

Genotype selection of *Pouteria sapota* (Jacq.) H.E. Moore & Stearn, under a multivariate framework

Selección del genotipo de *Pouteria sapota* (Jacq.) H.E. Moore & Stearn, bajo un enfoque multivariado

Renan Mercuri Pinto ^{1*}, Luiz Ricardo Nakamura ¹, Thiago Gentil Ramires¹, Ezequiel Abraham López Bautista ² and Carlos Tadeu dos Santos Dias¹

¹Universidade de São Paulo – USP. Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ, Piracicaba, São Paulo, Brazil.

²Universidad de San Carlos de Guatemala – USAC, Facultad de Agronomía, Guatemala. *Autor para correspondencia: renanmp@usp.br

Rec.: 28.03.2015 Acep.: 08.09.2015

Abstract

The *Pouteria sapota*, also popularly known as sapote or mamey sapote, is a fruit tree of sapotaceae family originally from the tropical region of Central America with a great importance due to the almost complete utilization of the tree (fruit, seeds and wood) by industries. Thus, the study of its features becomes indispensable for selecting the most promising genotypes to increase the profitability of its production. In this study, it was used a dataset of 63 sapote trees placed in the botanical garden of Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), located in Turrialba, Costa Rica. 17 quantitative characteristics were measured from trees, in order to evaluate the yield potential through the application of two multivariate statistical techniques: factor analysis (FA) and cluster analysis (CA). Firstly, the FA was performed and the 17 initial characteristics were reduced to four common factors that might describe particular characteristics like “fruit”, “seed”, “wood” and “leaf”. Thereafter, a CA was performed, with scores of FA, which allows the formation of five groups of trees with different traits. This methodology revealed the most promising trees in the economic point of view for every industry that uses the tree as raw material.

Keywords: Plant genetic resources, cluster analysis, factor analysis, multivariate analysis

Resumen

La especie *Pouteria sapota*, también conocida como zapote o zapote mamey es un árbol frutal que pertenece a la familia sapotaceae, originaria de la región tropical de América Central, de gran importancia debido al amplio uso del árbol (fruta, semilla y madera) para la industria. Debido a esto, el estudio de sus características se hace indispensable para seleccionar genotipos promisorios para incrementar la rentabilidad de su producción. En este estudio fue utilizada una base de datos consistente en registros de 63 árboles de zapote, plantados en el jardín botánico del Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), localizado en Turrialba, Costa Rica. 17 características cuantitativas fueron medidas a los árboles, con el objetivo de evaluar el rendimiento potencial por medio de la aplicación de dos técnicas estadísticas multivariadas: análisis factorial (AF) y análisis de conglomerados (AC). Inicialmente fue realizado el AF y 17 características iniciales fueron reducidas a cuatro factores comunes que pueden describir las características particulares como: “fruta”, “semilla”, “árbol” and “hoja”. Posteriormente fue aplicado el AC con los escores del AF, permitiendo la formación de cinco grupos de árboles con diferentes características. Esta metodología mostró ser prometedora para la selección de árboles con mejores características, con fines económicos para la industria.

Palabras clave: Análisis de conglomerados, análisis factorial, análisis multivariado, recursos fitogenéticos.

Introduction

The sapote [*Pouteria sapota* (Jacq.) H.E. Moore & Stearn], is a fruit tree of the sapotaceae family native to Central America, and wild populations of these trees are found from southern Mexico to Costa Rica and possibly in Northern South America. Besides being a fruitful open-pollinated species, it has an important economic effect due to the almost complete utilization of the tree (fruit, seeds and wood) by industries. For instance, fruits can be consumed in their natural state and its pulp can be used in the manufacture of jellies, ice cream and juice. Solís-Fuentes *et al.* (2015) reported a study showing that seeds can be used as constituents in mixtures with other natural fats, such as cocoa butter or mango seed fat. Seeds are also used as fuel and their oil can be used by cosmetic companies and also as a medicine to prevent baldness or reduce muscle pain (Alia-Tejacal *et al.*, 2007). Dermal infections are also controlled by the application of wood latex (Mutchnick & McCarthy, 1997).

This species grows in heavy clay soil (Puerto Rico), sandy clays (Guatemala) and even in sandy soils (Florida, USA), reaching 20 to 25m in average. The leaves are ovate or lanceolate and are concentrated at the top of the branches. The flowers are small and grow along leafless branches, each of them consists of five true and five false stamens, the pistil has only one stigma and the ovary has five carpels. The fruit contains one or more seeds and has a variety of forms that can weigh about 3 kg in some genotypes. The flesh varies in texture (red, orange or grey) and it is aromatic, sweet and soft when ripe (Bermejo *et al.*, 1994).

Due to the great utility of products taken from the sapote tree, several studies have emerged in the literature in different areas. Recently, Torres-Rodríguez *et al.* (2011), presented the antioxidant effects to promote health, based on phenolic composition of sapote fruit, which showed no significant changes with the fruit ripening. Haro *et al.* (2015), showed great industrial potential to use sapote pulp for soda production since it has a high concentration of sugar, high yield and sustainable with low contamination. Moo-Huchin *et al.* (2013) reported that the chemical composition of crude oil from the sapote seed might be an important source of vegetable oil for industrial uses.

Studies on genetic diversity include the characterization of germplasm collections, an activity that consists in registering the qualitative and quantitative characteristics that are inheritable. As Querol (1988), the characterization is the use of data to describe and thus differentiate these improvements of a given specie. The systematic

characterization allows knowing the variation within the collections and selects the best genotypes to grow them. An important statement is that the covariance of those characteristics must be studied in a multivariate approach as the identification of the best genotypes comes along by the consideration of all important variables in study simultaneously (Pinto *et al.*, 2015). Another interesting point is that certain multivariate techniques, such as the factor analysis, reduce the dimensionality of a given problem, simplifying it. Some applications using different types of crops with multivariate techniques are already available in many works, such as in Nakamura *et al.* (2013) who study coffee genotypes, Righetto *et al.* (2014), that present a multivariate application with cacao, Vargas *et al.* (2015), who perform a study with tomato genotypes, among others.

Hence, the main propose of this study was to provide an efficient technique to select most promising genotypes of *Pouteria sapota*, classifying them into groups according to their particular genetic traits, i.e., indicate to industries that use sapote as raw material, the best genotypes according to their activities combining two different multivariate techniques: factor analysis (FA) and cluster analysis (CA). These important techniques will be explained with more details in the next section.

Material and methods

The dataset in study is located in the Cabiria 6 Botanic Garden of the Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), based in Costa Rica (north latitude 9° 53', west longitude 83° 39' and 602 meters above sea level).

CATIE sapotaceae collection plants were introduced to the CATIE genbank between 1977 and 1983, from seeds collected from Mexico to Panama, planted at a distance of 8 x 6.5 m. During November 1994 to January 1995, the collections were visited, at the same time it was noted the condition of each tree. If a plant was abundant fruits or flowers in large quantity, no matter the size, it was tagged and was part of the study. The point of harvest for the characterization of the fruits was determined when these were orange inside. After harvested the fruits wrapped in newspaper, and were stored in the laboratory, until they reached the appropriate point of maturity to be evaluated.

The following variables were measured: FW: Fruit weight (g); FL: Fruit length (mm); FD: Fruit diameter (mm); PuT: Pulp thickness (mm); PeT: Peel thickness (mm); PW: Peel weight (g); PP: Pulp performance (%); NS: Number of seeds; SL: Seed length (mm); SD: Seed diameter (mm); SW: Seed weight (g); LL: Leaf length (cm); LW: Leaf width

(cm); TH: Tree height (m); TTD: Tree trunk diameter (cm); e, TD: Treetop diameter (m).

Due to the high number of variables, as in Righetto *et al.* (2014), we applied the factor analysis (FA), a multivariate statistical technique whose main objective is to describe the covariance relationships among the dataset variables in terms of a few unobservable variables, which are called common factors (Johnson and Wichern, 2007). The orthogonal factor model is given by

$$X_{px1} - \mu_{px1} = \Lambda_{pxm} F_{mx1} + \epsilon_{mx1}$$

Where: X is a vector of random variables, with mean μ and covariance matrix Σ ; Λ is the loading matrix, in which the higher the loading is, higher is the relationship between a given variable with a given factor. F is the factor vector; and, ϵ is the vector of random errors.

In order to reduce the dimensionality of the variables, we used the following criterion: we retain the first m components account for a percentage of total variation greater than 70%. This threshold is advised by Mardia *et al.* (1992), and respected by many authors as, e.g., Surita *et al.* (2007), Khayatnezhad *et al.* (2011) and Mollasadeghi *et al.* (2011). The main objective here is that all retained factors have practical meanings and, once they are obtained, we can name them following it. It is not uncommon to find the necessity of a rotation on the original axis (factors) in order to find a proper interpretation of the factors and, thus, a rotation method should be applied in this cases.

Once factors and scores were obtained, we applied the cluster analysis (CA) using the generalized euclidean distance, given by

$$d_{ij} = \left\{ \sum_{k=1}^n (x_{ik} - x_{jk})^2 \right\}^{0.5}$$

and Ward's minimum variance method, in order to obtain homogeneous groups of genotypes and so point those with the best agronomic and economic characteristics. This method was chosen since it uses a huge statistical plea based in the analysis of variance (Nakamura *et al.*, 2013) to group similar objects: firstly we calculate the proximity matrix, given by $P=0.5D$, where D is the distance matrix; then we just need to calculate the proximity value of a cluster RS with any group T by:

$$p_{(rs)t} = \frac{1}{(n_t + n_{rs})} [(n_t + n_r)p_{rt} + (n_t + n_s)p_{st} - n_t p_{(rs)}]$$

Where: n_t, n_s, n_r and $n_{(rs)}$ corresponds to the number of individuals from groups T, R, S and RS, respectively, and p_{rt} and $p_{(rs)}$ are the proximity

values of group R with T, group S with T and group R with S, respectively. After all clusters are formed, we can display a dendrogram (Johnson and Wichern, 2007). Examples of applications using this method can be obtained in Badu-Apraku *et al.* (2011), Kholghi *et al.* (2011), Lajús *et al.* (2013), among others.

Results and discussion

Table 1, shows some descriptive statistics: mean, standard deviation, coefficient of variation (CV %), skewness and kurtosis, of each variable in the dataset. As we can see, none of the variables present a high standard deviation when compared with its mean. In addition, only the variable number of seeds presents high skewness (1.30), i.e., there is an accumulation of observations in the right of its distribution, and only the variable leaf width presents high kurtosis (3.08), indicating a high number of observations around its mode. Although these values indicate a non-Gaussian distribution that is not an issue for some statistical analysis, like FA (Johnson & Wichern, 2007) since the Gaussian assumption does not interfere in the analysis process.

Table 1. Descriptive analysis of the dataset in study

Variable	Mean	Standard deviation	CV(%)	Skewness	Kurtosis
FW	380.81	128.25	33.68	0.46	-0.42
FL	98.38	16.04	16.31	0.19	-0.08
FD	83.36	10.16	12.19	0.14	-0.83
PuT	27.00	4.76	17.63	0.27	-0.93
PeT	1.93	0.41	21.27	0.01	-0.38
PW	52.38	16.74	31.96	0.51	-0.11
PP	74.62	4.88	6.53	-0.41	-0.11
NS	1.31	0.31	23.88	1.30	1.42
SL	62.22	7.18	11.55	0.19	-0.32
SD	33.17	3.08	9.28	-0.55	1.55
SW	40.28	9.28	23.03	0.21	-0.51
LL	25.32	7.05	27.83	0.63	0.39
LW	8.30	1.99	23.98	1.02	3.08
TH	6.82	1.96	28.73	0.62	-0.44
TTD	25.41	4.70	18.50	0.29	1.41
TD	6.28	1.46	23.19	0.18	0.16

FW: Fruit weight (g); FL: Fruit length (mm); FD: Fruit diameter (mm); PuT: Pulp thickness (mm); PeT: Peel thickness (mm); PW: Peel weight (g); PP: Pulp performance (%); NS: Number of seeds; SL: Seed length (mm); SD: Seed diameter (mm); SW: Seed weight (g); LL: Leaf length (cm); LW: Leaf width (cm); TH: Tree height (m); TTD: Tree trunk diameter (cm); e, TD: Treetop diameter (m) / CV(%): Coefficient of Variation

Once we obtained the descriptive statistics, we calculated the Kaiser-Meyer-Olkin (KMO) rate, which is a measure that evaluates if the FA is reliable in the dataset in study. In our dataset, the KMO value was 0.70, thus we performed the FA.

The FA was performed over the correlation matrix and we could observe that the cumulative percentage of variation explained using only four common factors was 72.33% (eigenvalues: 6.12, 2.50, 1.77 and 1.17, for first, second, third and fourth factors, respectively) and then this *m* number of factors was chosen using the threshold proposed by Mardia *et al.* (1992). Table 2, displays the rotated loading matrix using the Varimax criterion.

We can see from Table 2 that variables in study were well distributed over the four factors retained. Hence, highlighted loadings contributed to naming each of the factors: Factor 1: most important variables are FW, FD, PuT and PP, and thus we named this factor as “Fruit”; Factor 2: most important variables are FL, PW, SL, SD and SW, and thus we named this factor as “Seed”; Factor 3: most important variables are PeT, LL, TH, TTD and TD, and thus we named this factor as “Wood”; and Factor 4: most important variables are NS, LL and LW, and thus we named this factor as “Leaf”.

As the scores were obtained, we performed a CA through the Ward’s minimum variance method (Figure 1). According to Figure 1, five different clusters composed of similar genotypes were obtained, and their principal characteristics are: Group 1: large fruits – Trees: 1, 3, 6, 7, 8, 13, 20, 22, 24, 37, 46, 47, 54, 55, 57 and 62; Group 2: small fruits and large trees – Trees: 10, 25, 26, 34 and 35; Group 3: small leaf – Trees: 2, 9, 12, 14, 15, 16, 17, 19, 29, 30, 39, 40, 42, 44, 50, 52, 53, 56, 59, 60 and 61; Group 4: small trees and

Table 2. Varimax rotated factor matrix for variables in study

Variable	Factor 1	Factor 2	Factor 3	Factor 4
FW	0.81	0.52	0.20	0.01
FL	0.35	0.72	0.30	0.04
FD	0.88	0.26	0.16	0.01
PuT	0.82	0.19	0.28	-0.22
PeT	-0.07	0.45	0.54	-0.04
PW	0.50	0.67	0.36	-0.10
PP	0.92	-0.09	0.04	-0.05
NS	-0.31	-0.05	-0.38	0.66
SL	0.11	0.75	0.31	0.12
SD	0.06	0.68	-0.05	-0.06
SW	0.06	0.78	-0.09	0.48
LL	-0.03	0.08	0.67	0.56
LW	0.01	0.13	0.21	0.84
TH	0.31	0.03	0.59	-0.06
TTD	0.16	0.14	0.76	0.01
TD	0.20	0.10	0.80	0.04

FW: Fruit weight; FL: Fruit length; FD: Fruit diameter; PuT: Pulp thickness; PeT: Peel thickness; PW: Peel weight; PP: Pulp performance; NS: Number of seeds; SL: Seed length; SD: Seed diameter; SW: Seed weight; LL: Leaf length; LW: Leaf width; TH: Tree height; TTD: Tree trunk diameter; e, TD: Treetop diameter

large leaves – Trees: 4, 5, 23, 41, 43, 48, 58 and 63; and Group 5: large seeds – Trees: 11, 18, 21, 27, 28, 31, 32, 33, 36, 38, 45, 49 and 51.

Based on groups formed through the cluster analysis, whose averages and standard deviations of the variables in the respective groups are displayed in Table 3, it is possible to indi-

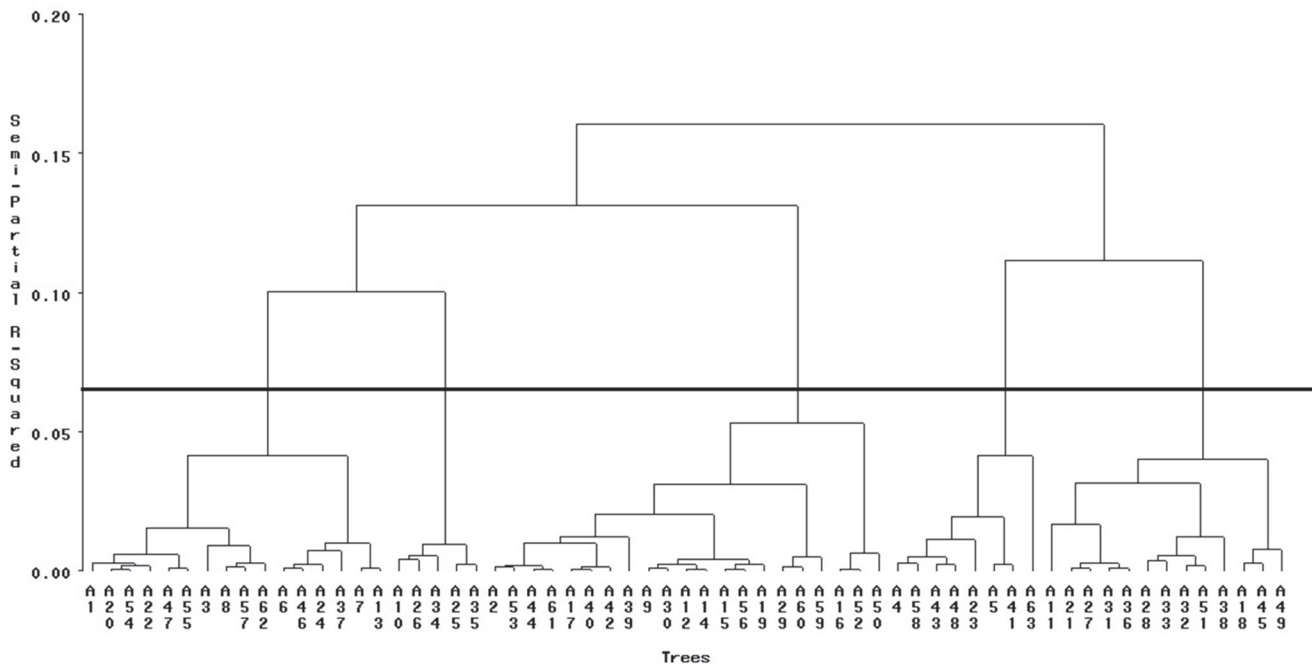


Figure 1. Dendrogram obtained from the Ward’s hierarchical method using scores from the previous applied factor analysis.

cate promising groups of genotypes according to the necessity of different types of industries since the applied methodology was able to group homogeneous genotypes in each cluster (standard deviations are not numerically high when compared to its means in Table 3). For instance, we can say that food industries should consider genotypes from Groups 1 and/or 5, since those trees produced, respectively, large fruits and large seeds when compared to the other groups and hence should be used on the manufacture of jellies, ice cream, juice and chocolate. Moreover, cosmetic companies should also choose Group 5, since the seed oil is the raw material in this type of industry. Finally, Group 2 may be the choice of carpentry companies, since that tree in this group were the largest ones.

Table 3. Mean and standard deviation (in parentheses) of each variable under study for each of the five groups identified

Variable	Group 1	Group 2	Group 3	Group 4	Group 5
FW	491,07 (87,96)	338,66 (39,91)	298,29 (94,70)	272,20 (73,27)	461,45 (113,69)
FL	103,12 (11,17)	104,48 (10,01)	88,36 (12,89)	85,05 (10,36)	114,57 (13,32)
FD	94,17 (5,68)	76,72 (3,56)	78,76 (8,95)	76,13 (5,60)	84,48 (9,49)
PuT	31,54 (3,11)	24,36 (1,36)	26,17 (4,36)	21,41 (2,27)	27,23 (4,15)
PeT	1,98 (0,52)	2,42 (0,13)	1,83 (0,36)	1,80 (0,19)	1,95 (0,40)
PW	62,91 (15,05)	56,70 (6,36)	42,96 (12,45)	36,55 (8,36)	62,72 (16,08)
PP	79,29 (2,64)	70,70 (1,95)	73,35 (5,13)	70,60 (3,39)	74,92 (3,63)
NS	1,17 (0,14)	1,28 (0,11)	1,16 (0,18)	1,93 (0,24)	1,35 (0,26)
SL	62,04 (4,51)	67,86 (7,19)	58,26 (4,70)	56,60 (6,02)	70,12 (6,04)
SD	33,13 (3,53)	33,16 (1,16)	32,47 (2,57)	31,26 (3,31)	35,52 (2,43)
SW	39,40 (6,53)	42,38 (7,64)	33,10 (6,09)	42,48 (10,42)	50,82 (5,50)
LL	28,03 (6,50)	35,04 (4,80)	21,01 (4,79)	24,76 (3,97)	25,57 (8,18)
LW	8,82 (1,17)	10,16 (1,23)	6,95 (1,44)	9,85 (3,02)	8,15 (1,69)
TH	8,31 (1,95)	7,42 (1,98)	6,45 (1,87)	5,25 (1,15)	6,31 (1,38)
TTD	27,66 (4,69)	31,90 (3,64)	24,22 (2,90)	21,65 (4,98)	24,40 (4,08)
TD	7,31 (1,37)	8,16 (1,37)	5,92 (0,82)	4,80 (1,09)	5,80 (1,17)

FW: Fruit weight; FL: Fruit length; FD: Fruit diameter; PuT: Pulp thickness; PeT: Peel thickness; PW: Peel weight; PP: Pulp performance; NS: Number of seeds; SL: Seed length; SD: Seed diameter; SW: Seed weight; LL: Leaf length; LW: Leaf width; TH: Tree height; TTD: Tree trunk diameter; e, TD: Treetop diameter

It is also noteworthy that, although this methodology was applied in the selection of promising genotypes of sapote trees, it can also be applied on other crops. Therefore, it is a powerful biometric tool for agroindustry that use any fruit as a raw material.

Conclusions

The factor analysis was a powerful tool used in this work, since it performed really well the reduction of the original number of variables in the dataset (from 17 quantitative characteristics to only four common factors) since it did not lose a lot of the original information (72.33% of the original variance was explained by this few factors).

The cluster analysis returned five homogeneous clusters and thus, it was possible to decide which genotypes are better for each kind of industry that uses sapote as a raw material.

Acknowledgements

The authors gratefully acknowledge grant from CAPES and CNPq (Brazil).

References

- Alia-Tejaca. I. Villanueva-Arce. R. Pelayo-Zaldivar. C. Colinas-León. M.T. López-Martínez. V. & Bautista-Baños. S. (2007). Postharvest physiology and technology of sapote mamey fruit (*Pouteria sapota* (Jacq.) HE Moore & Stearn). *Postharvest Biol Tec*, 45, 285-297. doi:10.1016/j.postharvbio.2006.12.024.
- Badu-Apraku. B. Oyekunle. M. Akinwale. R.O. Fontem-Lum. A. (2011). Combining ability of early-maturing white maize inbreds under stress and nonstress environments. *Agron J*, 103, 544-557. doi: 10.2134/agronj2010.0345.
- Bernejo. J.E.H. León. J. (1994). Neglected crops: 1492 from a different perspective. No. 26. Food & Agriculture Org. 341 p.
- Haro. I.R. Cotrina. G.S. Plasencia. P.T. Castillo. M.S. (2015). Potencial industrial de la pulpa de *Pouteria sapota* para la preparación de néctar de calidad. *Rebiol*, 34(2), 5-12.
- Johnson. R.A. Wichern. D.W. (2007). Applied multivariate statistical analysis. 6th ed. New Jersey: Upper Saddle River. 773 p.
- Khayatnezhad. M. Zaefizadeh. M. Gholamain. R. (2011). Factor analysis of yield and other traits of durum wheat under drought stress and no stress conditions. *Plant Ecophy*, 3(1), 23-27.
- Kholghi. M. Bernousi. I. Darvishzadeh. R. Pirzad. A. Maleki. H.H. (2011). Collection, evaluation and classification of Iranian confectionery sunflower (*Helianthus annuus* L.) populations using multivariate statistical techniques. *Afr J Biot*, 10, 5444-5451. doi: 10.5897/AJB10.2146.
- Lajús. C.R. Miranda. M. Scheffer-Basso. S.M. Carneiro. C.M. Escosteguy. P.A.V. (2013). Leaf tissues proportion and chemical composition of *Axonopus*

- jesuiticus* x *A. scoparius* as a function of pig slurry application. *Ciência Rural*, 44(1), 276-282. doi: 10.1590/S0103-84782013005000154.
- Mardia. K.V. Kent. J.T. Bibby. J.M. (1992). Multivariate analysis. London: Academic Press. 518 p.
- Mollasadeghi. V. Shahryari. R. Imani. A.A. Khayat-nezhad. M. (2011). Factor analysis of wheat quantitative traits on yield under terminal drought. *Am-Euras J Agric & Environ Sci*, 10(2), 157-159.
- Moo-Huchin. V. Estrada-Mota. I. Estrada-León. R. Cuevas-Glory. L.F. Sauri-Duch. E. (2013). Chemical composition of crude oil from the seeds of pumpkin (*Cucurbita spp.*) and mamey sapota (*Pouteria sapota* Jacq.) grown in Yucatan, Mexico. *J Food Sci*, 11(4), 324-327. doi:10.1080/19476337.2012.761652.
- Mutchnick. P.A. McCarthy. B.C. (1997). An ethnobotanical analysis of the tree species common to the subtropical moist forests of the Petén, Guatemala. *Economic Botany*, 51(2), 158-183. doi: 10.1007/BF02893110.
- Nakamura. L.R. Bautista. E.A.L. Quaresma. E.S. Dias. C.T.S. Miranda. E.F.O. (2013). Seleção de genótipos promissores de café: uma abordagem multivariada. *Rev Bras Biom*, 31(4), 516-528.
- Pinto. R.M. Campos. D.H.S. Tomasi. L.C. Cicogna. A.C. Okoshi. K. Padovani. C.R. (2015). Multivariate Analysis for Animal Selection in Experimental Research. *Arq Bras Card*, 104 (2), 97-103. doi: 10.5935/abc.20140219.
- Querol. D. (1988). Recursos genéticos nuestro tesoro olvidado. Aproximación técnica y socioeconómica. Industrial Gráfica S.A., Lima, Perú. 218 p.
- Righetto. A.J. Nakamura. L.R. Bautista. E.A.L. Dias. C.T.S. (2014). Aplicación de técnicas estadísticas multivariadas para el agrupamiento de materiales genéticos de cacao (*Theobroma cacao* L.), *Tikalía*, 32(1), 47-62.
- Solis-Fuentes. J.A. Ayala-Tirado. R.C. Fernández-Suárez. A.D. & Durán-de-Bazúa. M.C. (2015). Mamey sapote seed oil (*Pouteria sapota*). Potential, composition, fractionation and thermal behavior. *Grasas y Aceites*, 66(1), 1-10. doi: 10.3989/gya.0691141.
- Surita. C.A. Gloaguen. T. Montes. C.R. Dias. C.T.S. (2007). Assessment of soil solution chemicals after tannery effluents disposal. *Am J Appl Sci*, 4(1), 1063-1070. doi: 10.3844/ajassp.2007.1063.1070.
- Torres-Rodríguez. A. Salinas-Moreno. Y. Valle-Guadarrama. S. Alia-Tejagal. I. (2011). Soluble phenols and antioxidant activity in mamey sapote (*Pouteria sapota*) fruits in postharvest. *Food Res Int*, 44(1), 1956-1961. doi: 10.1016/j.foodres.2011.04.045.
- Vargas. T.O. Alves. E.P. Abboud. A.C.S. Leal. M.A.A. Carmo. M.G.F. (2015). Diversidade genética em acessos de tomateiro heirloom. *Hort Bras*, 33(2), 174-18. doi: 10.1590/S0102-053620150000200007.