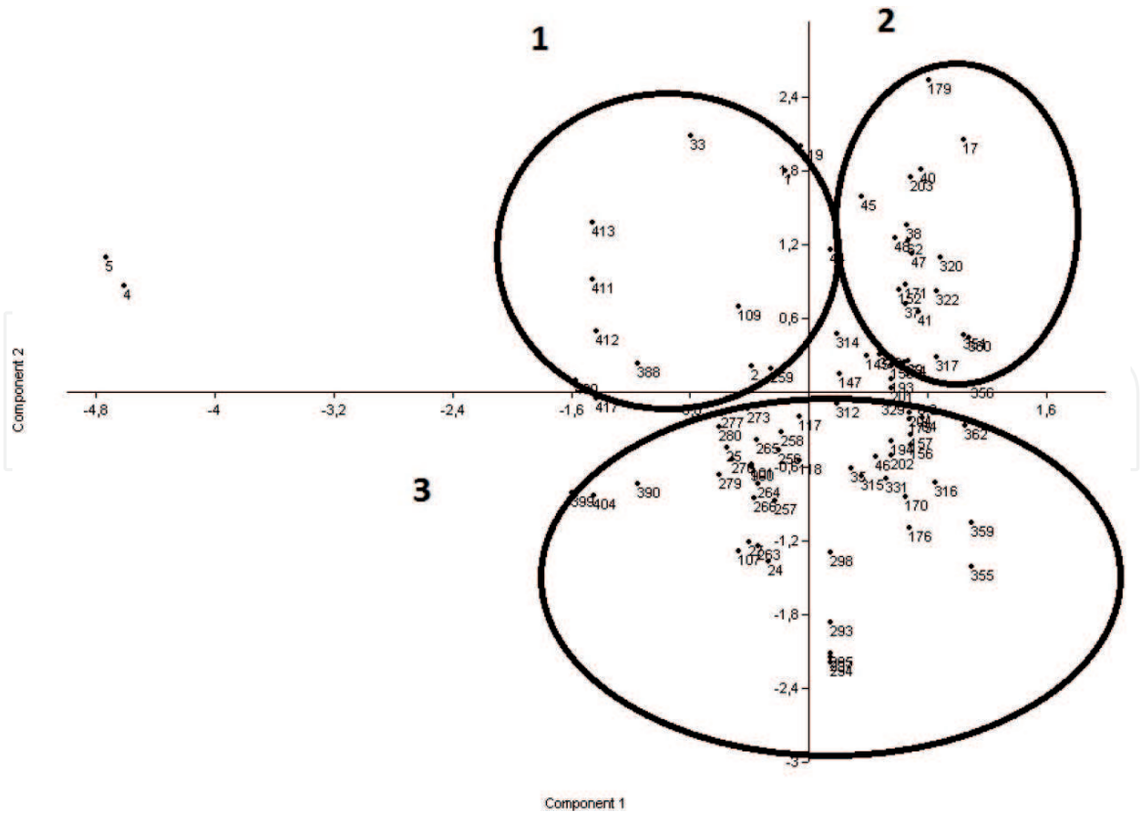


Figure 7.
PCA analysis made in criollo avocados with morphological traits.

4. Discussion

Molecular characterization was made here in 90 criollo trees from Northeast Colombia (Antioquia) by using 14 microsatellites previously standardized by Alcaraz and Hormaza (2007). These criollo trees presented similar genetic diversities to avocado germplasm banks from Spain [14] and US [1]. These germplasms are mainly composed by avocado cultivars from different countries such as Mexico, Guatemala, Israel, US, Chile, Ecuador, Canary Islands, amongst other countries. Also, genetic diversity found within criollo trees was higher than the diversity obtained with 6 microsatellites in avocados from Veracruz (Mexico) by Galindo-Tovar and Milagro (2011). Differences between these two studies could be due to criollo avocados from Antioquia being the product of multiple hybridizations through insect pollinators whereas Mexico samples were mostly composed of avocado cultivars. Cañas-Gutierrez et al. (2015) made a molecular analysis of 111 avocados from Antioquia with 38 AFLP and observed that criollo avocados were molecularly highly polymorphic. They also found that these trees shared AFLP bands with the varieties Fuerte, Hass and Reed demonstrating their hybrid origin.

Genetic population structure was also found amongst 11 Antioquian municipalities, where criollo avocadoes were collected, as FST = 0.05439 ($p < 0.001$). Similar results were obtained by Cañas-Gutiérrez et al. (2015) by using AFLP. The outcome obtained with these molecular markers could be explained due to differences in Holdridge zones amongst 11 municipalities. Avocadoes collected in these 11 sites were clustered in three groups by the PCA analysis. In contrast, only K = 2 groups were estimated with STRUCTURE HARVESTER based on molecular data only.

Person correlations between expected heterozygocities (estimated for 14 microsatellites) vs. elevation (altitude) and expected heterozygocities vs. topography were only significant for the former. This result suggests the importance of Holdridge zones (altitudes) in the genetic differentiation of criollo avocadoes from Colombia. Also, the Mantel test was significant suggesting isolation by distance in criollo Antioquian avocado. This outcome might be explained due to avocado farm producers sharing seedlings between neighbors, increasing genetic similarities between close orchards compared to distant orchards.

Also, a genetic characterization was made in 87 criollo trees based on Sanger DNA sequencing by using three nuclear genes: endo-1,4-D-glucanase (Cell), chalcone synthase (CHS), and serine–threonine-kinase (STK). These loci were previously standardized by Chen et al. (2008) in 21 sylvester avocado cultivars from Mexico, Ecuador, Costa Rica and the Dominican Republic and by Chen et al. (2009) in 33 domesticated cultivars. In these two separated studies, the authors estimated K = 3 avocado clusters with STRUCTURE and assigned their cultivars and domesticated avocadoes to the three botanical races.

Concerning the results related to the heterozygocities found in the criollo avocadoes. These results can be due to *P. americana* being a diploid species with 24 chromosomes (2n = 2x = 24) [37–39], meaning that during cell division the species segregates two alleles per locus producing homozygous and heterozygous genotypes. Chen et al. (2008) and Chen et al. (2009) re-sequenced four nuclear loci: endo-1,4-D-glucanase (Cell), chalcone synthase (CHS), flavanone-3-hydroxylase (F3H), and serine–threonine-kinase (STK) with AS (allele-specific)-PCR procedure. This last method facilitated the identification of heterozygous genotypes from each cultivar. In contrast, in this study, Sanger sequencing was implemented. This method required chromatograph observations to detect heterozygous genotypes [29, 40]. Chen et al. (2015) made a phylogeographic study of five Rhizophora species (mangrove) with five nuclear genes (22,454, 23,056, 22,274, 23,714 and C49) and sequenced four R. apiculate populations with seven nuclear genes (C22, 22,274, 23,056, 23,852, 22, 23,186 and 22,066). They found that the genus Rhizophora diverged into two clades: one composed by R. mangle and R. racemosa and the other by *R. apiculata*, *R. stylosa* and *R. mucronata*. In the case of *R. apiculate*, the presence of double peaks in nuclear genes allowed them to discover hybrids between *R. stylosa* and *R. apiculata* as they represented genotypes that were the product of the crosses between these two mangrove species. Likewise, heterozygous detection was made by Wei et al. (2017) with the nuclear gene PAL (phenylalanine ammonia-lyase) and the species *Camellia flavida*, an endangered population of yellow camellia from southwest China. Wei et al. (2017) used chloroplast and nuclear sequences to improve the taxonomic classification of *C. flavida* and found that the species was differentiated into three groups named: *C. flavida:* var. flavida 1, var. flavida 2, and var. patens. These two studies were useful to implement the same data analysis performed in criollo avocadoes made in this work.

In criollo avocadoes, results of the percentage of heterozygous genotypes showed that 20% of heterozygous were observed in locus Cell, 43% in locus CHS and 49% in locus STK. These outcomes were slightly lower to Chen et al. (2008) as they estimated heterozygocities of 40%, 70.6% and 50% for these three loci respectively from sylvester avocado cultivars and values of 48.4%, 69.2% and 68% from domesticated cultivars. Higher gene diversity estimators (number of haplotypes, Hd, π, Θ) were found in both Cell and CHS loci compared to haplotypes reported in the GenBank, but lower estimators were found for locus STK. These results might be due to differences between samples origins, sample sizes (sequence length) and selection pressures [17]. Also, it is important to mention that the Tajima test was only significant for locus Cell meaning that the gene might be under purifying selection pressure or the population is in expansion in Antioquia.

DNAsp results showed that some criollo haplotypes from Antioquia clustered with the Mexican and Guatemalan sequences (and therefore their races) reported from the three loci by Chen et al. (2008) and (2009). Our results also showed K = 2 for the concatenated sequenced loci with STRUCTURE HARVESTER. The first cluster was integrated by avocado sequences from Antioquia together by sequences reported by Chen et al. (2008) and (2009) from Ecuador, Guatemala, Mexico, California, and crosses between Mexico x Guatemala races and the second cluster by 3 samples from Antioquia that were genetically apart from the rest. In contrast, samples from Dominican Republic sequences deposited in the GenBank did not cluster with any criollo samples, demonstrating that they are not genetically similar to the West Indian race as these avocados were classified within this race by Chen et al. (2008) and (2009). Bayesian phylogeny obtained with concatenated sequences produced two main clades, one was composed of Antioquian avocados and the other by sequences reported previously [7, 18]. This outcome was similar to the results obtained with STRUCTURE HARVESTER as the tree clustered criollo avocados within one group that was closely related to two samples that were previously classified as Guatemalan races.

Concerning morphological data analyzed here, most continuous traits presented large Coefficient of Variation (CV) estimations suggesting that most morphological traits are highly variable, for this reason, they were not useful to classify criollo trees within the three botanical races [41]. According to Montes-Hernández et al. (2017), CV values superior to 20% indicate a high morphological variation within vegetal species. During avocado samplings, fruit collections were not carried out, for this reason; other morphological traits were used to classify criollo trees within the three botanical races. They were: stem form, leaves length and width and trunk surface. According to 44, the Mexican race exhibits small leaves, the Guatemalan race medium size leaves and the West Indian race large leaves. Leaves length mean was 18. 58 cm for criollo avocados and ranged from 9.80 cm to 30.80 cm. Leaves' width mean was 8.78 cm and ranged from 4.70 cm to 14. 50 cm. The trunk surface is usually smooth in both Mexican and Guatemalan avocado races, whereas in the West Indian race is rough [42]. Results on the trunk surface showed that 83.5% of criollo trees were rough and the rest were smooth. Finally, another trait that is relevant to distinguish the three races is elevation ranges as Mexican and Guatemalan races are usually found in high altitudes whereas West Indian races are found in low altitudes [7]. In this study, criollo avocados were distributed from 864 to 2.281 m. a. s. l. suggesting they are within Mexican and Guatemalan races ranges of distribution. The Mexican race is mainly sowed from 1.000 to 2.000 m. a. s. l. And the Guatemalan race above 2.000 m. a. s. l [43]. Furthermore, results from PCA analysis demonstrate that the elevation is the most relevant landscape variable discovered here as it plays an important role in the clustering pattern obtained for 90 criollo trees.

In sum, results obtained here showed that criollo avocados are highly diverse in genetics and morphology. Morphological traits failed to differentiate criollo trees within the 3 botanical races. Population genetic structure was found with Antioquia and also that elevation differences between sampling sites played an important role in the genetic differentiation of the criollo trees studied in this work. This result was also corroborated by PCA analysis. STRUCTURE HARVESTER showed K = 2. Criollo samples were genetically similar to Guatemalan and Mexican races. Further analyses are necessary for the species related to its genetic characterization, particularly next-generation sequencing studies.

5. Conclusions

This work has shown the importance of the genetic analysis of crops of economical importance based on diverse molecular markers where microsatellites are

relevant to show the genetic variability of the species and current gene flow of a species while DNA sequencing generates information related to the genetic history of the species though comparisons of DNA sequences hold in databases vs. DNA sequences obtained in an investigation. Landscape analysis also requires the use of morphological traits as they can be used to explain the genotypic and phenotypic interactions in a species. According to the results obtained in this study, we found that Colombian avocado named "criollo" is the product of multiple hybridizations between natural trees enhanced by pollinators according to the results obtained from microsatellites and the high genetic diversity found in our avocado. Also, morphological variation showed that the species in Colombia is diverse and a product hybridization given the intermixed traits found on 90 criollo trees. Finally, DNA sequencing of three nuclear genes showed that Colombian avocados are genetically closer to the Mexican and the Guatemalan races of avocado and for this reason grafting between criollo trees and Hass variety is possible since both avocados are produced within the same altitudes.

Acknowledgements

This project was financially supported by Universidad Nacional de Colombia, sede Medellín to Clara Inés Saldamando Benjumea. Code No 29424. Collection and genetic access were provided for this study by ANLA (Autoridad Nacional de Licencias Ambientales) to UNAL (Universidad Nacional de Colombia) No 0255 given on March 14, 2014.

Authors contributions

GPC made the collections and molecular analysis. JM and CISB analyzed the data. RAI trained GPC, GPC and CISB wrote the manuscript.

DOI: http://dx.doi.org/10.5772/intechopen.101194

Author details

Clara Inés Saldamando-Benjumea[1,2*], Gloria Patricia Cañas-Gutiérrez[3], Jorge Muñoz[2] and Rafael Arango Isaza[1,2]

1 Escuela de Biociencias, Facultad de Ciencias, Universidad Nacional de Colombia, Medellín, Colombia

2 Plant Biotechnology Unit, Corporation for Biological Research, Medellín, Colombia

3 Plant Health and Biological Control Unit, Corporation for Biological Research, Medellín, Colombia

*Address all correspondence to: cisaldam@unal.edu.co

References

[1] Ashworth VETM, Clegg MT. Microsatellite markers in avocado (*Persea americana* Mill.): genealogical relationships among cultivated avocado genotypes. The Journal of Heredity. 2003;**94**(5):407-415. DOI: 10.1093/jhered/esg076

[2] Galindo-Tovar ME, Arzate -Fernández. Consideraciones sobre el origen y primeras dispersiones del aguacate (Persea americana, Lauraceae). Cuadernos de Biodiversidad. 2010;**33**:11-15

[3] Corona-Jacome EC, Galindo-Tovar ME, Lee-Espinosa HE, Landero-Torres I. Diversidad genética del aguacate (*Persea americana* Mill.) en cuatro zonas de su área de dispersión natural. Agro Productividad. 2016;**6**:80-85

[4] Bergh B, Ellstrand N. Taxonomy of the avocado. Calif Avocado Society Year Book. 1986;**70**:135-145

[5] Davenport TL. Avocado flowering. Horticultural Reviews. 1986;**8**:257-289

[6] Sharma CB, Martinez GC. The component triacylglycerols of avocado fruit-coat. Journal of American Oil Chemistry Society. 1972;**49**:229-232

[7] Chen H, Morrell PL, De la Cruz M, Clegg MT. Nucleotide diversity and linkage disequilibrium in wild avocado (Persea americana Mill.). The Journal of Heredity. 2008;**99**:382-389. DOI: 10.1093/jhered/esn016

[8] Bergh BO, Lahav E. Avocados. In: Janick J, Moore JN, editors. Fruit Breeding. Vol 1. Tree and Tropical Fruits. New York: Wiley; 1996. pp. 113-166

[9] Chanderbali AS, Albert VA, Ashworth VTEM, Clegg MT, Litz RE, Soltis DE, et al. *Persea americana* (avocado): bringing ancient flowers to fruit in the genomics era. Bioassays. 2008;**30**:386-396. DOI: 10.1002/bies.20721

[10] Galindo-Tovar M, Milagro P. Genetic relationships within avocado (*Persea americana* Mill.) in seven municipalities of central Veracruz characterized using microsatellites markers. Tropical and Subtropical Agroecosystems. 2011;**13**(3):339

[11] Schaffer B, Gil PM, Mickelbart MV, Whiley AW. In: Schaffer B et al., editors. Ecophysiology The Avocado: Botany. Production and Uses. 2nd ed. CAB International; 2013. pp. 168-199

[12] Cañas-Gutiérrez GP, Galindo-López LF, Arango- Isaza RE, Saldamando-Benjumea CI. Diversidad genética de aguacate (*Persea americana*) en Antioquia. Colombia. Agronomía Mesoamericana. 2015;**26**(1):129-143. DOI: 10.15517/am.v26i1.16936

[13] Mijares P, López L. Variedades de aguacate y su producción en México. México: Memoria Fundación Salvador Sánchez Colín CICTAMEX S.C; 1998

[14] Alcaraz M, Hormaza I. Molecular characterization and genetic diversity in an avocado collection of cultivars and local Spanish genotypes using SSRs. Hereditas. 2007;**144**:244-253. DOI: 10.1111/j.2007.0018-0661.02019x

[15] Gross-German E, Viruel MA. 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 & Genomes. 2013;**9**:539-555. DOI: 10.1007/s11295-012-0577-5

[16] Corona-Jacome EC, Galindo-Tovar ME, Lee-Espinosa HE, Landero-Torres I. Diversidad genética del

aguacate (Persea americana Mill.) en cuatro zonas de su área de dispersión natural. Agro Productividad. 2016;**6**:80-85

[17] Freeland J. Molecular Ecology. Jonh Wiley & Sons: West Sussex; 2005, p. 388

[18] Chen H, Morrell PL, De la Cruz M, Clegg MT. Tracing the geographic origins of major avocado cultivars. The Journal of Heredity. 2009;**100**:56-65. DOI: 10.1093/jhered/esn068

[19] Nei M, Kummar S. Molecular Evolution and Phylogenetics. Oxford: Oxford University Press; 2000. p. 333

[20] Manel S, Schwartz MK, Luikart G, Taberlet P. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology & Evolution. 2003;**18**:189-197. DOI: 10.1016/S0169-5347(03)00008-9

[21] IPGRI (International Plant Genetic Resources Institute). Descriptors for Avocado (Persea spp.). Rome, Italy: International Plant Genetic Resources Institute; 1995. p. 58

[22] Peakall R, Smouse PE. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes. 2006;**6**:288-295. DOI: 10.1111/j.1471-8286.2005.01155.x

[23] Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity. 1995;**86**:248-249. DOI: 10.1093/oxfordjournals.jhered.a111573

[24] Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;**155**:945-959

[25] Librado P, Rozas J. dnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;**25**:1451-1452. DOI: 10.1093/bioinformatics/btp187

[26] Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Molecular Ecology. 2005;**14**:2611-2620. DOI: 10.1111/j.1365-294X.2005.02553.x

[27] Earl DA, von Holdt BM. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources. 2012;**4**:359-361. DOI: 10.1007/s12686-011-9548-7

[28] Jakobsson M, Rosenberg NA. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007;**23**:1801-1806. DOI: 10.1093/bioinformatics/btm233

[29] Chen Y, Hou Y, Guo Z, Wang W, Zhong C, Zhou R, et al. Applications of multiple nuclear genes to the molecular phylogeny, population genetics and hybrid identification in the mangrove genus Rhizophora. PLoS One. 2015;**10**(12):e0145058. DOI: 10.1371/journal.pone.0145058

[30] Clark AG. Inference of haplotypes from PCR-amplified samples of diploid populations. Molecular Biology and Evolution. 1990;**7**:111-122. DOI: 10.1093/oxfordjournals.molbev.a040591

[31] Nei M, Kummar S. Molecular Evolution and Phylogenetics. Oxford: Oxford University Press; 2000. p. 333

[32] Posada D. jModeltest: Phylogenetic model averaging. Molecular Biology and Evolution. 2008;**25**:1253-1256. DOI: 10.1093/molbev/msn083

[33] Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, et al. BEAST 2: A software platform for bayesian evolutionary analysis. PLoS

Computational Biology. 2014;**10**: e1003537. DOI: 10.1371/journal.pcbi.1003537

[34] Rambaut A. FigTree v1.4.2. [WWW Document]. 2014. Available from: http://tree.bio.ed.ac.uk/software/figtree/ [Accessed: June 9, 2019]

[35] Hammer OD, Harper AT, Ryan PD. PAST: Palaeontological statistics software package for education and data analysis. Palaeontologia Electronica. 2001;**4**:1-9 http://palaeo-electronica.org/2001_1/past/issue1_01.htm

[36] OhsawaT TY, Saito Y, SawadaY IHY. Altitudinal genetic diversity and differentiation of Quercus crispula in the Chichibu mountains, Central Japan. International Journal of Plant Sciences. 2007;**168**:333-340

[37] Garcia A. Cytogenetic studies in the genus Persea (Lauraceae). I. Karyology of seven species. Canadian Journal of Genetics and Cytology. 1975;**17**:173-180

[38] Ibarra-laclette E, Méndez-Bravo A, Pérez-Torres CA, Albert VA, Mockaitis K, López-Gómez R, et al. Deep sequencing of the Mexican avocado transcriptome, an ancient angiosperm with a high content of fatty acids. BMC Genomics. 2015;**16**:599. DOI: 10.1186/s12864-015-1775-y

[39] Boza EJ, Tondo CL, Ledesma N, Campbell RJ, Bost J, Schnell RJA, et al. Genetic differentiation, races and interracial admixture in avocado (*Persea americana* Mill.), and Persea spp. evaluated using SSR markers. Genetic Resources and Crop Evolution. 2018;**65**: 195-1215

[40] Wei SJ, Lu YB, Ye QQ, Tang SQ. Population genetic structure and phylogeography of camellia flavida (Theaceae) based on chloroplast and nuclear DNA sequences. Frontiers in Plant Science. 2017;**8**:718. DOI: 10.3389/fpls.2017.00718

[41] Montes-Hernández SJ, de la Torre-Vizcaino D, Heredia-García E, Hernández-Martínez M, Camarena-Hernández MG. Caracterización morfológica de germoplasma de aguacate MEXICANO (Persea americana var. drymifolia. LAURACEAE). Interciencia. 2017;**42**(3):175-180

[42] Barrientos-Priego AF, López-López L. Historia y genética del aguacate. In: Téliz D, González H, Rodríguez J, Dromundo R, editors. El aguacate y su manejo integrado. México: Mundi-Prensa; 2000. pp. 19-31

[43] Barrientos PAF, López LL. Historia y genética del aguacate. In: Memoria 2002 de la Fundación Salvador Sánchez Colín. Coatepec de Harinas, México: Centro de Investigaciones Científicas y Tecnológicas del Aguacate en el Estado de México; 2002. pp. 100-121